

## Editorial

Causality, Unintended Consequences and  
Deducing Shared Causes

Steven M. Shugan\*

Warrington College of Business, University of Florida, 201B Bryan Hall, P.O. Box 117155,  
Gainesville, Florida 32611, steven.shugan@cba.ufl.edu

Despite warnings against inferring causality from observed correlations or statistical dependence, some articles do. Observed correlation is neither necessary nor sufficient to infer causality as defined by the term's everyday usage. For example, a deterministic causal process creates pseudorandom numbers; yet, we observe no correlation between the numbers. Child height correlates with spelling ability because age causes both. Moreover, order is problematic—we hear train whistles before observing trains, yet trains cause whistles.

Scientific methods specifically prohibit inferring causal theories from specific observations (i.e., effects) because, in part, many credible causes are perfectly consistent with available observations. Moreover, actions inferred from effects have more unintended consequences than actions based on sound deductive causal theories because causal theories predict multiple effects. However, an often overlooked but key feature of these theories is that we describe the cause with more variables than the effect. Consequently, inductive processes might appear deductive as the number of effects increases relative to the number of potential causes. For example, in real criminal trials, jurors judge whether sufficient evidence exists to infer guilt. In contrast, determining guilt in criminal mystery novels is deductive because the number of clues (i.e., effects) is large relative to the number of potential suspects (i.e., causes). We can make inferential tasks resemble deductive tasks by increasing the number of effects (i.e., variables) relative to the number of potential causes and seeking a shared cause for all observed effects. Moreover, under some conditions, the method of seeking shared causes might approach deductive reasoning when the number of causes is strictly limited. At least, the resulting number of possible causal theories is far less than the number generated from repeated observations of a single effect (i.e., variable).

*Key words:* causality; common cause; shared causes; deduction; inference; scientific method; structural equations; scientific method

---

**Taboos on Causal Claims**

Most scientific disciplines, people, and organizations implicitly claim insights into causal processes. Those insights allow learning, decisions, predictions, diagnoses, and attributions (e.g., giving credit and placing blame). Causality and so-called causal models (e.g., Fornell and Larcker 1981, Bagozzi and Yi 1988) remain popular. We researchers, however, seem schizophrenic about causality. We refer to causal theories as mathematical models, never mentioning causality. Maruyama (1998, p. 9) states: “the term ‘causal modeling’...fairly quickly fell into disfavor

in the minds of at least a subset of social scientists, who objected to any use of the term ‘causal’ with nonexperimental data. The term ‘causal’ was replaced by the less controversial and very descriptive term ‘structural equations.’” Statistical textbooks emphatically avoid references to causality, focusing on correlation and statistical dependence, yet these textbooks refer to spurious correlations as if some correlations are causal. Some research implies that independent variables cause dependent variables (Shugan 2006), although the difference involves error rather than causality. Some researchers claim experimental methods can detect causality, whereas others warn of confounding variables (e.g., Cox 1992, Greenland et al. 1999). Some researchers claim that although single equations capture only correlations or, more generally, statistical dependence, multiple equations (i.e., causal structural models) can somehow make correlations causal. Some researchers claim to build only predictive models while simultaneously appealing to casual explanations. To avoid the terms cause

---

Editorial pages are not part of the regular *Marketing Science* page budget. We thank the INFORMS Society of Marketing Science for paying for all editorial pages. We also thank the Society for granting every page supplement requested by the current editor.

We welcome and often post responses to editorials. Please see [mktsci.pubs.informs.org](http://mktsci.pubs.informs.org).

\*Steven M. Shugan is the Russell Berrie Eminent Scholar and Professor.

and effect, we refer to exogenous and endogenous variables, but these terms are also problematic, as their definitions remain “subtle and sometimes controversial” (Greene 1997, p. 712). Macroeconomist Kevin Hoover (2001, p. 7) states: “causal language has fallen into disrepute...many modern macroeconomists, instinctively or self-consciously avoid it.” Despite new creative approaches to examining causality, including nonparametric structural equation models (Pearl 2003), graphical methods (Pearl 1995, Krider et al. 2005), and triangulating methodologies (Deshpande 1983), many researchers only insinuate causality while avoiding the term.

One explanation for avoiding causality is that the scientific method provides clever tools (e.g., randomized experiments) for testing causal theories derived from basic principles. Experiments can also test multiple causes (e.g., Gourville and Soman 2005). However, sometimes, we seek to uncover the cause of an effect (e.g., a new product failure, a stock market advance). This inductive objective is problematic. Inductive approaches are fraught with danger because of their inherently unscientific underpinnings. For example, there exist infinite explanations that fit any observation and, of course, deriving and testing a theory on the same observation is tautological. Moreover, actions based on inductive reasoning often cause unintended consequences because we focus on changing the effect, rather than on inferring a deductive causal theory that would allow prediction of multiple consequences.

Scottish economist David Hume (Jones 1969, p. 319) argued that causality was only a perception without physical reality. Perhaps we create causality only when we measure it. Hence, causality is similar to beauty, taste, anger, inspiration, epiphany, insight, thought, anticipation, and perception of cadence with no physical reality. Researchers avoid causal claims for good reason.

## Why Causality Lacks a Conceptual Definition

After thousands of years, at least since the inception of Aristotle’s four causal types, scholars continue to debate the meaning and implications of causality. Incredibly, this deceptively simple concept apparently remains elusive, as misconceptions, conflicting approaches, and fundamental disagreements thwart the investigation of causality. Rather than reviewing the gargantuan literature on causality, let us consider a few examples illustrating some conceptual and practical difficulties relating to causality.

- Suppose Kim, who likes very few flavors, buys a candy because it is chocolate. Consequently, if the candy were not chocolate, Kim should not buy. If something causes an effect, the absence of the cause

should eliminate the effect (i.e., often called a counterfactual). However, if the candy were vanilla, Kim would still buy, because Kim likes vanilla. The cause appears unnecessary for the effect.

- Suppose Lee, a salesperson, loses a client because he fails to provide sufficient service support. Consequently, not providing support appears to cause the client’s loss. However, the firm’s CEO also fails to provide that support. The dubious conclusion is that the CEO not providing service must have caused the client’s loss. Maybe lack of service support itself is the cause, but maybe the client’s unreasonably high expectations regarding support is the cause. The cause appears ambiguous.

- In the preceding example, we could blame the CEO for hiring Lee and, consequently, the CEO might be the true cause for losing the client. However, the board of directors hired the CEO, so perhaps, the board is the cause. Wait, the shareholders voted for the board and, hence, the shareholders are the cause. Maybe the company founders who determined the organization structure are the cause. Every cause appears to have its own cause, *ad infinitum*.

- Similarly, if all events have causes, everything is part of the same causal chain, and we have no influence over causality. For example, from the moment a firm is founded (and before), every decision by management is predictably caused by past events. Hence, to break the chain, some events must not be causes. Unfortunately, we lack an unambiguous method of distinguishing events with causes from those without causes.

- Suppose a buyer purchases one unit of our product. The purchase is certainly not the cause of successful sales. However, if no consumers purchase, there would be no success. Perhaps all buyers must be the cause, but one buyer alone could cause success by purchasing a sufficient quantity. Perhaps quantity is the cause, but then successful sales merely cause successful sales.

- A sales decrease must always follow a previous increase; otherwise, sales would always be zero. Hence, sales increases are necessary and sufficient (assuming bounded sales) for sales decreases. However, few would argue that sales increases cause sales decreases. Similarly, day fails to cause night, despite night always following day. It seems that necessary sequences still contradict causality.

- Suppose astute future psychologists can predict perfectly human behavior. Given that psychologists remain human, they can predict their own behavior, including their own predictions. Given that everything is predictable, events apparently inevitably unfold with destiny, and no cause exists beyond history itself. Everything in the past simply causes the

future. The creation of the firm, the product, distribution, the firm's history of decisions, the buyer's past experiences, and so on, cause the buyer's purchase. Related to determinism, this conclusion contradicts the concept of any exogenous decision and has troubled many researchers, at least since French mathematician Pierre-Simon Laplace's thought experiments involving determinism (Earman 1986, pp. 5–7).

- Similarly, causality might be simply a linear time line. For example, we construct a function  $S(t)$  that describes sales as a function of time  $t$ . The function predicts the future, the present, and the past sales. The function captures all the relevant decisions (advertising, the economy, prices, etc.) as a function of time  $t$ . The concept of causality is irrelevant and merely confuses matters.

- We theorize that latent unobservable utilities cause buyers to purchase products. Although direct observations of utility are impossible, we infer them from behavior. However, perhaps invisible, undetectable, mythical fairies cause purchase behavior, and we implicitly infer their utilities. Certainly, claiming an unobservable cause is problematic and unscientific, because causal theories with unobservable (i.e., hidden) variables are often tautological.

- Understanding causality only has value if causes and effects reoccur. However, no two situations are identical in every respect. We might drive home every night. Although we appear to return to the same situation or state, we have not. On the trip home, we have interacted and influenced other drivers, pedestrians, vendors (gas stations, drugstores) and, perhaps, have had profound effects unknown to us.

- Similarly, suppose firms find that lowering product quality causes profits to increase; however, they are unable to guarantee that the effect will persist. It might take time for current customers to find alternative offerings, for competitors to enter the market, and for customers to recognize lower durability and reliability. More importantly, everyone might learn. Hence, causes never have the same effects.

- We dismiss obviously necessary causes. For example, we require the stellar sun for the buyer to exist, but we seldom consider that the stellar sun causes purchasing behavior.

- No one cause appears necessary. There are hundreds of ways, for example, to cause a buyer to purchase a product.

- No cause appears sufficient. A buyer will fail to purchase a product unless we have a buyer, a product, needs, money, distribution, oxygen for the buyer, and so on.

- Causal questions often lack answers. For example, we ask whether advertising increases sales. Advertising does and does not. Worse, sales can increase or decrease without advertising.

- The concept of randomness necessitates a complete lack of causality because, if there were a cause, the variable would not be random.

- Capturing all intuitive definitions of causality is difficult, if not impossible.

## Causes Can Follow Effects in Time

David Hume viewed causality as psychological impressions created from repeatedly observing the same sequence of events so that "on its next occurrence, the first impression of the sequence will again be followed by the second" (Jones 1969, p. 319). For example, when one is first viewing a movie, scenes occur without causality (e.g., a bus is late and two people meet). After repeated viewing, the moviegoer infers that the late bus causes the meeting.

Most researchers accept Hume's sequential ordering, so that causes precede effects. Reichenbach (2000, p. 27) attempts to define causality without appeal to sequential orderings, but acknowledges that "...when we are asked how to distinguish the cause from the effect, we usually say that... the cause is the one that precedes the other in time." However, some researchers discuss true backward causality (Dowe 2007).

Unfortunately, we can observe effects before their causes, at least, the intuitive causes. For example, we hear train whistles before the train, but the train is the cause. Retail orders for Christmas occur before Christmas, but Christmas causes the orders. Emergency workers observe phone calls to 911 before observing people in distress, but 911 calls do not cause the distress (see Shugan (2006) for other examples). Experimentation could theoretically disentangle causality, but observed causes can follow observed effects in time.

## Probabilistic Causality—Not a Panacea

Some researchers argue that probabilistic causality could potentially solve the causality enigmas (e.g., Granger 1988, Pratt and Schlaifer 1988, Zellner 1988). These researchers debate whether probabilistic uncertainty implies truly random discrete processes, as encountered in quantum mechanics, or Bayesian uncertainty reflecting the decision maker's ignorance (Popper 1950). Either interpretation allows inference of causal relationships and unobserved common causes from covariance matrices (Spirtes et al. 1988, 65–71).

Despite the merits of probabilistic causality, if understanding deterministic causality is frustrating, then introducing uncertainty, random disturbances, and sampling error clouds understanding, at best. At worst, probabilistic interpretations could hide fundamental logical inconsistencies similar to the manner in

which appeals to uncertainty can sustain bad theories (Shugan 2006).

For example, we might argue that a suspected pathogen causes a disease, but the pathogen is absent from some afflicted individuals while found in some nonafflicted persons. Correlation or statistical dependence can be irrelevant or, worse, misleading. The suspected pathogen might be innocuous while the true pathogen populates all afflicted individuals and only afflicted individuals. Maybe if the brakes fail to always stop the car, we need a better causal theory of braking. Retreating to probabilistic interpretations of causality can prolong the rejection of false theories or inhibit the development of better theories. Although probabilistic interpretations allow creative refinement of nearly correct ideas, many renowned researchers specifically caution against proposing any theory derived from specific empirical observations, and they argue that such proposals violate the fundamental tenets of scientific methodology (e.g., Einstein 1919, Popper 2002). Even advocates of probabilistic approaches warn that searching for causality “is hard to achieve except when large effects are involved” (Cox and Wermuth 2004, p. 303), making errors relatively small. Simultaneity is another problem. It is difficult, for example, to disentangle whether loyalty causes multiple purchases or the reverse or both (Gupta and Zeithaml 2006). Probabilistic causality is insufficient.

### Causality Without Correlation

Statistical textbooks are replete with examples of correlation without causality. For example, children’s spelling skills correlate with foot size, but age causes both. The number of churches correlates with the number of crimes, but city size causes both. Shark attacks correlate with barbecues, but warm weather causes both. Hence, common causes create correlations. Moreover, the causal direction is often unclear (e.g., Moorman et al. 2005, Keller and Lehmann 2006). Alcoholic drinking might cause depression, or depression might cause drinking (Shugan 2006).

In contrast, most textbooks incorrectly assert that observed correlation is necessary for causality. The existence of pseudorandomness (uncorrelated numbers generated by entirely deterministic causal processes) seems sufficient to prove observed correlation is unnecessary. For an everyday example, consider the case of a skilled driver steering an automobile down a curved two-way highway. With expert steering, the driver maintains an equal distance between the car and the median strip, keeping this distance constant. Consequently, there is no observed correlation between that distance and the steering wheel’s alignment, despite general agreement that steering is precisely the cause of the observed distance.

Often, understanding and properly managing control variables will eliminate or reverse expected observed correlations. Increasing quality should cause higher sales, but firms raise prices, eliminating observed correlations between quality and sales. Eating low fat foods should cause weight loss, but consumers increase their consumption, eliminating the expected observed correlation. Perhaps these paradoxical examples result from choosing the wrong variables or from ignoring the right variables. However, causality exists without observed correlations.

### Causality Without Change

Granger (1969) and others have proposed clever tests for revealing causal relationships by measuring changes in causes and effects. For example, when (1) increases in advertising expenditures precede increases in consumption, and (2) increases in consumption are unrelated to subsequent increases in advertising expenditures, then we conclude that advertising causes consumptions, rather than the reverse. Parenthetically, for advertising, Granger tests reveal evidence for causality in both directions (Ashley et al. 1980). Granger tests can also reveal whether objective quality causes perceived quality (Mittra and Golder 2006).

However, neither the cause nor the effect needs to change. The sun’s gravity causes the earth’s orbit, yet the sun’s gravity remains constant. Providing great, consistent service causes the firm’s sales to grow; yet, constant service causes changes in sales. An airplane pilot causes a constant altitude by continually changing the controls. A wall causes a roof to remain stationary. Although experiments might pursue manipulations involving change, rather than controlling for change, the manipulation can create change not present in reality.

More generally, change itself depends on the observer’s perspective. A buyer on vacation observes a change in the layout of a local Walgreen’s store (compared to home). However, local buyers observe no change. Microsoft Windows XP users find favorable changes in Windows Vista while Apple Computer users find unfavorable changes. Attempting to hold perspective constant ultimately requires that all observers experience the same history. Causality can exist without change.

### Why Causes Have More Detail than Effects

When building a model or theory, we make 90% of the assumptions when we select which variables to observe and, implicitly, which not to observe. The selected variables assume possible causes and effects. More importantly, we implicitly assume that history

repeats itself when the selected variables take the same values.

The claim is that two situations are alike (if not identical) when the variables assume the same values. Therefore, knowledge about one situation allows us to predict others. The proposed relationships between the variables should persist, regardless of changes in unobserved variables.

As noted earlier, no two situations are identical. History never truly repeats itself because it always grows longer. Whereas the selected variables might return to previous values, many other variables assume new values. Moreover, previously absent variables might appear. For example, launching a new product today differs from launching one tomorrow, because a day has passed. People grow older; new people are born; technology advances; some people learn while others forget. Markets change by the day, if not by the minute.

Only on our small set of selected variables are two situations identical, at least with respect to interrelationships. For example, we return home each night to the same location, but we interacted with the environment (e.g., breathed the air, created sounds, disturbed some animals, fed some insects), influenced other people (e.g., purchased items, encountered congested traffic, changed the future of others) and, certainly, aged physically. Given that situations always differ, both causes and effects must involve a limited number of observed variables. Moreover, the relationships between these variables must persist, regardless of changes in unobserved variables. Satisfying this requirement is easier when we describe the cause in more detail than the effect. Consequently, the cause is less likely to recur than the effect because we place greater limits on admissible causes, i.e., situations when we make predictions.

For example, we predict a single variable, sales, from a cause consisting of advertising, price, distribution, product features, and so on. Hence, the cause requires that many variables assume specific values. The effect, in contrast, requires prediction of only one variable. Having more detailed causes eases the task of deducing effects, because we potentially use more information to predict less. Moreover, we put more conditions on the cause because a situation causes only one outcome but many situations produce the same outcome. The reverse is also true. When inferring causes from effects, adding more detail to effects eases the task of inferring causes.

## A Simple Definition of Causality

Researchers, attempting to celebrate the general applicability of their theories, promote terminology resembling everyday language. Statisticians use the term

“significance” but quickly disavow any association with the everyday meaning of practical or important. Accountants use the term “asset value” but disavow any association resembling the everyday meaning of worth. The concept of “work” in physics dramatically differs from the everyday meaning. The terms have precise unambiguous definitions despite their appeal to generality.

When defining terms, we should sacrifice the relevance (and ambiguity) of everyday language for the more precise (and narrow). When defining causality, for example, Granger (1980) notes that “causation has no generally accepted definition. . . . Provided I define what I personally mean by causation, I can use the term. I could, if I so wish, replace the word cause throughout my lecture by some other words, such as ‘oshkosh’ or ‘snerd,’ but what would be gained?”

Common language often lacks mathematical precision. For example, mathematical statements are either true or false. Verbal statements, in contrast, can be inconsistent, meaningless, ambiguous, or nonsensical. Statements such as “yellow is large,” “I am lying,” and “ugly can be beautiful” are neither true nor false. The same is true of evidence for a causal theory. Some evidence is neither consistent nor inconsistent but merely irrelevant or meaningless.

Therefore, let us adopt the narrow definition of causality inspired by Bayesian decision theory (e.g., see Buzzell 1962, Green 1963, Savage 1972, Berger 1980). Decisions are actions that cause different outcomes. Causes are actions defined on control variables, and effects are outcomes with particular characteristics. For example, the decision (the cause) is launching a new product or not. The effects are outcomes possessing particular characteristics (gains in reputation, short-term profits, long-term profits, cash flow, sales growth, etc.).

Decision makers could be uncertain about which actions produce which outcomes, but that uncertainty reflects insufficient information (i.e., Bayesian uncertainty) about causality while outcomes are deterministic. Exhaustive observation would resolve all uncertainty. Decisions recur when one is facing the same actions and outcomes.

## Experimentation and Controls

Arguably, the greatest advance for determining causality, and perhaps science as well, was English statistician Sir Ronald Fisher’s ingenious randomized experiment. Today, most researchers argue that only controlled experiments reveal causality. Holland (1986, p. 959) states that “it takes two causes to define an effect” and his motto is “no causation without manipulation.” Box et al. (1978, p. 495) state: “Broadly speaking, to find out what happens when you change

something, it is necessary to change it.” Even Danish physicist Hans Christian Ørsted’s famous method involving thought experiments (Sorensen 1992, p. 57) are routine for theoretical scientific inquiries into causality (Bishop 1999)

Experiments examine effects with and without the cause. The control (i.e., without the cause) tests whether the cause is necessary. Then, we test whether the cause is sufficient. For example, to examine the effect of advertising, we must both not advertise and advertise. Naturally occurring data lacks essential information because, in this example, decision makers choose advertising levels based on the perceived effect, thereby eliminating necessary observations. If we observe a perfect relationship between firm market share and firm profits, we remain ignorant of whether larger market shares cause larger profits, larger profits cause larger market shares, good management causes both, or public data are missing for profitable smaller share companies (perhaps, because they are private or have been acquired).

However, experiments seem less necessary when deductive causal theories make sufficiently unexpected predictions (i.e., inconsistent with other theories) that allow falsification. For example, consider physicist Sir Arthur Stanley Eddington’s confirmation of physicist Albert Einstein’s predictions that light would bend around the sun during a solar eclipse as predicted by the theory of relativity (e.g., Will 1993, p. 171). Similar predictions provide evidence of causality without experimental data. Hence, nonexperimental predictions, when nonobvious and unique, might be useful.

Moreover, constructing control groups can be difficult, if not impossible. First, observer effects can cause measurement to alter observation. For example, television networks and other programmers change content when Nielsen measures audience size (e.g., Lowry 2002).

The Heisenberg uncertainty principle, often confused with observer effects, at least until 1935 (Bohr 1935, McKerrow and McKerrow 1991), states that gaining accuracy on one variable’s measurement decreases accuracy for other variables. Although the uncertainty principle only applies to duality properties of subatomic particles, there are analogous situations. For example, we are unable to measure simultaneously an automobile’s ability to accelerate and brake at a point on a hilly road. We must first measure acceleration, then braking. To measure acceleration, more accurately, we spend more time accelerating (say, from 10 to 20 miles per hour). However, more time causes the automobile’s position to shift and makes the braking measurement less accurate because the road’s gradient and the engine’s state have changed. Braking and

acceleration are conjugate properties of the automobile. Both are real; both can be measured; but not together.

This problem also inflicts experimental controls. The control can never be simultaneous with the cause. For example, we are unable to launch and not launch a new product in a city at the same time. We must assume, for example, that without a launch, control and test cities remain identical over time. The longer the experiment, the more likely the cities diverge.

## Determining Effects and Unintended Consequences

The scientific method requires that we deduce causal theories from principles (Einstein 1919, Popper 2002), rigorously derive testable observable implications (i.e., effects), scrutinize the consistency of deduced effects with subsequent observation, and always make predictions before making observations (i.e., not inferring explanations after making observations). Perhaps it is best to avoid inadvertent tester bias, and different researchers should propose and test theories. For example, Einstein et al. (1935), using a thought (Gedanken) experiment, deduced a disturbing prediction of quantum mechanical theory regarding non-local causality. Years later, John Stewart Bell (1964) conceptually designed an experiment for testing the prediction. Years later, Alain Aspect et al. (1981) creatively implemented the experiment.

The scientific method allows straightforward deduction of many effects. For example, compensation theories predict salesperson behavior on many variables (Basu et al. 1985, Lal and Staelin 1986). Different commission schemes, in contrast to fixed-salary schemes, should produce different time allocations for higher potential customers, higher margin products, travel, postsales support, and many other observable effects (Hauser et al. 1994).

Moreover, good causal theories should predict all effects or consequences of causal actions (i.e., actions intended to cause an effect). Inference, in contrast, limits our attention to one effect rather than predicting all effects of causal actions. Unintended consequences (i.e., unexpected effects) become problematic when inferring causes from effects.

For example, legislation mandating worthwhile safety features in automobiles raises the price of automobiles, encourages buyers to delay purchases, increases the average age of automobiles, increases the number of older, badly maintained automobiles on the road, and so on. Economic theory predicts all of these consequences.

Unexpected consequences often occur when one is inferring an action to remedy undesirable features of a single effect. For example, to remedy last minute unsold seats, an airline might offer regular

last-minute discounts. Unexpected effects of discounting might include some passengers delaying purchases to obtain discounts (Desiraju and Shugan 1999, Xie and Shugan 2001). Similarly, increasing sales volume with price promotions causes other effects, such as encouraging forward buying, inventorying items and causing some buyers not to purchase at regular prices. A reasonable causal theory of buyer behavior would predict these unintended consequences. In contrast, inferring causal actions from single effects often spawns unintended effects.

## Determining Causes and the Perils of Inference

Although scientific methods warn against inference, we must sometimes find a cause from an effect, particularly an unexpected effect such as a precipitous sales decline or an epidemic outbreak. Something is unexpected when our extant causal theory fails. Consequently, we perilously infer a cause to determine appropriate future actions without a causal theory.

As noted earlier, the scientific method makes a clear distinction between induction and deduction. Inductive inference allows invaluable activities (e.g., testing, learning, etc.). However, inference without a priori theory is dangerous and generally considered unscientific, because inference is prone to uncorrectable errors. There are numerous, if not infinite, explanations for any specific observation (e.g., a stock market decline). More seriously, inferential procedures cement incorrect inchoate theories because confirmation bias amplifies confirming features of the observation reinforcing erroneous inferences. The process deteriorates.

Rather than reviewing the related gargantuan literature, consider a few examples of the perils of inference.

- Philosopher Carl Gustav Hempel's famous proof shows that observing white shirts provides inferential evidence that ravens are black, because observing a black raven is logically equivalent (i.e., the contrapositive) to observing a not raven that is not black (Hempel 1945). Simply stated, any observation not conflicting with a theory becomes inferential support for the theory. Deductive theory, in contrast, narrows relevant evidence to very specific falsifiable predictions, perhaps concerning raven genetics.

- Unlike deduction, inductive inferences are rarely replicable. Experts can infer different conclusions from the same observations.

- Inductive inference is subject to many psychological biases. Psychologist Peter Cathcart Wason's famous experiments (Wason 1968) found that most people are unable to infer extremely simple underlying processes from sequences of numbers (e.g., 2, 4, 6). His four-card experiments suggest great difficulty correcting wrong inferences.

- We observe many calendars in nonleap years—None contains February 29<sup>th</sup>. Any inference about February 29<sup>th</sup> in other years is wrong. Inference only allows generalizations within the same population with no method for knowing that population (e.g., other calendars in the same year, English-language calendars, etc.).

- Holland (1986, p. 959), for example, notes "... the 'cause' of a given effect is always subject to revision as our knowledge about the phenomenon increases. For example, do bacteria cause disease? Well, yes... until we dig deeper and find that it is the toxins the bacteria produce that really cause the disease; and this is really not it either. Certain chemical reactions are the real causes... and so on, ad infinitum." The inferred cause appears subjective.

- Pete made a bad marketing decision causing an undesirable effect, so Pete is the cause and eliminating Pete prevents the future effect. However, Pete has learned from the experience and he would do better with future decisions. Valid inferences might not persist and become useless for prediction.

- Very high ticket prices for a celebrity's concert upsets loyal fans. Focusing on the effect (upset loyal fans), the celebrity lowers ticket prices for future concerts and brokers buy all tickets before the fans. Inference of causal actions from specific events often spawns many unexpected consequences.

Deductive causal theory can solve all of these problems by predicting when and why particular effects will occur. Testable predictions test the theory. More importantly, deductive theories produce fewer unexpected effects because they allow predictions on many variables in many domains. For example, trying to reduce salesperson travel expenses, we observe expenses and find high meal costs, so we limit spending on meals. Deductive theories of incentives predict unintended consequences on many variables. For example, salespeople should be reluctant to entertain new clients; salespeople should spend their budgets on less productive activities; the best salespeople might take jobs at other employers; and so on. Maybe better solutions include redesigning sales territories, negotiating hotel rates, or providing salespeople with better information regarding effective entertainment strategies.

Parenthetically, unlike dictionary definitions of inferential induction that emphasize inferring general principles from specific observations, statistical technologies define inference differently. Estimation technology usually uses predetermined formulae (not inferred from observation) to calculate parameters that maximize the consistency of the theory with observations and, subsequently, report that consistency. End-users can make their own inferences. Bayesian technologies use predetermined formulae

that allow decision makers to assimilate new information. Bayesian inference only dictates changes in uncertainty. Although both technologies claim property rights to the word “probability” and advocate conflicting definitions, the two technologies are compatible because they are suitable for different tasks. The highly criticized Bayesian prior probability distribution might hinder the task of falsification while being indispensable for decision-making tasks. Neither technology truly infers general theories from specific observations.

### A Unique Cause

There are possibly infinite true causes for an effect when causes themselves have causes. For practical purposes, let the relevant cause imply the most economical action, i.e., the action that minimizes the cost function across all actions achieving the desired features of resulting effects. For example, actions preventing new product failures include having no new products, disbanding our business, or changing a product attribute. Actions preventing epidemic outbreaks include eliminating people, relocating people, or stopping transmission. The last actions seem the most economical.

The most economical action within a causal chain is usually closest to the effect. For example, the president hires the marketing vice president, who hires a market research firm, which uses a forecasting model, which encourages a new product launch that subsequently fails. The most economical action would involve the forecasting model, provided it achieves the desired effect. For precise definitions, the terminology (i.e., primitives) of decision theory suffices. Finally, although there may be no unique economical action that works, assuming so would be counter-productive if it did exist.

### Making Inference Deductive (or Near Deductive)

Although scientific methods specifically prohibit inferring causal theories from specific observations because many credible causes can be perfectly consistent with available observations, when the number of possible causes is relatively small, the task becomes deductive. For example, consider the typical inferential crime mystery in which a suspected perpetrator (i.e., the cause) commits a crime (i.e., an effect). The criminal trial parallels statistical inference. Prosecutors present evidence. Jurors decide whether sufficient evidence exists to infer guilt, given a reasonable doubt (i.e., a type I error). As an inferential process, mistakes are possible. Juries sometimes convict the innocent and exonerate the guilty. Hence, the prosecutor’s so-called proof is a misnomer, only representing inductive consistency between the prosecutor’s beliefs and

observed evidence. Parenthetically, courts explicitly prohibit Bayesian analysis by rarely allowing the use of a suspect’s past criminal history to form a Bayesian prior probability distribution.

In contrast, the criminal mystery novel seems similar. The reader observes the evidence. The reader infers guilt or innocence. Readers can be wrong. However, this task is deductive. Deductive logic and a few simple mathematical equations reveal, without error, the perpetrator’s identity.

The critical difference between real criminal trials and mystery novels is the limited set of potential perpetrators relative to the number of clues. Hence, the number of potential causal theories (i.e., suspects) is smaller than the number of observed effects (i.e., clues). We transform an inferential task into a deductive task by increasing the number of effects relative to the number of potential causes.

Deductive prediction commonly involves using many observable causal variables to deduce conclusions about a smaller number of effects. For example, increasing product awareness, availability, quality, reliability, and durability increases sales. Hence, live variables predict one (i.e., sales).

Inductive inference lacks this property. Induction often involves going from the few to the many. Hence, dramatically increasing the number of observed effects (i.e., variables) might allow deduction. Similarly, more clues allow unique identification of the guilt perpetrator. Analogously, inferential statistics become descriptive statistics when we obtain a census for the entire population. Perhaps increasing the number of observed effects (i.e., variables) will allow inference of a single cause, making the process resemble deduction. For example, physician Jerome Groopman (2007) advocates encouraging patients to provide richer sets of symptoms (i.e., effects), rather than inferring the most likely cause of prominent symptoms.

Note that increasing the number of observed effects differs from observing repeated instances of the same effect. Observing more effects increases the number of observed variables, rather than the number of observations of the same variables. For example, repeated observations of sales at different outlets differ from chronicling all unexpected observations.

### Shared Causes

By expanding the number of observed effects (or variables), we attempt to infer a shared cause for all the observed effects. By not trying to better fit the cause to a single effect (or single variable), we avoid the usual deficiencies associated with inductive approaches such as overlying on situation-specific factors (i.e., over-fitting). If a causal action occurs, then it should cause multiple observable effects beyond the specific outcome of interest. These other predictable effects might be simultaneous with the effect of

interest or nearby (in time and space). Better theories provide richer predictions and, consequently, produce a greater variety of specific predictions. For example, the rich theory of rational consumer expectations makes predictions on every consumer decision in many different situations.

Having a sufficient number of observed effects might greatly limit the number of consistent causal theories. For example, other unexpected observations occurring with sales declines might include loss of customers, low purchase rates, low levels of customer satisfaction, many product returns, bad customer perceptions, distributor concerns, and so on.

Although finding shared causes fails to replicate a deductive process, it does increase the number of observed effects relative to the number of potential causes, at least, obvious causes. Moreover, under some conditions, the method of seeking shared causes might converge to deductive reasoning when the number of causal theories is strictly limited. At least, the resulting set of causal theories is far less than the number generated from repeated observations of a single effect. Unfortunately, some relevant causal theories might be inconspicuous. Hence, although observing many effects might simulate deduction, errors remain possible.

### Shared Causes—A New Product Example

Consider two previous examples—the new product failure and the epidemic. In the case of the new product, we can look for other effects. For example, have there been other recent failures? The products from other firms might have failed; other products from the same firm might have failed. That might indicate specific causes beyond defective product attributes, such as a faulty new product development process, a flawed distribution channel, a poorly run organization, an unsound corporate strategy, an unworkable environment, and so on. However, more importantly, as we increase the number of observed events, we should be able to limit the number of candidate causal theories. For example, if many other new products succeeded before and after the failure, the cause must be more specific to this product.

The strategy is finding a shared cause for all observed effects. Increasing the number of observed effects narrows the number of potential shared causes. Rather than trying to explain an effect by seeking other instances of the effect, we seek observations of other local effects. This process should ultimately eliminate all but one shared cause. Seeking a shared cause simulates a more deductive approach. Of course, seeking a shared cause remains inferential unless the number of candidate causes is both limited and known.

### Shared Causes—An Epidemic Example

Plague Bacillus inflicted massive human disaster across the world, devastating and decimating both cities and countries. Candidate causes for the disease included the star's unfavorable constellations, comets, supernatural powers, and poisoning of wells by Jews... "who paid for this with tortures inflicted on them by the panicked population" (Gross 1995, p. 7609). Centuries later, facing government resistance, physician Alexandre Emile John Yersin visited plague-ravaged Hong Kong. An English priest, Father Vigano, magnanimously supplied Yersin with a small shack for a crude laboratory. Bribing two English sailors (Marriott 2004, p. 136) at great personal risk, Yersin gained access to the hospital morgue. Micro-observation of abnormalities in plague fatalities consistently revealed rod-shaped microbes. Yersin tested the causal theory by injecting the microbes into guinea pigs that subsequently died with large quantities of the microbe.

Unlike other researchers who focused on laboratory observation of fatalities, Yersin also made extensive environmental observations, finding other effects of the plague. Among recorded observations were large numbers of dead rats. In plague-ravaged Long-Tcheon Indochina, bacteriologist Paul-Louis Simond, inspired by Yersin, also expanded observation to outside the laboratory. Unlike others, he examined blisters on workers' feet, finding high levels of plague bacilli (Marriott 2004, p. 236). He observed geographic separation between outbreaks (Marriott 2004, p. 237). He observed that, "one day, in a wool factory, employees arriving in the morning noticed a large number of dead rats on the floor. Twenty laborers were ordered to clean the floor of the dead animals. Within 3 days, 10 of them developed plague, whereas none of the other employees became ill" (Gross 1995, p. 7610). Making further observations on still more variables (e.g., the nature of the contact, when the contact occurred), Simond found a shared cause—contact with recently deceased rats (i.e., rats that were not living but not cold dead). Investigating all events occurring during death, he found that "immediately after the rat dies, when its corpse cools off, the rat flea (*Xenopsylla cheopis*) leaves the dead rat and jumps on other healthy rats or, if no rats are available, they jump on men or women." Using very crude laboratory equipment (e.g., paper bags), Simond tested the causal theory. Simond, for example, showed that inflicted rats with no fleas did not transmit the plague to other rats.

Yersin and Simond isolated the shared cause by expanding the number of observed effects. Other researchers (e.g., Shibasaburo Kitasato) also isolated microbes but limited observations to the laboratory. The cause of transmission revealed itself (after many

centuries of arduous effort by investigators) only after expanding the number of observed effects.

## Conclusions

Researchers avoid causal claims for very good reasons. Causality often lacks both conceptual and operational definitions capturing the everyday intuitive definitions for causality. Moreover, intuitive causal relationships seem to violate scholarly definitions. For example, observed causes can follow observed effects in time. Causality can exist with or without observed correlation between the cause and the effect. Causality can exist without observing any change. Perhaps it is best to adopt a specific definition of causality—a causal action produces effects having particular features. We use the most economical (i.e., lowest-cost) causal action. Our causal models, then, predict the effects of actions.

When building causal models for prediction, we make 90% of the assumptions when we select which variables to observe and, implicitly, which not to observe. Causes and effects are limited to those variables. An often overlooked but key feature of modeling is that we describe the cause with more variables than the effect. Qualitatively, we describe the effect in less detail than the cause. This is a key difference between inference and deduction.

For example, a real criminal trial involves inference. Jurors decide whether sufficient evidence exists to infer guilt, given reasonable doubt (i.e., a type I error). In contrast, determining guilt in a criminal mystery novel is deductive because the number of clues (i.e., effects) is large relative to the number of potential suspects (i.e., causes). We transform an inferential task into a deductive task by increasing the number of effects relative to the number of potential causes.

One reason scientific methods specifically prohibit inferring causal theories from specific observations is that many credible causes are perfectly consistent with available observations, as evidenced by most political and scientific debates. This is the well-known shortcoming of inductive reasoning.

However, when the number of possible causes is sufficiently small, the task becomes deductive. Consequently, inductive tasks might also approach deductive tasks when we expand the number of observed effects (i.e., variables). Note that increasing the number of effects is not equivalent to observing repeated instances of the same effect. Observing more effects increases the number of observed variables rather than the number of observations of the original variables. Yersin and Simond solving of the Bacillus plague puzzle provides an example.

By increasing the number of observed effects (i.e., variables), we limit the number of possible shared

causes, i.e., causes that predict all of the observed effects. A sufficient number of observed effects might greatly limit the number of consistent causal theories. Although the method of seeking shared causes fails to replicate deductive processes, it might approach one. Moreover, under some conditions, the method of seeking shared causes might converge to deductive reasoning when the number of candidate causal theories is strictly limited. At least the resulting number of causal theories is far less than the number generated from repeated observations over a single effect or a much smaller number than the universe of candidate causal theories.

## References

- Ashley, R., C. W. J. Granger, R. Schmalensee. 1980. Advertising and aggregate consumption: An analysis of causality. *Econometrica* 48(5) 1149–1167.
- Aspect, A., J. Dalibard, G. Roger. 1982. Experimental test of Bell's inequalities using time-varying analyzers. *Physical Rev. Lett.* 49(25) 1804–1807.
- Bagozzi, R. P., Y. Yi. 1988. On the evaluation of structural equation models. *J. Acad. Marketing Sci.* 16(1) 74–94.
- Basu, A. K., R. Lal, V. Srinivasan, R. Staelin. 1985. A theory of sales-force compensation plans. *Marketing Sci.* 4(Fall) 267–291.
- Bell, J. S. 1964. On the Einstein-Podolsky-Rosen paradox. *Physics* 1(3) 195–200.
- Berger, J. O. 1980. *Statistical Decision Theory and Bayesian Analysis*. Springer-Verlag, New York.
- Bishop, M. A. 1999. Why thought experiments are not arguments. *Philosophy Sci.* 66(4) 534–541.
- Bohr, N. 1935. Can quantum-mechanical description of physical reality be considered complete? *Physical Rev.* 48(8) 696–702.
- Box, G. E. P., E. Pelham, W. G. Hunter, J. S. Hunter, S. J. Hunter, W. G. Hunter. 1978. *Statistics for Experimenters: An Introduction To Design, Data Analysis, and Model Building*. John Wiley and Sons, Inc., New York.
- Buzzell, R. D. 1962. Decision theory and marketing management. *J. Marketing* 26(3) 7–16.
- Cox, D. R., N. Wermuth. 2004. A comment on the coefficient of determination for binary responses. *Internat. Statist. Rev.* 72(3) 285–305.
- Cox, D. R. 1992. Causality: Some statistical aspects. *J. Roy. Statist. Soc., Ser. A* 155(2) 291–301.
- Deshpande, R. 1983. Paradigms lost: On theory and method in research in marketing. *J. Marketing* 47(4) 101–110.
- Desiraju, R., S. Shugan. 1999. Strategic service pricing and yield management. *J. Marketing* 63(1) 44–56.
- Dowe, P. 1996. Backwards causation and the direction of causal processes. *Mind* 105(418) 227–248.
- Earman, J. 1986. *Primer on Determinism. Series in Philosophy of Science*, Vol. 32. Springer-Verlag, New York.
- Einstein, A., B. Podolsky, N. Rosen. 1935. Can quantum-mechanical description of physical reality be considered complete? *Physical Rev.* 47(10) 777–780.
- Einstein, A. 1919. Induction and deduction in physics. *Berliner Tageblatt* (December 25) [Translation by Alfred J. Engel. 2002. *The Collected Papers of Albert Einstein: Volume 7. The Berlin Years: Writings, 1918–1921*. Princeton University Press, Princeton, NJ, 221–236.]
- Fornell, C., D. F. Larcker. 1981. Evaluating structural equation models with unobservable variables and measurement error. *J. Marketing Res.* 18(1) 39–50.
- Gourville, J. T., D. Soman. 2005. Overchoice and assortment type: When and why variety backfires. *Marketing Sci.* 24(3) 382–395.

- Granger, C. W. J. 1969. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica* 37(3) 424–438.
- Granger, C. W. J. 1980. Testing for causality: A personal viewpoint. *J. Econom. Dynam. Control* 2(1) 329–352.
- Granger, C. W. J. 1988. Some recent development in a concept of causality. *J. Econometrics* 39(1-2) 199–211.
- Green, P. E. 1963. Bayesian decision theory in pricing. *J. Marketing* 27(1) 5–14.
- Greene, W. H. 1997. *Econometric analysis*. Prentice Hall, New York.
- Greenland, S., J. M. Robins, J. Pearl. 1999. Confounding and collapsibility in causal inference. *Statist. Sci.* 14(1) 29–46.
- Groopman, J. 2007. *How Doctors Think*. Houghton Mifflin, Boston, MA.
- Gross, L. 1995. How the plague bacillus and its transmission through fleas were discovered: Reminiscences from my years at the Pasteur Institute in Paris. *Proc. National Acad. Sci. U.S.A.* 92(17) 7609–7611.
- Gupta, S., V. Zeithaml. 2006. Customer metrics and their impact on financial performance. *Marketing Sci.* 25(6) 718–739.
- Hauser, J. R., D. I. Simester, B. Wernerfelt. 1994. Customer satisfaction incentives. *Marketing Sci.* 13(4) 327–350.
- Hempel, C. G. 1945. Studies in the logic of confirmation. *Mind (A quarterly review of Psychology and Philosophy)* 54(213) 1–26.
- Holland, P. W. 1986. Statistics and causal inference. *J. Amer. Statist. Assoc.* 81(396) 945–960.
- Hoover, K. 2001. *Causality in Macroeconomics*. Cambridge University Press, Cambridge, UK.
- Jones, W. T. 1969. *Hobbes To Hume: A History of Western Philosophy*. Harcourt, Brace, and World, New York.
- Keller, K. L., D. R. Lehmann. 2006. Brands and branding: Research findings and future priorities. *Marketing Sci.* 25(6) 740–759.
- Krider, R. E., T. Li, Y. Liu, C. B. Weinberg. 2005. The lead-lag puzzle of demand and distribution: A graphical method applied to movies. *Marketing Sci.* 24(4) 635–645.
- Lal, R., R. Staelin. 1986. Salesforce compensation plans in environments with asymmetric information. *Marketing Sci.* 5(3) 179–198.
- Lowry, B. 2002. Pressing the flesh in the ratings war. *Los Angeles Times* (February 13) 18.
- Marriott, E. 2004. *Plague: A Story of Science, Rivalry, and the Scourge That Won't Go Away Laws of Imitation*. Henry Holt and Company, New York.
- Maruyama, G. M. 1998. *Basics of Structural Equation Modeling*. Sage, Thousand Oaks, CA.
- McKerrow, K. K., J. E. McKerrow. 1991. Naturalistic misunderstanding of the Heisenberg uncertainty principle. *Educational Res.* 20(1) 17–20.
- Mitra, Debanjan, Peter N. Golder. 2006. How does objective quality affect perceived quality? Short-term effects, long-term effects, and asymmetries. *Marketing Sci.* 25(3) 230–247.
- Moorman, C., R. Du, C. F. Mela. 2005. The effect of standardized information on firm survival and marketing strategies. *Marketing Sci.* 24(2) 263–274.
- Pearl, J. 1988. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Elsevier Science and Technology Books, Amsterdam, The Netherlands.
- Pearl, J. 1995. Causal diagrams for empirical research. *Biometrika* 82(4) 669–688.
- Pearl, Judea. 2003. Statistics and causal inference: A review. *Test* 12(2) 281–345.
- Popper, K. R. 1950. Indeterminism in quantum physics and in classical physics. Part I. *British J. Philos. Sci.* 1(2) 117–133.
- Popper, K. R. 2002. *The Logic of Scientific Discovery*. Routledge, London, UK.
- Pratt, J. W., R. Schlaifer. 1988. On the interpretation and observation of laws. *J. Econometrics* 39(1-2) 23–52.
- Reichenbach, H. 2000. *Direction of Time*. Dover Publications Inc., New York.
- Savage, L. J. 1972. *The Foundations of Statistics*. Dover Publications Inc., New York.
- Shugan, S. M. 2006. Errors in the variables, unobserved heterogeneity, and other ways of hiding statistical error. *Marketing Sci.* 25(3) 203–216.
- Sorensen, R. A. 1992. *Thought Experiments*. Oxford University Press, Oxford, UK.
- Spirtes, P., C. Glymour, R. Scheines. 2001. *Causation, Prediction, and Search. Adaptive Computation and Machine Learning Series*. MIT Press, Cambridge, MA.
- Wason, P. C. 1968. Reasoning about a rule. *Quart. J. Experiment. Psych.* 20(3) 273–281.
- Will, C. M. 1993. *Theory and Experiment in Gravitational Physics*. Cambridge University Press, Cambridge, UK.
- Xie, J., S. M. Shugan. 2001. Electronic tickets, smart cards, and online prepayments: When and how to advance sell. *Marketing Sci.* 20(3) 219–243.
- Zellner, Arnold. 1988. Causality and causal laws in economics. *J. Econometrics* 39(1-2) 7–21.