Selected Paper
Number 62

Backward and Forward Thinking in Decision Making

Hillel J. Einhorn and
Robin M. Hogarth

Graduate School
of Business

The University
of Chicago

Hillel J. Einhorn, professor of behavioral science, founded the Center for Decision Research in 1977 and served as the center's director until 1983, when Robin Hogarth accepted the position of director. A member of the faculty of the Graduate School of Business since 1969, Einhorn came to Chicago the year he received his Ph.D. in industrial psychology from Wayne State University. Both his bachelors and his master's degrees are from Brooklyn College of the City University of New York. He has been a visiting professor at Carnegie-Mellon University and the Hebrew University of Jerusalem. Grants from the National Institute of Mental Health, the Spencer Foundation, the Illinois Department of Mental Health and Developmental Disabilities, and the Office of Naval Research have permitted Einhorn to pursue his research on such subjects as uncertainty, preference reversals, and ambiguity and choice. From 1970 to 1985, Einhorn contributed close to forty articles to a variety of publications. Many of these articles were written in collaboration with Robin Hogarth.

Robin M. Hogarth, professor of behavioral science and director of the Center for Decision Research, received his Ph.D. from the Graduate School of Business in 1972. After appointments at INSEAD (The European Institute of Business Administration) and the London Business School, he returned to Chicago as a faculty member in 1979. Hogatth is an associate editor of ***Management Science*** and the ***International Journal of Forecasting.*** His research has focused mainly on aspects of individual decision making and, in particular, on the assessment of uncertainty. While at INSEAD, he also headed a large study on the effects of management training. He is the author of three books and has published articles in many leading scholarly journals. Together with Einhorn, he is joint principal investigator for projects funded by the Office of Naval Research.

Backward and Forward Thinking in Decision Making

On a recent airing of public television's "Wall Street Week," Louis Rukeyser asked Milton Friedman what he would do about the Federal Reserve Board, an object of Friedman's criticism for many years. Friedman's response was immediate: "I would get rid of them." Surprised, Rukeyser pressed further: "What would you use to replace them!" "A computer," responded Friedman. He then went on to explain that the money supply should be set by consistently applying a simple rule or algorithm. This method, he argued, would provide for more stability and certainty in determining economic policy. Whatever the merits of Friedman's proposal, it strikes at the heart of a conflict many managers experience in making decisions. Namely, should decisions follow some analytically derived principle or rule, or should one simply trust judgment? Since suggesting the replacement of the Federal Reserve Board by a computer invokes reactions of disbelief ("Is this a joke?"), judgment is clearly favored. In this paper, however, we argue that good decision making requires both judgments and algorithms. Specifically, we argue that judgments are most useful in diagnosis (thinking backwards), while algorithms are most useful for predictions and forecasts (thinking forwards). Unfortunately, awareness of this division of labor is rare in our daily lives since we continually shift between forward and backward thinking. As stated by Kierkegaard, "Life is lived forwards, but it is understood backwards."

To understand the nature of forward and backward thinking requires that we think about how we think. This is not easy to do; for example, you have just read the last sentence and thought about it. How did you do it? The difficulty of considering these issues lies in the fact that whereas decisions are the observable end products of complex thought processes, the components of such processes are typically hidden from view.

Going Forwards by Going Backwards

Decision making is a forward-looking activity that depends critically on being able to predict future events. However, when making predictions about the future, one first looks backwards in time to understand the determinants of one's present position. Backward thinking is useful for finding a metaphor, model, or theory that will help in looking forwards. Moreover, success in backward thinking depends on skills of causal or diagnostic reasoning. These include, for example, identifying relevant causal variables, linking those variables in causal chains or scenarios, assessing the strength of those chains, and evaluating alternative explanations. Hence, reasoning about the past involves cognitive activities such as recognizing similarities, imagining how things might have evolved differently, synthesizing information from different sources, and drawing on experience via memory. Moreover, this process is largely intuitive in nature. Situations are appraised quickly, often without conscious awareness of the underlying process. As stated by Michael Polanyi, "We know more than we can tell."

Ironically, this often leads to great confidence in judgments that we cannot explicitly justify to ourselves or to others.

The nature of forward reasoning, on the other hand, is quite different. Having identified relevant factors within a causal model of the situation, the task is to weight and combine the variables in order to make a prediction. This requires the use of a rule or strategy to integrate multiple pieces of information according to their predictive accuracy into a single forecast or prediction. Thus, a critical difference between backward and forward reasoning is that the former inevitably depends on judgment while the latter requires an algorithm to combine information into an accurate prediction. We will consider backward thinking in more detail before discussing predictive judgments.

On Backward Thinking

Backward thinking involves three interrelated phases: (1) finding relevant variables, (2) linking variables into causal chains or scenarios, and (3) assessing the plausibility of such chains or scenarios. We first consider how people do these tasks before making suggestions as to how to do them better. To illustrate the nature of backward thinking, consider the following situation.

> Imagine that you lived several thousand years ago and belonged to a tribe of methodologically sophisticated cave-dwellers. Your methodological sophistication is such that you have available to you all present day means of the methodological arsenal-details of the principles of deductive logic, probabihty theory, access to computational equipment, etc. However, your level of substantive knowledge lags several thousand years behind your methodological sophistication. In particular, you have little knowledge about phys-

ics, chemistry or biology. In recent years, your tribe has noted an alarming decrease in its birth-rare. Furthermore, the tribe's statistician estimates that unless the trend is shortly reversed, extinction is a real possibility. The tribe's chief has accordingly launched an urgent project to determine the cause of birth. You are a member of the project team and have been assured that all means, including various forms of experimentation with human subjects, will be permitted to resolve this crucial issue. (Einhorn & Hogarth, *Journal of Forecasting.* 1982, p. 23)

What is a "relevant" causal factor? Much psychological research indicates that processes of perception and judgment are sensitive to differences or deviations from present states and reference points. Therefore, diagnostic curiosity is usually triggered by noting that something is abnormal or unusual. Research has shown, for example, that unexpected events-such as a win by an underdog, a loss by a favored team, more or less profits than expected-induce people to generate causal explanations. Rarely does one seek the cause of why one feels "average," why traffic flowed normally, or why some accident is typical. Note that our cave dwellers are concerned because the decline in the birth rate is unusual and remedial action is needed.

In searching for the cause of some effect, people usually direct their attention to prior deviations or abnormal states of comparable length and strength. Our cave dwellers, for example, probably asked themselves what unusual event (or events) preceded the decline in births. Moreover, if the effect is large (that is, of substantial duration and/or magnitude), the suspected cause is thought to be of comparable size. There is a strong notion that similar causes have similar effects. Indeed, John Stuart Mill noted that this deeply rooted belief "not only reigned supreme in the ancient world, but still possesses almost undisputed dominion over many of the most cultivated minds." Mill thought that such a belief was erroneous, and many cases exist in which similarity has been misleading (for example, the primitive belief that fowl excrement can cure ringworm because the latter looks like the former). On the other hand, it is difficult to imagine how one could search for variables without some notion of similarity.

Of particular importance is the non-literal use of similarity through analogy and metaphor. In providing models of phenomena, analogies and metaphors direct attention to particular factors and processes. Moreover, their use engages prior knowledge so that explanations for new or poorly understood phenomena can be integrated with what is already known. For example, in trying to understand how the brain works, one could consider it as a computer, a muscle, or a sponge. Note how the various metaphors direct attention to different, but known, variables and processes. A computer metaphor suggests informational input, retrieval, and computational processes; a muscle model suggests that processes are strengthened and weakened with amount of use and that thinking can be a strain; a sponge metaphor suggests a passive absorption of information. The choice of a particular metaphor is crucial since it directs attention to a limited set of variables, thereby excluding others.

Cues-to-causality and causal chains. The search for causally relevant variables goes hand in hand with the assessment of the strength of causal candidates. Both involve what we call "cues-to-causality," which are probabilistic indicators that suggest the presence of causal relations. These cues are: temporal *order* (causes precede effects in time), *contiguity in time and* space (one typically seeks causes that are close to the effect in time and space), *correlation* (causes tend to covary with effects); and *similarity* of cause and effect (including physical similarity, analogy and metaphor, and congruity of length and strength between cause and effect). Although the presence of "cues-to-causality" does not prove the existence of causal relations, the cues do serve two important functions. First, they indicate likely directions in which to seek relevant variables. Second, they constrain the number of causal scenarios or chains that can be constructed to link causal candidates to their supposed effects. To illustrate, how likely is it that sunspots cause price changes on the stock market? (Before dismissing this as an absurd hypothesis, note that the eminent nineteenth-century economist Jevons believed in a causal relation between sunspots and business cycles.)

In order to link sunspots to stock prices, one has to construct a causal chain or scenario that meets various constraints. For the sake of discussion, assume that sunspots occur before price changes (temporal order is preserved), that there is a positive correlation between sunspot activity and price changes (correlation is significant), and that price changes occur six months after sunspot activity (low contiguity in time). To bridge both the time lag and distance gap between sunspots and price changes, one needs to construct a causal chain linking these events. If this cannot be done, no causal relation will be seen. On the other hand, consider the following chain: sunspots affect weather conditions, which affect agricultural production, which affect economic conditions, which affect profits, which affect stock prices. Note that the cues-to-causality constrain the possible causal chains that can be imagined. This is most obvious in considering the cue of temporal order. If X does not precede Y in time, it cannot be a cause of Y. The linkage is also constrained by both the cues of contiguity in time and space, and by congruity (length and strength of cause and effect). Consider the role of contiguity in time and space in the above example. In order to link these events at all, the spatial and temporal gaps must be bridged by positing a change in weather, etc. Imagine, however, that price changes occur immediately after sunspot activity, rather than six months later. The closeness in time between the two events precludes the weather/economic conditions scenario since that scenario requires a time delay between cause and effect. Hence, to link sunspots and price changes, one would have to construct another scenario that is consistent with high contiguity in time. Similarly, imagine taking up smoking and getting lung cancer the next week. It seems highly unlikely that smoking is the cause since a causal chain with high contiguity in time is not easy to construct.

Now consider incongruity between cause and effect; that is, small causes/big effects or big causes/small effects. To account for this, the causal chain must involve some type of "amplification" in the first case and "dampening" in the second. For example, consider the incongruity between germs and illness, as exemplified in the germ theory of disease. When Pasteur advanced this theory it must have seemed incredible to his contemporaries that invisible creatures caused death, plagues, and so on. Without knowledge of how germs enter the body, multiply, disseminate through the system, and are communicated to others, there was no causal chain that could amplify such small causes to produce the large effects.

Finally, the number and strength of alternative explanations play a key role in causal thinking. Alternative explanations reduce the strength of any particular causal candidate or scenario. Indeed, it has been suggested that the primitive notion of a cause involves asking oneself the question, "Would Y have occurred if X had not?" The greater the number and strength of alternative explanations underlying a "yes" answer, the lower the causal relevance of X for Y.

The description of the processes and factors that comprise backward thinking suggests several ways to improve one's diagnostic skills.

(1) Use multiple metaphors. A significant feature of causal/diagnostic thinking is the remarkable speed and fluency with which people generate explanations and accommodate discrepant facts into expanded hypotheses. Since backward thinking is so fluent, a single metaphor can often be generated quickly. However, since metaphors are imperfect, relying on a single image can lead to errors. As stated by the general semanticists, "The map is not the territory." We therefore recommend the method of multiple metaphors as a way of guarding against the premature adoption of a single model.

Instead of focusing on a single metaphor, experiment with several. Consider, for example, how one might think about complex organizations such as graduate schools of business. One could think of these institutions as finishing schools (where students mature appropriately prior to corporate life), as military academies (preparing students for economic warfare), as churches (providing indoctrination in particular "schools of thought"), as diploma mills (providing a certification function), as medical schools (preparing students to become interns and subsequently professionals), or perhaps as job shops (tooling students to perform specific tasks). Each of these metaphors leads to considering quite different variables and issues. Therefore, while no single model is correct, each directs attention to different factors, thereby providing a more complete picture of the phenomenon. Various metaphors could, moreover, be used to characterize differences between schools. Indeed, some consulting firms use animal metaphors, for example, "cash cows" and 'lean dogs," in discussing how to manage the different parts of complex business enterprises. It is interesting to speculate how their advice might change were different metaphors adopted.

(2) Avoid reliance on a single cue. When a single cue is used to infer causality, serious errors in diagnosis can occur. To illustrate, consider the following extension of our caveman scenario. Imagine that you notice that children are physically similar to men and women who live together. Moreover, this leads you to formulate the hypothesis that sexual intercourse causes pregnancy. Note that this hypothesis is not obvious since contiguity in time between cause and effect is

low, knowledge of a causal chain linking intercourse and pregnancy is missing, the cause is not sufficient for the effect, and many factors correlated with intercourse are difficult to rule out (for example, sitting under palm trees in full moonlight; we note in this regard a letter to "Dear Abby" in which a teenage girl asked if holding hands with her boyfriend could get her pregnant!). In any event, imagine that you decided to test your hypothesis with an experiment. From a sample of 200 couples, 100 are assigned at random to an intercourse condition, and 100 to a nonintercourse condition. After some time, you obtain the results shown in the following table.

		Pregnant: Yes	Pregnant: No	
Intercourse	Yes	20	80	100
	No	5	95	100
		25	175	200

In light of modern knowledge, readers may be surprised to note that the pregnant-no intercourse cell in the table contains 5 entries. However, unbeknownst to our cavemen, this represents measurement error in the data (due to faulty memory of respondents, lying, etc.). Be that as it may, you calculate the correlation implicit in the table and find that it is .34. Since this correlation is at best modest, you conclude that intercourse is not a major factor in causing pregnancy.

There are several implications of this example. First, whereas elementary statistics books correctly warn that correlation does not imply causation (the case of so-called “spurious correlation”), it is also the case that causation does not imply correlation (a case we have named “causalation”). Second, the reliance on any single cue-to-causality can lead to substantial errors of judgment. Hence, statistics and other forms of logic that rely on single measures can take one only so far. Indeed, the importance of multiple cues in diagnostic reasoning is crucial.

(3) Use the cues-to-causality creatively. A great benefit of the cues-to-causality is that they structure our perceptions, thereby helping us interpret ambiguous information. However, it is important to realize that the human ability to interpret information unambiguously comes at the cost of novelty and originality. The cues direct attention to what is “obvious,” thereby reducing alternative interpretations that can lead to innovation and creativity. Indeed, the hallmark of creative insights is that they are accompanied by surprise. When novel interpretations are appropriate to the task, such new understanding subsequently appears both simple and self-evident. This suggests that one way to facilitate creative thinking might be to go specifically against the cues. For example, when dealing with a complex outcome, search for a dissimilar and simple causal candidate rather than a similar and complex one.

(4) Assess causal chains. While the construction of a causal chain is necessary to link potential causes and effects, the plausibility or strength of the chain may vary greatly. For example, the chain between sunspots and stock prices is weak because there are many links and each is uncertain. In

fact, if the strength of a chain is a multiplicative function of the strength of its individual links, then any chain is only as strong as its weakest link; and longer chains are generally weaker than shorter ones. Research has shown that people do not always grasp these notions. In particular, complex scenarios in which outcomes are described with much detail are often regarded as more coherent and thus mote likely than simple ones. Hence, one must be careful to evaluate chains according to the number and strength of their individual links. This is particularly important in planning where the success of each stage in a sequence of activities depends on the success of prior stages.

(5) Generate and test alternative explanations. As noted above, people exhibit great fluency in diagnostic thinking. However, a cost of this fluency is that it can lead to adopting erroneous beliefs in the form of myths and superstitions. While it is true that these become discredited over time, the interim periods during which they hold sway can be lengthy. The history of medicine, for example, is full of such theories as bloodletting that were once popular and even thought to be scientifically sound. It takes little imagination to realize that the same fate could await many of our most cherished theories about economics and business.

To guard against myths and superstitions, we recommend the use of either real or thought experiments. Note that the posing and answering of the counterfactual question may involve doing a real experiment or what is called a “thought” experiment. In the former, one compares the effect of X on Y with that of not-X (the control group) on Y. In this way, the counterfactual question is answered. To assess the effectiveness of personnel selection or advertising, for example, one could conduct experiments by randomly selecting employees or stopping advertising completely. While such experiments are typically infeasible, partial experiments could provide much useful information. Thus, one could randomly admit a small percentage of candidates, and advertising could be stopped in selected time periods or areas.

When real experiments are not possible, one can nevertheless imagine situations in which the effect occurs without the suspected cause. In such thought experiments, the construction of "possible worlds" and imaginary scenarios is crucial for judging causal significance. Consider, for example, whether a particular advertising campaign caused an observed increase in sales. By asking and then answering the question, "Would the sales increase have occurred without the advertising campaign?", a better estimate of the causal relation between sales and advertising can be obtained. Good thought experiments include a second question, similar in spirit, "Could sales not increase if we advertise?" Systematically posing these questions can increase the power of thought experiments by generating information analogous to that available from real experiments.

Algorithms in Forward Reasoning

The literature on people's ability to engage in accurate forward thinking presents a uniformly depressing picture. Accuracy in prediction is typically low and inferior to predictions from even simple statistical models. These statements are based on systematic evidence that has been accumulating for over forty years. Moreover, it has been gathered in such diverse areas as clinical psychology, medicine, scholastic performance, marriage counseling, bank-lending operations, production scheduling, and various economic activities.

The typical reaction to these results is surprise or disbelief. After all, compared to human judgment, the disadvantages of statistical models are obvious. First, a model is limited by the number of variables or factors it can incorporate. Being incomplete, it must make errors in prediction. On the other hand, human judgment can consider each case as unique, employing as many factors as are necessary to predict that case. Second, whereas the model is static, the human judge learns from experience, presumably improving predictive performance over time. Third, even if models are better than humans in predicting, the difference is small. Thus, the marginal increase in predictability is not worth the cost of building a model. We now consider each of these arguments.

(1) Models must make errors. Use of a formal model implies trade-offs. Foremost among these is the fact that a model will make errors since it is an abstraction that cannot possibly capture the full richness of the relations between variables. Human judgment, on the other hand, can sometimes capitalize on idiosyncratic features that are difficult to model. However, since unaided judgment also results in errors, it is important to consider the nature of both types of errors. If we assume that the goal of prediction is to maximize predictive success over many cases, the predictions of humans and models differ in that the latter are perfectly

consistent; that is, models don't get bored, tired, or distracted. Thus, models are not subject to inconsistency and random error. On the other hand, models will lead to errors because they are incomplete. The important question therefore centers on which method leads to less overall error.

How can the acceptance of error, via the use of a formal model, lead to less overall error? The results of psychological experiments on probability learning illustrate how. In these studies, either a red or a green light is illuminated on each of a number of trials, and subjects are asked to predict which light will go on. If the prediction is correct, subjects are given a cash reward; if the prediction is wrong, there is no reward. The appearance of red and green lights is generated by a random process such that the red light will appear a given percentage of times, say 60%, and the green light 40%. Subjects are not told about the characteristics of this process but have the opportunity of learning about it by participating in the task. The major result of these experiments is something called "probability matching"; namely, subjects respond to the lights in the same proportion as they occur. In this case, subjects predict red about 60% of the time, and green 40%. The expected payoff for such a strategy can be calculated as follows: since subjects predict red on 60% of the trials and red occurs on 60%, they will be correct (and receive the payoff) on 36% of the trials. Similarly for green: subjects predict green on 40% of the trials and green appears 40% of the time. Hence, 16% of the trials will be correctly predicted. Therefore, the overall accuracy of predictions will be 52% (that is, 36% + 16%). Now consider how well subjects would do by using the simple rule of always predicting the more likely color. Such a strategy explicitly accepts error; however, it also leads to 60% correct predictions. Since 60% is greater than 52%, subjects would make more money if they accepted error and consistently used a simple rule. Indeed, such a rule maximizes their wealth in this situation. However, most subjects try to predict perfectly and engage in a futile attempt to discern patterns that are diagnostic of the (nonexistent) rule they believe determines the onset of the lights. We note the analogy to the stock market without further comment.

(2) Models are static. The criticism that models are necessarily static is not true. First, models can and should be updated as new information is acquired. Second, models are now being developed that learn from the outcomes of predicted events. While this work is still in its early stages, it suggests that models may be able to learn from experience. Third, the literature on people's ability to learn from feedback in probabilistic environments is not encouraging. Part of the difficulty in learning occurs when predictive judgments are made in order to take actions. In this case, outcomes only provide ambiguous feedback regarding the quality of predictions. For example, if the president takes strong measures to counteract the prediction of an economic slowdown, consider the difficulties of learning from the various outcomes. Imagine the outcome of "no recession." This could result from an incorrect forecast and an ineffective action or from an accurate forecast accompanied by a highly effective action. Similarly, a recession could indicate an accurate forecast with an ineffec-

tive action or an inaccurate forecast coupled with an action that caused the malady it was designed to prevent. Whereas some actions are taken to counteract the prediction of undesirable events, other actions contribute to the very outcomes predicted. For example, rumors (that is, predictions) that a bank will fail may well cause the failure if people act on the rumors by withdrawing their deposits. Hence, to learn about our predictive ability requires separating the quality of predictions from the effects of actions based on those very predictions.

(3) Models are not worth the cost. As a general statement, it is impossible to evaluate the claim that the marginal increase in predictability from using models does not outweigh the extra costs of building them. However, if enough predictions are made, even small increases in accuracy can produce large benefits. For example, in the late 1970s, AT&T carried out a study to determine the characteristics of good and bad credit risks. The results of this study were incorporated in a set of decision rules that were subsequently implemented to determine which new customers should be required to provide deposits. In developing these rules, AT&T went through a period in which credit was granted not only to customers who would have previously been classified as good risks but also to those who would have been in the bad-risk category. Thus, the rules derived from the study were validated across the whole range of customer characteristics. The results of implementing the decision rules led to an estimated annual reduction of $137 million in bad debts. Whereas no data have been made public on the cost of creating and maintaining the model, it is difficult

to believe that these could be significant relative to the estimated savings.

While many phenomena we seek to predict are inherently complex, the rules for forward reasoning need not match this complexity. On the contrary, many successful applications have involved simple combinations of just a few variables. Sometimes these rules have been derived from modeling an expert's past judgments, sometimes by simply averaging past decisions, and sometimes by just aggregating a few variables thought to be relevant.

Conclusion

We have argued that decision making involves two types of thought processes, backward and forward reasoning, which are confounded in everyday experience. Moreover, lack of awareness of this confounding can have serious negative consequences. For example, while we have argued that forward reasoning is best served by the use of explicit rules or algorithms, causal notions can exert undue influence on predictive judgments. In particular, when people take action in situations where outcomes are generated by random processes, they are sometimes subject to what psychologists call "illusions of control." Thus, people tend to believe that lottery tickets they personally select have a greater chance of winning than those selected for them by lottery administrators. Similarly, in more complex situations the cognitive activities of planning and forecasting can lead to illusions of control and overconfidence by underestimating the importance of random factors in the environment. Since illusions of control are conceptually identical to superstitious behavior, the joint presence of randomness and cues-to-causality make some superstitious behavior inevitable. Nevertheless, one should be on guard to minimize such behavior, and we strongly recommend maintaining a distrustful attitude toward all undocumented claims of predictive accuracy, whether derived from models, experts, or both. A simple way to remember this advice is to consider the "seersucker theory" of prediction. This theory has only one axiom: For every seer there is a sucker.

The need to distinguish the roles of backward and forward inference is well captured by the following statement made in an important paper on how to improve predictive ability, "...the whole trick is to decide what variables to look at and then know how to add." The "trick," however, is a difficult one that requires complex backward thinking. Moreover, it is easy to underestimate the remarkable fluency with which people reason causally and perform the backward inferences necessary to develop models and metaphors in the first place. This can be illustrated by the difficulties computer scientists face in trying to build programs that simulate the understanding process by means of "artificial intelligence." A recent example concerns a program written to simulate the comprehension of newspaper headlines. The program was provided with background knowledge and a set of rules to rewrite the headlines. One such headline was, "World shaken. Pope shot." The computer interpreted this as, "Earthquake in Italy. One dead."

Finally, although the psychological study of judgment and decision making has clarified the distinction between backward and forward reasoning, it seems appropriate to conclude with a metaphor that captures the interdependence of both modes of thought. In Roman mythology, the god Janus was the porter of heaven and the guardian deity of gates. As such, he was commonly represented as having a head with two faces-one facing forwards and the other backwards. What better way to represent the psychology of decision making?