Hindsight and Causality

David Wasserman
Institute for Philosophy and Public Policy
University of Maryland

Richard O. Lempert
Department of Sociology, University of Michigan
University of Michigan Law School

Reid Hastie
Center for Research on Judgment and Policy
University of Colorado, Boulder

When people know how an event turned out, they are usually unable to reproduce the judgments they would have made without outcome knowledge. Furthermore, they are unaware of their inability to recapture their pre-outcome state of mind. This tendency to overestimate what they would have known without the outcome knowledge is called "hindsight." An experiment explored the moderating effects of the type of cause to which the outcome was attributed on the magnitude of the hindsight effect. When the outcome was attributed to unforeseeable "chance" factors, such as an unexpected storm or an earthquake, the hindsight effect was virtually eliminated. When no causal attribution was provided or when a plausible "deterministic" cause (human skill or lack of skill) was cited, subjects' judgments showed sizable hindsight effects. These findings are interpreted as supporting Fischhoff's "creeping determinism" hypothesis and as providing evidence that the hindsight effect is a by-product of adaptive learning from feedback.

Hindsight is prevalent in everyday judgments. Events in the past usually appear simple, comprehensible, and predictable in comparison with events in the future. Everyone has had the experience of believing that he or she "knew all along" the outcome of a legal trial, business investment, political election, or football game. And everyone has reacted with skepticism to similar claims from someone else.

Fischhoff (1975) introduced a laboratory analogue of natural hindsight judgment situations that set the basic laboratory paradigm for subsequent experiments studying the hindsight phenomenon. In Fischhoff's paradigm, subjects first received information about a target event—for example, an obscure historical event such as the 19th-century wars between the British and the Gurkha of Nepal. Second, some subjects were told that a particular outcome had "actually occurred" and other subjects were not. Third, all subjects were asked to estimate the probability of each outcome, and the informed subjects were asked to make their estimates *as if they had not received the outcome information.* The "hindsight bias" was operationalized as the tendency for some subjects with outcome knowledge (hindsight) to claim that they would have estimated a probability of occurrence for the reported outcome that was higher than they would have estimated in foresight (without the outcome information).

It is important to note that the hindsight bias does not refer to all retrospective increases in the probabilities assigned to events. The bias is a projection of new knowledge into the past accompanied by a denial that the outcome information has influenced judgments. Thus, subjects who learn of an outcome in a hindsight experiment and claim that they "would have known it all along" are fooling themselves.

Authors' Note: Address correspondence to Reid Hastie, Psychology Department, Campus Box 345, University of Colorado, Boulder, CO 80309-0345.

PSPB, Vol. 17 No. 1, February 1991 30-35

A number of theoretical explanations have been provided for the hindsight effect by Fischhoff and others. The simplest explanations cite the potential experimental demand characteristics of the method, attributing the effect to subjects' efforts to appear cooperative or perspicacious. These explanations do not account for most of the hindsight effects reported in the experimental literature (see Wood, 1978).

The most strongly supported explanation for the effect is the process that Fischhoff labeled "creeping determinism" in his seminal article: "Upon receipt of outcome knowledge judges immediately assimilate it with what they already know about the event in question. In other words, the retrospective judge attempts to make sense, or a coherent whole, out of all that he knows about the event" (1975, p. 297). For example, after reading Fischhoff's paragraph on the British-Gurkha war and learning from the experimenter of a British victory, the subject might assimilate the outcome to the case information by adding semantic links signifying causal relations between events in the case information and the outcome. In essence, the subject would "rewrite the story" of the war so that its beginning and middle would be connected causally to its end (e.g., Pennington & Hastie, 1986, 1987, 1988; Trabasso & van den Broek, 1985).

The "creeping determinism" hypothesis is consistent with more of the results in the hindsight literature than any other explanation that has been proposed (Hawkins & Hastie, 1990). Furthermore, the "creeping determinism" account is consistent with cognitive process accounts that have been invoked to explain several related judgment and memory phenomena: plausibility judgments as a basis to respond on fact retrieval and recognition memory tests (Reder, 1987); "perseverance" effects in social judgment (Ross, Lepper, & Hubbard, 1975); the role of explanation in predictions of social events (Sherman, Skov, Hervitz, & Stock, 1981); and subjects' inability to discount information once it has been integrated into an impression of another person (Schul & Burnstein, 1985; Wyer & Budesheim, 1987) or into a legal judgment (e.g., Casper, Benedict, & Kelly, 1988; Hans & Doob, 1975).

Fischhoff's label "creeping determinism" suggests that causal ("deterministic") relationships may play an especially prominent role in the inferences that subjects make after receiving information about the outcome of the event, inferences that underlie the hindsight biases. There is liberal evidence in the text comprehension literature that causal inferences are a particularly common on-line inference in the minds of subjects comprehending text describing social events (Fletcher & Bloom, 1988; Myers, Shinjo, & Duffy, 1987; Trabasso & van den Broek, 1985).

The present experiment was designed to assess the significance of deterministic, causal relationships in the production of hindsight effects in judgment. We attempted to vary the degree to which a subject considered deterministic factors when thinking about the reported outcome of an event to be judged. Our design included a replication of Fischhoff's control conditions in which no outcome feedback was provided (foresight condition) and a replication of Fischhoff's basic hindsight condition in which an outcome was given to the subjects but with no specific information about its causes. In addition, our design included experimental conditions designed to manipulate the availability of specific deterministic or "chance" causal explanations for the outcome of the event. For example, in one of our sets of materials based on Fischhoff's British-Gurkha historical scenario, outcomes were provided to subjects (either "British victory" or "British defeat") accompanied by a statement of the primary cause of the outcome: either a "deterministic" attribution (something about the armed forces of the winning side) or a "chance" attribution (an unexpected and unseasonable rainstorm).

Our hypothesis was that subjects told the outcome of the battle, but nothing about its causes, would "rewrite" their mental representations of the case information to make sense of the outcome, looking for causal, deterministic links between the outcome and factors in the case materials. We thought that much the same process would occur for subjects given our "deterministic" feedback along with the outcome. However, when subjects were told that the outcome was due to a "chance" event, we thought their tendency to rewrite the causal relations in the scenario would be inhibited. Under these conditions we expected that the hindsight effect, changes in the retrospective probability ratings in the direction of the reported outcome, would be diminished or eliminated. Thus, our hypothesis is based on the assumption that hindsight effects occur when subjects reason about events like Fischhoff's historical scenarios because they elaborate their long-term memories of an event by adding *links* between the outcome and possible causal precursors.

Our operationalization of the "deterministic" versus "chance" distinction was motivated by the notion of "foreseeability" that is central in tort law. Our chance explanations cited causal events that were unforeseeable given information in the case materials (e.g., unseasonable monsoon rains; an earthquake). In contrast, our

"deterministic" explanations cited plausible effects of factors whose relevance to the outcome would be expected by our subjects (e.g., troop discipline; scientific know-how).

METHOD

Overview

Our basic experimental method followed Fischhoff (1975) quite closely. Subjects were presented written case materials describing historical events. Some subjects were provided with information about the outcomes, some with information about the outcomes plus the causes of the outcomes, and some with no outcome or cause information. All subjects were asked to make ratings of the probabilities that they *would have assigned* to the possible outcomes of the cases, as if they had not heard any information about the actual outcomes. Subjects also made ratings of the importance of various factors mentioned in the case materials as cues to judging the outcomes. These importance ratings served two purposes: first, as a check on the manipulation of causal information for those subjects who were given explanations of the outcomes; second, as converging evidence for the "creeping determinism hypothesis" that subjects had revised their mental representations of information in the case on receipt of outcome information.

Subjects

Subjects were 278 undergraduate students at Northwestern University who participated in the study in partial fulfillment of the requirements of an introductory psychology course.

Design

Each subject read material from one of two scenarios, the first depicting an impending battle between a British regiment and a band of Gurkha rebels in Nepal in 1814, the second describing a gold prospecting venture undertaken by a pair of geologists in the 1970s. Subjects received either no outcome information, only outcome information, or outcome information coupled with a "deterministic" or a "chance" explanation for the outcome. Subjects receiving outcome information were told of either a British or a Gurkha victory or of the success or failure of the prospecting venture. These manipulations yielded a seven-cell design for each of the two scenarios, for a total of 14 experimental conditions.

Materials

Typewritten descriptions, approximately 250 words in length, were constructed for each scenario. The British-Gurkha scenario was based on the materials used by Fischhoff (1975), and the gold prospecting scenario was devised by the authors. Each scenario included seven identifiable factors that might serve as determinants for the outcome that would follow the scenario events. The British victory subjects were told the British won either as a result of an unexpected monsoon ("chance" explanation) or because of the discipline of the British troops ("deterministic" explanation). British defeat was attributed to the opposite consequences of the chance monsoon or to the lack of discipline of the British troops. The second scenario, describing a gold prospecting venture in Alaska, led to a success or failure outcome that was attributed either to an unexpected earthquake ("chance" explanation) or to the competence of the young geologists engaged in the enterprise ("deterministic" explanation).

Procedure

The experimental session lasted about 10 min; subjects participated in a large-group testing session during the meeting of an introductory psychology course. Each subject received an experimental booklet with a cover page of instructions. Subjects were told they would be reading an account of a historical event and would be asked to make judgments about that event. They then read one of the two scenarios, received outcome information according to their experimental condition, and made a series of judgments about the case materials.

Subjects were asked to indicate which outcome had been presented, in order to check that their interpretations of the presented outcome agreed with the experimenters' (all subjects responded correctly). Then the subjects were instructed:

> Some subjects are asked to read the accounts of this battle [prospecting venture] but are NOT told the outcome. Your task is to put yourself in their shoes and to attempt to judge the likelihood of each of the following outcomes by writing a probability value of from 0 to 100% next to each outcome. The TOTAL of all your probability judgments should equal 100%.

The final page of the experimental booklet consisted of a list for each domain of seven factors that had been presented in the scenario. Subjects were asked to rate each factor on a 7-point scale to indicate how relevant or important they considered the factor to be in determining the given outcome. The seven factors referred to background conditions reported in each scenario. For example, items for the British-Gurkha battle included terrain, precipitation and weather conditions, leadership and tactics, and troop discipline. Subjects paced

TABLE 1: Mean Estimates of Probability of British Victory or Gold-Mining Success (on a 100-Point Rating Scale)

Type of Explanation	British-Gurkha Battle		Gold Prospecting Venture	
	British Victory	British Defeat	Success	Failure
"Deterministic"	57	18	58	45
None	50	21	54	43
"Chance"	40	33	37	35
No outcome disclosed ("foresight")	43		45	

themselves through the experiment, and no subject took longer than 15 min to complete the task.

RESULTS

Retrospective Estimates of Probability

The mean estimates of probability of success for the battle and gold prospecting scenarios are presented in Table 1. Subjects who were given no outcome information predicted British victory or success in the prospecting venture at close to the midpoints on the scale, 43 and 45 respectively. This demonstrates that our scenarios were not skewed to favor one outcome or another.

An analysis of variance was conducted on the 12-cell design, excluding the no-outcome-feedback conditions, with type of causal explanation provided ("deterministic," none, "chance") and type of outcome (success, failure) and scenario (British-Gurkha, prospecting venture) as the independent variables and subjects' probability ratings on the 0 to 100 scale as the dependent variable. There was a highly significant effect for the outcome variable (success vs. failure), $F(1, 201) = 31.03$, $p < .0001$, $MS_e = 451$. More important, for the predictions we made concerning the moderating effects of explanation type (deterministic, none, chance), we obtained a significant interaction between outcome and explanation type, $F(2, 201) = 3.78$, $p < .02$, $MS_e = 451$. Inspection of the means in Table 1 shows that, as predicted, we obtained definite and approximately equal-magnitude outcome feedback (hindsight) effects under conditions where subjects were provided with deterministic explanations for the outcome or no explanation for the outcome. In contrast, the difference between success and failure feedback conditions when accompanied by a chance explanation was sharply diminished (success-failure differences of 7 and 2 out of a possible 100 scale points, for the two scenarios respectively). (Contrasts with one degree of freedom comparing the hindsight conditions with the "no outcome disclosed" control conditions were significant for the "deterministic" explanation and no-explanation conditions [success outcome: $F(1, 276) = 7.80$, $MS_e = 537$, $p < .01$; failure outcome: $F(1, 276) = 8.38$, $p < .01$, $MS_e = 536$] but not for the "chance" explanation conditions [success outcome: $F(1, 276) = 1.23$, n.s., $MS_e = 550$; failure outcome: $F(1, 276) = 2.72$, n.s., $MS_e = 547$].)

There is one somewhat puzzling comparison among means in our results for the gold prospecting scenario. Subjects receiving no outcome feedback gave a mean rating of 45 for success of the prospecting venture, whereas subjects' ratings in hindsight, after they were told the venture had failed, were 45, 43, and 35 in the "deterministic," no-explanation, and "chance" explanation conditions; none of these means is significantly different from the mean found when no outcome was disclosed. The puzzling result is that failure feedback subjects did *not* show a hindsight effect such that their ratings of the probability of success were depressed below the ratings by subjects who had received no outcome feedback. Essentially, this result is a "no hindsight effect" result for failure feedback subjects in the gold prospecting scenario. In contrast, an elevated probability-of-success rating was obtained for the gold prospecting scenario for subjects given success feedback, at least in the "deterministic" and no-explanation treatments (as we had predicted).

Our speculation is that some subjects treat failure as a "nonevent" not requiring an explanation beyond the mere acknowledgment that failure occurred. This result may reflect a "nonevent" status for the failure outcome analogous to the manner in which certain other "nonoccurrence" events have been treated by subjects in other judgment studies. For example, Fischhoff and others have found that informing a subject that an outcome *did not* occur produces small hindsight adjustments of retrospective probabilities, in contrast to substantial adjustments following feedback that an outcome *did* occur (Fischhoff, 1977; Fischhoff & Beyth, 1975; Wood, 1978).

Importance Ratings of Background Factors

The mean ratings of importance of the seven factors cited in each of the two sets of scenario materials are summarized in Tables 2 and 3. We did not state specific predictions concerning the pattern of importance ratings that should result from our manipulations of outcome and causal explanation information. However, there were significant effects of outcome and/or causal explanation factors on 11 out of 14 factors. For the most part these ratings make good sense in terms of our hypothesis: Importance ratings increased for factors that were cited as causal in the explanations provided by the ex-

TABLE 2: Average Rated Importance of Seven Factors as Determinants of the Outcome in the British-Gurkha Scenario

	British Victory			*British Defeat*			
Causal Factor	*Determ.*	*Chance*	*No Expl.*	*Determ.*	*Chance*	*No Expl.*	*No Outcome ("Foresight")*
1. Size of forces	4.4	3.9	4.5	2.6	2.5	3.1	4.4
2. Rough terrain	4.1	4.2	3.8	5.6[a]	4.5	5.6	4.7
3. Temperature	3.9	3.6	3.4	4.6	2.5	5.4	4.1
4. Precipitation and weather	3.3	4.9[a]	2.8	3.9	5.7[a]	4.8	3.4
5. Weapons, war animals	4.6	4.8	5.3	3.9	4.8	4.6	4.5
6. Leadership and tactics	5.6[a]	4.1	5.8	5.4[a]	3.8	5.1	4.6
7. Troop discipline	5.4[a]	4.0	5.7	3.6	3.2	4.6	4.5

NOTE: Ratings could range from 0 to 6.
a. Factor cited in outcome materials in that condition.

TABLE 3: Average Rated Importance of Seven Factors as Determinants of the Outcome in the Gold Prospecting Scenario

	Success			*Failure*			
Causal Factor	*Determ.*	*Chance*	*No Expl.*	*Determ.*	*Chance*	*No Expl.*	*No Outcome ("Foresight")*
1. Industry veterans	2.7	1.6	4.1	3.4	3.0	4.9	4.5
2. Untested scheme	2.9[a]	3.5	4.5	4.6[a]	4.6	4.6	4.6
3. General riskiness	3.0	4.0	2.1	4.4	4.4	3.1	2.6
4. Capital	3.9	4.3	4.1	3.8	4.7	4.7	4.5
5. Geology and weather	3.9	4.0[a]	3.6	3.9	5.3[a]	3.1	4.1
6. Feel for mining	2.5	1.9	2.4	4.3[a]	3.0	2.8	2.8
7. Industry rivals	2.8	2.6	2.8	2.5	3.5	3.0	2.5

NOTE: Ratings could range from 0 to 6.
a. Factor cited in outcome materials in that condition.

perimenter and decreased (after explanation feedback) for factors that were not cited. Furthermore, many ratings of importance shifted with outcome information only, and these shifts were also in reasonable directions. For example, the "size of forces" factor for the British-Gurkha war showed a dramatic significant effect of outcome feedback; importance ratings increased with British victory.

DISCUSSION

Our finding that the magnitude of the hindsight effect depends on the type of explanation that subjects are given for the outcome of an event provides support for the "creeping determinism" hypothesis first posed by Fischhoff (1975). This conclusion is consistent with an interpretation of the hindsight bias as a side effect of adaptive learning from outcome feedback. When one learns of a deterministic, causal explanation for an outcome, it should affect one's subsequent judgments of events in the domain under consideration. However, we would also note that subjects who are given a "chance" explanation for the outcome also learn something important about the judgment of historical events.

The label "bias" does not derive from the fact that our subjects quite reasonably changed their minds about the importance of certain factors that produced outcomes in the domains of our stimulus scenarios. Rather, the label "bias" derives from subjects' tendency to make these changes, to fail to recapture their pre-outcome information orientation, and to be unaware of the change in their model for judgments in the domain. It is interesting that we might label our "chance" explanation subjects as relatively unbiased and accurate in comparison with subjects in the "deterministic" and no-explanation/outcome-feedback conditions. Presumably, the "chance" explanation subjects did not change their generic model for judgments in the colonial wars or gold prospecting domain. The subjects may have even been chastened by feedback that indicated that prediction in the domain could involve unpredictable chance factors.

Our interpretation of the difference between "deterministic" and "chance" explanation conditions is similar to an interpretation of the results of a recent series of experiments reported by Hoch and Loewenstein (1989). Hoch and Loewenstein found that hindsight effect magnitudes were reduced, and "reverse hindsight" effects could even be obtained, when the hindsight feedback was highly surprising to subjects. They conjectured that the experience of surprise elicited by some outcomes led

subjects to rethink the judgment and to render conservative retrospective probability judgments in hindsight. This pattern of results is similar to the results of our experiment, and the Hoch and Loewenstein explanation may apply, especially if we think that our "chance" explanations were more likely to produce surprise reactions than our "deterministic" explanations.

Explanations of our results that invoke hypotheses citing motivational (e.g., experimental demand) factors, anchor and adjustment heuristics, and evidence-sampling biases are not consistent with our results. Subjects in the "deterministic" and "chance" explanation conditions were equally aware of and attentive to the experimenter-provided causal information. However, only subjects provided with "deterministic" feedback or allowed to generate their own causal explanations (in the no explanation treatments) changed their schemata for judgment and therefore showed hindsight biases in their retrospective probability judgments. Thus, our research suggests that the hindsight bias is closely related to other forms of induction and generalization that depend on subjects' inferences about systematic causal factors in the events under consideration.

REFERENCES

Casper, J. D., Benedict, K., & Kelly, J. R. (1988). Cognitions, attitudes, and decision making is search and seizure cases. *Journal of Applied Social Psychology, 18,* 93-113.

Fischhoff, B. (1975). Hindsight ≠ foresight: The effect of outcome knowledge on judgment under uncertainty. *Journal of Experimental Psychology: Human Perception and Performance, 1,* 288-299.

Fischhoff, B. (1977). Perceived informativeness of facts. *Journal of Experimental Psychology: Human Perception and Performance, 3,* 349-358.

Fischhoff, B., & Beyth, R. (1975). "I knew it would happen"—remembered probabilities of once-future things. *Organizational Behavior and Human Performance, 13,* 1-16.

Fletcher, C. R., & Bloom, C. P. (1988). Causal reasoning in the comprehension of simple narrative texts. *Journal of Memory and Language, 27,* 235-244.

Hans, V. P., & Doob, A. N. (1975). Section 12 of the Canada Evidence Act and the deliberations of simulated juries. *Criminal Quarterly, 18,* 235-253.

Hawkins, S. A., & Hastie, R. (1990). Hindsight: Biased judgments of past events after the outcomes are known. *Psychological Bulletin, 107,* 311-327.

Hoch, S. J. & Loewenstein, G. F. (1989). Outcome feedback: Hindsight *and* information. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 605-619.

Myers, J. L., Shinjo, M., & Duffy, S. A. (1987). Degree of causal relatedness and memory. *Journal of Memory and Language, 26,* 453-465.

Pennington, N., & Hastie, R. (1986). Evidence evaluation in complex decision making. *Journal of Personality and Social Psychology, 51,* 242-258.

Pennington, N., & Hastie, R. (1987). Explanation-based decision making. In *Proceedings of the Ninth Annual Conference of the Cognitive Science Society* (pp. 682-690). Hillsdale, NJ: Lawrence Erlbaum.

Pennington, N., & Hastie, R. (1988). Explanation-based decision making: Effects of memory structure on judgment. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14,* 521-533.

Reder, L. M. (1987). Strategy selection in question answering. *Cognitive Psychology, 19,* 90-138.

Ross, L., Lepper, M. R., & Hubbard, M. (1975). Perseverance in self-perception: Biased attributional processes in the debriefing paradigm. *Journal of Personality and Social Psychology, 5,* 880-892.

Schul, Y., & Burnstein, E. (1985). When discounting fails: Conditions under which individuals use discredited information in making a judgment. *Journal of Personality and Social Psychology, 49,* 894-903.

Sherman, S. J., Skov, R. B., Hervitz, E. F., & Stock, C. B. (1981). The effects of explaining hypothetical future events: From possibility to probability to actuality and beyond. *Journal of Experimental Social Psychology, 17,* 142-158.

Trabasso, T., & van den Broek, P. (1985). Causal thinking and the representation of narrative events. *Journal of Memory and Language, 24,* 612-630.

Wood, G. (1978). The knew-it-all-along effect. *Journal of Experimental Psychology: Human Perception and Performance, 4,* 345-353.

Wyer, R. S., Jr., & Budesheim, T. L. (1987). Person memory and judgments: The impact of information that one is told to disregard. *Journal of Personality and Social Psychology, 53,* 14-29.