12.5 Limitations of Inductive Techniques

What do the methods explained in the preceding sections actually do for us? John Stuart Mill believed that they were instruments with which we may discover causal connections, and also that they were canons with which causal connections may be proved. On both counts he overestimated their power. Inductive techniques are indeed of very great importance, but their role in science is more limited than Mill supposed.

One substantial difficulty arises from the fact that, in formulating these methods, Mill made the assumption that one can identify cases “having only one circumstance in common” or other cases “having every circumstance in common save one.” But these expressions must not be taken literally; any two objects will have many circumstances in common however different they may appear; and no two things can ever differ in only one respect—one will be farther to the north, one will be closer to the sun, and so on. Nor could we even examine all possible circumstances to determine if they differ in only one way. What the scientist has in mind as he applies these techniques are not all circumstances, but the sets of relevant circumstances—whether there is only one relevant circumstance in common, or all relevant circumstances save one in common. That is, we apply the methods to the circumstances that have some bearing on the causal connection in question.

Which are those circumstances? We cannot learn which factors are relevant using the methods alone. In order to use the methods we must come to the context in which they are to be applied with some analysis of causal factors already in mind. The caricature of the “scientific drinker” illustrates this difficulty: He drinks Scotch and soda one night, bourbon and soda the next night, and on the following nights brandy and soda, then rum and soda, then gin and soda. What is the cause of his intoxication? Repeatedly inebriated, he swears never to touch soda again!

This scientific drinker did apply the method of agreement in accordance with the rules—but his doing so was to no avail because the factors that really are relevant in those antecedent circumstances had not been identified and therefore could not be manipulated. Had alcohol been specified as one of the factors common to all the cases, it would have been possible to eliminate soda very quickly, of course, using the method of difference.

The heroic investigation of the causes of yellow fever, discussed earlier in connection with the method of difference, confirmed the conclusion that the fever is spread by the bite of an infected mosquito. We know that now, just as we know now that it is alcohol and not soda that causes drunkenness. But the yellow fever experiments required insight and imagination as well as courage;
the notion that the fever was spread by mosquitoes was originally thought to be silly, or absurd, or was not thought of at all. Circumstances in the real world do not come wearing tags marked “relevant” or “irrelevant.” The testing of mosquito bites as cause required some earlier sorting of possibly relevant factors, to which the inductive methods might then be applied. With that prior analysis in hand, the methods can prove exceedingly helpful—but the methods by themselves, without some hypotheses in the background, are not sufficient instruments for scientific discovery.

Nor can the methods by themselves constitute rules for proof. Their application proceeds always on the basis of some antecedent hypotheses about causal factors, as noted just above, and because all circumstances cannot have been considered, attention will be confined to those believed to be the possible causes in question. However, this judgment regarding which circumstances are to be investigated may prove to have been in error. Medical scientists, for a very long time, did not consider dirty hands even as possible agents of infection, and so could not identify such dirtiness as the cause of disease.* Investigation is stymied when the investigators fail to break down the circumstances before them into the appropriate elements, elements that cannot be known in advance. Because the analyses presupposed by the application of the methods may be incorrect, or inadequate, the inferences based on those analyses may also prove to be mistaken. This dependence of induction on the merit of the underlying hypotheses shows that inductive techniques cannot by themselves provide the proof of causation that Mill had hoped for.

Yet another problem should be borne in mind: The application of inductive methods always depends on observed correlations, and even when the observations have been made accurately, they may be incomplete and therefore deceptive. The greater the number of observations, the greater is the likelihood that the correlation we observe is the manifestation of a genuine causal law—but no matter how great that number is, we cannot infer with certainty a causal connection among instances that have not yet been observed.

---

*The failure of physicians to wash their hands (because they did not understand how infectious diseases were spread) resulted in untold misery and uncounted deaths over centuries, especially from puerperal (or childbed) fever, which was carried on the hands of doctors from mother to mother, until the proof of that disastrous causal connection was given by the Hungarian physician Ignac Semmelweis, in the middle of the nineteenth century. See Sherwin B. Nuland, *The Doctors’ Plague* (New York: W. W. Norton, 2003).
These limitations illuminate once again the great gulf between deduction and induction. A valid deductive inference constitutes a proof, or demonstration, but every inductive inference is, at best, highly probable and never demonstrative. Therefore Mill’s claim that his canons are “methods of proof” must be rejected, along with the claim that they are “the methods of discovery.”

Nevertheless, the techniques explained in this chapter are central in much of science and are very powerful. Because it is impossible for investigators to take all circumstances into account, the application of the methods must always suppose one or more causal hypotheses about the circumstances under investigation. Being unsure which factor(s) are the cause(s) of the phenomenon under investigation, we often formulate alternative hypotheses and subject each to testing. What the five methods of induction, being mainly eliminative in nature, enable us to determine is this: If some specified analysis of the antecedent circumstances is correct, one of these factors cannot be (or must be) the cause (or part of the cause) of the phenomenon in question. This may be deduced, and the deduction may be valid, but the soundness of that argument will always depend on the correctness of the antecedent analysis that had been supposed.

The methods of induction are splendid, but they can yield reliable results only when the hypothesis that they seek to confirm (or falsify) does identify correctly the circumstances that are causally relevant. The methods permit the deduction of those results only when that hypothesis has been assumed as a premise in the argument. The nature of the power these methods give us may now be seen. They are not paths for discovery; they are not rules for proof. They are instruments for testing hypotheses. The statements of these inductive techniques, taken together, describe the general method of controlled experiment, which is a common and indispensable tool in all of modern science.

So important is the role of hypotheses in systematic empirical investigations that the enterprise of devising and testing hypotheses may be regarded as the method of science—to which we turn in the next chapter.

**EXERCISES**

Analyze each of the following investigations, or arguments, and indicate which of the methods of causal reasoning—Mill’s methods—are being used in each of them.

1. Teens who lose their virginity earlier than their peers are more likely to shoplift, destroy property, or sell drugs than their virgin counterparts, according to a recent national study of 7000 teenagers. Those who had sex early were 20 percent more likely to engage in delinquent acts one
year later compared to those whose first sexual experience occurred at the average age for their school. Those who waited longer than average to have sex had delinquency rates 50 percent lower a year later compared to average teens. Waiting appears to have a protective effect. “We’re not finding that sex itself leads to delinquency; sex itself is not always a problem behavior,” writes co-author and Ohio State sociologist Stacy Armour. However, “the timing of sexual initiation does matter. Kids go off on a different trajectory if they’re having sex early.”


2. Strong evidence has been presented that a diet low in folic acid [a trace vitamin in the B complex] during pregnancy increases the chances of giving birth to a premature baby of lower than normal birth weight. Dr. Theresa Scholl [of the University of Medicine and Dentistry of New Jersey] studied the outcomes of pregnancy for 832 women from the inner city of Camden, N.J., to determine the influence of dietary and supplementary consumption of folic acid. “We found that the women who consumed less than 240 micrograms per day of folic acid had about a two to threefold greater risk of preterm delivery and low birth weight,” she said. She reported that even small increases in the women’s serum folic acid concentrations by the 28th week decreased the odds of preterm delivery as well as the chance of having a baby of low birth weight. Of the 219 women in the low-folic-acid category (receiving less than 240 micrograms a day), 44 had preterm, low birth weight infants. “The risks declined in direct relationship to increased serum levels of folic acid, showing that low intake is a risk factor throughout pregnancy,” Dr. Scholl concluded.


3. The sequence of DNA units in the genome of humans and in that of chimpanzees is 98.8 percent identical; humans and chimps shared a joint ancestor as recently as five million years ago. Relatively few genes, therefore, must define the essence of humanity, and biologists have long supposed that if they could identify genes that have changed in the evolutionary advance leading from that joint ancestor, they would better understand the genetic basis of how people differ from chimpanzees, and hence what makes humans human.

This project received a significant boost in 2001 when a large London family with barely intelligible speech was found to have
mutations in a gene called FOXP2. Chimpanzees also have an FOXP2 gene, but theirs is significantly different from ours. The human version shows signs of accelerated evolutionary change in the last 100,000 years, which suggests that the gene acquired a new function that helped to make human speech possible.

—Reported by Dr. Michelle Cargill of Celera Diagnostics, Alameda, CA, and Dr. Andrew Clark, of Cornell, in Science, 11 December 2003

4. A simple, inexpensive and surprisingly powerful combination of treatments that all but wiped out malaria in a group of HIV-positive children in a recent study in Uganda was described at a very recent medical conference in Los Angeles. The combination—taking one inexpensive antibiotic pill each day and sleeping under an insecticide-treated mosquito net—reduced the incidence of malaria by 97 percent compared with a control group. The study, conducted by Dr. Anne Gasasira of Makerere University in Kampala, Uganda, found that among 561 healthy children who were not HIV-infected and who did not take the antibiotic and sleep under bed nets, there were 356 episodes of malaria. This compared with 4 episodes among 300 children who were known to be HIV-infected and received both treatments. “The findings were shockingly dramatic,” said Dr. Elaine Abrams, a professor of pediatrics and epidemiology at Columbia University.

—Reported at the 14th Conference on Retroviruses and Opportunistic Infections, Los Angeles, 28 February 2007

5. Some theories arise from anecdotal evidence that is difficult to confirm. In The Left-Hander Syndrome (New York: Bantam Books, 1992), Stanley Coren sought to evaluate the common belief that left-handed persons die sooner than right-handers. But death certificates or other public records very rarely mention the hand preferred by the deceased. What could serve as a reliable data source with which that hypothesis could be tested? Coren searched baseball records, noting which hand baseball pitchers threw with, and then recording their ages at death. Right-handed pitchers, he found, lived on average nine months longer than lefties. Then, in a follow-up study, he and a colleague telephoned the relatives of people named on death certificates in two California counties, to ask which hand the deceased favored. Right-handed people (that study found) lived an average of nine years longer than lefties.
6. It has long been recognized that taller adults hold jobs of higher status and, on average, earn more than other workers. A large number of hypotheses have been put forward to explain the association between height and earnings. In developed countries, researchers have emphasized factors such as self esteem, social dominance, and discrimination. In this paper, we offer a simpler explanation: On average, taller people earn more because they are smarter. As early as age 3—before schooling has had a chance to play a role—and throughout childhood, taller children perform significantly better on cognitive tests. The correlation between height in childhood and adulthood is approximately 0.7 for both men and women, so that tall children are much more likely to become tall adults. As adults, taller individuals are more likely to select into higher paying occupations that require more advanced verbal and numerical skills and greater intelligence, for which they earn handsome returns. Using four data sets from the US and the UK, we find that the height premium in adult earnings can be explained by childhood scores on cognitive tests. Furthermore, we show that taller adults select into occupations that have higher cognitive skill requirements and lower physical skill demands.


7. Does the position of the arm, when blood pressure is being checked, make any difference? Researchers at the University of California at San Diego, using automated cuffs, took six readings from one hundred emergency room patients whose problems did not involve their circulatory systems. Their blood pressure was measured standing, sitting, and lying down; in each position it was measured with the arm straight out from the body and with the arm held at the side. They found that the position of the arm had a bigger effect on the readings than the position of the body. When the arm was parallel to the body readings were higher by as much as 14 millimeters of mercury. Dr. David A. Guss, one of the authors of the study, said that no single position was more accurate, “the most important thing is to use a consistent position from measurement to measurement.”

—From the Annals of Internal Medicine, reported in The New York Times, 6 January 2004

8. Near the end of the Middle Ages, a few theologians (the “scientists” of that time) persuaded a king of France to give them permission for an
experiment that had been forbidden by the Roman Catholic Church. They were allowed to weigh the soul of a criminal by measuring him both before and after his hanging. As usually happens with academics, they came up with a definite result: the soul weighed about an ounce and a half.


9. Undoubtedly the outstanding point of departure of industrial social psychology was the series of studies performed in the Hawthorne plant of the Western Electric Company, starting in 1927. These were conducted by three Harvard professors, Elton Mayo, F. J. Roethlisberger, and T. N. Whitehead, and by W. J. Dickson of Western Electric. The original aim of the studies was to obtain concrete data on the effects of illumination, temperature, rest periods, hours of work, wage rate, etc., upon production. A group of six girls, average workers, were chosen for the experiment; their task was the assembly of telephone relays. Almost from the beginning, unexpected results appeared: The production rate kept going up whether rest periods and hours were increased or decreased! In each experimental period, whatever its conditions, output was higher than in the preceding one. The answer seemed to lie in a number of subtle social factors.

. . . As Homans summarizes it, the increase in the girls’ output rate “could not be related to any change in their conditions of work, whether experimentally induced or not. It could, however, be related to what can only be spoken of as the development of an organized social group in a peculiar and effective relation with its supervisors.”

—S. Stansfeld Sargent and Robert C. Williamson, Social Psychology, 1966

10. Does noise have an adverse effect on those subjected involuntarily to it? When the airport at Munich, Germany, moved, researchers from the University of Hamburg, the University of Gavle in Sweden, and Cornell University took that rare opportunity to conduct a prospective study on the effects of noise, measuring the performance of students near the old airport and near the new one, before and after the move. The reading skills of students in both groups were tested, along with short-term and long-term memory, as reported in the journal Psychological Science, in October 2002. After the move, improvements in memory and reading were found among students near the old airport, while among students living near the new airport, reading skills and memory performance declined.
High levels of noise do interfere with learning and development, those researchers concluded—but the brighter side of their findings was this: Most of the learning damage done by noise appeared to reverse itself when the noise was removed.

11. The mood changes that many people experience during the shorter days of winter have a physiological basis in the brain, according to a study reported in the British medical journal, *The Lancet*, in January 2003. One hundred healthy volunteers, ages 18 to 79, allowed researchers to draw blood samples, at different times of the year, from their jugular veins, to get blood as close to the brain as possible. The researchers then correlated levels of brain chemicals, especially serotonin, with the weather data—temperature, air pressure, rainfall, and sunlight—at the times of blood collection. Only sunlight had causal impact; serotonin levels were found to be lowest in the three months of winter, but varied depending on the brightness of the day. “Our findings [the researchers wrote] are further evidence for the notion that changes in release of serotonin by the brain underlie mood seasonality and seasonal affective disorder.”

12. Prof. Norbert Schwartz, of the University of Michigan, conducted the following experiment. He tested the attitudes of people who had just used a University of Michigan copying machine in which, for some subjects, he had planted a dime which they found, while for others there was no windfall dime. After using the copier, subjects were asked how happy they were about life. Those who had found a dime were consistently more upbeat about “their lives as a whole,” and about the economy and many other matters. “We found,” said Prof. Schwartz, “that a dime can make you happy for about twenty minutes. Then the mood wears off.”


13. The largest and longest-running study of American child care has found that keeping a preschooler in a day care center for a year or more increased the likelihood that the child would become disruptive in class—and that this effect persisted through the sixth grade. Every year spent in such centers for at least 10 hours per week was associated with a 1 per cent higher score on a standardized assessment of problem behaviors completed by teachers. Parents’ guidance, and their genes, had the strongest influence on how children behaved—but this finding about the impact of day care centers held up regardless of the
12.5 Limitations of Inductive Techniques

child’s sex, or family income, and regardless of the quality of the day care center.


14. Speed kills. A report from the Insurance Institute for Highway Safety, issued in November of 2003, concluded that increased speed limits on Interstate highways led to nearly 1,900 additional deaths in 22 states from 1996 to 1999. The report is based, oddly, on a study by the Transport Safety Authority of New Zealand, working in the United States, which showed that, when the Federal cap on speed limits was placed at 65 mph the number of deaths on U.S. highways decreased. But almost immediately after the repeal of that Federal cap on speed limits the number of deaths in the states that did not retain the 65 mph limit increased markedly, while the number of deaths in those states that retained the 65 mph limit did not increase. Drivers in states with higher speed limits, the study showed, drive faster, and where the driving is faster the number of traffic fatalities goes up.

—“Study Links Higher Speed Limits to Deaths” The New York Times, 24 November 2003

15. A 16-year study followed 8,867 non-smoking male professionals with normal body weight who participated in vigorous daily exercise and ate a healthy diet. Those who drank one-half to two normal servings of wine, beer, or hard liquor a day had a 41 to 62 percent reduction of heart attack risk compared with those who drank no alcohol at all. It seems clear that in moderate quantities alcoholic drinks reduce the likelihood of heart attack. This effect is found not only in those with heart disease. The lead author of the study writes: “Even in the lowest risk people, we still find a lower risk associated with moderate drinking.”

—Kenneth Mukamal, “Alcohol Consumption and Risk for Coronary Heart Disease in Men with Healthy Lifestyles,” Archives of Internal Medicine, 23 October 2006

16. For heart patients, “noetic” intervention, like prayer, and therapy relying on music, imagery and touch (MIT), is defined as “an intangible healing influence brought about without the use of a drug, device, or surgical procedure.” 748 patients with coronary heart disease who were to undergo percutaneous coronary intervention (a type of stenting procedure), or elective cardiac catheterization, were enrolled
at one of nine study sites between 1999 and 2002. To test the efficacy of noetic intervention, patients were randomized into four groups: one group (189 patients) received both offsite intercessory prayer and MIT therapy; a second group (182 patients) received intercessory prayer only; a third group (185 patients) received MIT therapy only; the fourth group (192 patients) received neither the intercessory prayer nor the MIT therapy. The interventional heart procedures were conducted according to each institution’s standards practice, with a six-month period of follow-up. The prayer portion was double blinded, meaning that the patients and their care team did not know which patients were receiving intercessory prayer. The prayer groups for the study were located throughout the world and included Buddhist, Muslim, Jewish and many Christian denominations. 89 percent of the patients in this study also knew of someone praying for them outside of the study protocol.

As reported by the Duke University Medical Center, the researchers found no significant difference among the four treatment groups. Distant prayer and the bedside use of music, imagery and touch did not have a significant effect upon the primary clinical outcome of these patients undergoing medical interventions.

—“First Multicenter Trial of Intercessory Prayer,” The Lancet, 16 July 2005

17. The impulse to share does not come naturally to one who is thinking about money. Psychologists found that subconscious reminders of money prompted people to become more independent in their work, and less likely to seek help from others or to provide it. In one experiment 52 undergraduates unscrambled sets of jumbled phrases; one group untangled phrases that were often about money, like “high salary paying,” while another solved word puzzles that did not refer to money. Researchers then had the students work on a difficult abstract puzzle and offered to give help if they wanted it. Those who had been thinking about money worked on the problem by themselves an average of more than 70% longer than the others. Students “primed” to have money on their minds, while clearly self-reliant, were less likely than peers who had not been so primed to lend assistance, twice as slow to help another confused student, and about twice as stingy when asked to donate money to help needy students.

SUMMARY

In this chapter we have examined the concept of cause, the nature of causal connections, and the methods used to establish causal laws.

In Section 12.1, we examined various meanings of “cause.”

In Section 12.2, we explained the supposition of the uniformity of nature, and the generality of causal laws.

In Section 12.3, we discussed induction by simple enumeration.

In Section 12.4, we recounted and illustrated the principal techniques of inductive inference, called Mill’s methods, explaining their essentially eliminative nature. These five methods are

1. The method of agreement
2. The method of difference
3. The joint method of agreement and difference
4. The method of residues
5. The method of concomitant variation

In Section 12.5, we explained the limitations and the strengths of these inductive techniques, concluding that, although they cannot do all that John Stuart Mill had claimed for them, they are profoundly important as the intellectual instruments with which scientific hypotheses are confirmed or disconfirmed.

End Notes

1David Hume, An Enquiry Concerning Human Understanding (1748), sec. IV.


3Reported by Connor O’Shea of the Duke University Medical Center, at the meetings of the European Society of Cardiology in August 2000.


5Dr. Steven Grant, quoted in Science News, 27 January 2007.


The source of these data is the U.S. Census Bureau; the analysts are Gordon B. Dahl of the University of Rochester and Enrico Moretti of the University of California at Los Angeles, reporting online in *Slate*, in October 2003.