Propensities and Counterfactuals: The Loser That Almost Won

Daniel Kahneman and Carol A. Varey
University of California, Berkeley

Close counterfactuals are alternatives to reality that "almost happened." A psychological analysis of close counterfactuals offers insights into the underlying representation of causal episodes and the inherent uncertainty attributed to many causal systems. The perception and representation of causal episodes is organized around possible focal outcomes, evoking a schema of causal forces competing over time. We introduce a distinction between two kinds of assessments of outcome probability: dispositions, based on causal information available prior to the episode; and propensities, based on event cues obtained from the episode itself. The distinction is critical to the use of *almost*, which requires the attribution of a strong propensity to the counterfactual outcome. The final discussion focuses on characteristic differences between psychological and philosophical approaches to the analysis of counterfactuals, causation, and probability.

The question of how people think of things that could have happened but did not has attracted increasing interest among psychologists in recent years (J. T. Johnson, 1986; Kahneman & Miller, 1986; Kahneman & Tversky, 1982a; Landman, 1987; D. T. Miller, Turnbull, & MacFarland, 1990; Wells & Gavanski, 1989; Wells, Taylor, & Turtle, 1987). As philosophers have long known, the study of counterfactuals cannot be separated from a conception of causality, and an understanding of causality requires a conception of possibility and conditional probability. Counterfactual assertions rest on causal beliefs, and causal attributions invoke counterfactual beliefs, for example, about what would have happened in the absence of a putative cause. Some counterfactual assertions assign degrees of probability or plausibility to unrealized outcomes, many causal beliefs are probabilistic, and judgments of probability often draw on impressions of causal tendencies or propensities. The present article is concerned with a psychological analysis of this nexus of issues.

Our study began with an attempt to understand the psychology of assertions of the form "*X* almost happened," which we call *close counterfactuals.* An important characteristic of such assertions is that they are not expressed as a conditional with a specified antecedent, as counterfactual conditionals are. The close counterfactual does not invoke an alternative possible world, but states a fact about the history of this world—namely that things were close to turning out differently than they did.

Our approach combines some elementary phenomenological observations and an equally elementary linguistic inquiry into the conditions under which close counterfactual assertions are appropriate. The genre is not unknown in psychology: Heider (1958) and Schank and Abelson (1977), in particular, have successfully carried out ambitious exercises in this vein. Studies of what people mean when they say that "John went to the restaurant" or when they use the words *can* and *try* have contributed significantly to an understanding of how people think about events and actions. In this article we examine the use of the word *almost* in a speculative attempt to explore how people think about counterfactuals, probability, and causation.[1] The present analysis is restricted to cases in which "*X* almost happened" implies that *X* could have happened. We ignore figurative uses of *almost* in which it is used to denote "coming close" without implication of possibility, as in "at that bend the train almost touches the embankment." We also restrict our discussion of *almost* to cases in which either the actual outcome or the close counterfactual is an achievement (see Lyons, 1977; G. A. Miller & Johnson-Laird, 1976; Vendler, 1967)—a change of state that occurs at a particular moment, usually as the culmination of a longer causal episode. We analyze the beliefs that a speaker expresses by the assertion that an

This article is based on a talk presented at the Conference on Counterfactual Thinking and Other Forms of Mental Simulation, Brown County, Indiana, April 1989. The research was supported by the Air Force Office of Scientific Research, Air Force Systems Command, U.S. Air Force, under Grant AFOSR-88-0206 to Daniel Kahneman.

We are grateful to Maya Bar-Hillel, Baruch Fischhoff, Dale Griffin, Lenore Kahneman, Igal Kvart, Dale Miller, and Amos Tversky for helpful comments and conversations; to Anne Treisman for her insights throughout the project; and to three anonymous reviewers for useful suggestions.

Correspondence concerning this article should be addressed to Daniel Kahneman, Department of Psychology, Tolman Hall, University of California, Berkeley, California 94720.

[1] The importance of the word *almost* as an indication of cognitively and emotionally relevant alternatives to outcomes that actually materialized was pointed out by Heider (1958, pp. 141–144), who drew attention to an instructive passage in Henry Fielding's *Tom Jones* (1749/1975, Vol. 2, p. 691); see also Hofstadter (1979, pp. 634–643).

Journal of Personality and Social Psychology, 1990, Vol. 59, No. 6, 1101–1110
 0022-3514/90/$00.75

individual almost died, or almost missed a deadline, and examine what such beliefs imply to us about the cognitive representation of uncertain events and of causal propensities.

This article develops the following ideas: (a) Counterfactuals, causes, and (some) probabilities are treated as facts about the world, not as constructions of the mind. (b) The absence of perfect hindsight indicates that people attribute inherent uncertainty to causal systems—what happened is not treated as necessary or inevitable. (c) The perception and representation of causal episodes is organized around possible outcomes of the episode. (d) Probabilities of outcomes can be assessed on the basis of advance knowledge (dispositions) or of cues gained from the causal episode itself (propensities). The distinction is critical to the use of *almost,* which requires the attribution of a strong propensity to the counterfactual outcome. (e) Cues to propensity are the temporal or causal proximity of the focal outcome and indications of rapid progress through a causal script. (f) A general schema of causal forces competing over time is applicable to many achievement contexts. (g) There are characteristic differences between a psychological and a philosophical approach to the analysis of probability, causality, and counterfactuals.

The Counterfactual Stance

The statement "*X* almost happened" implies several ancillary beliefs. It commits the speaker to the belief that another observer with the same information would agree with the counterfactual assertion, as would be the case with public, objective facts. The close counterfactual also implies that *X* could have happened, denying the necessity or inevitability of what actually happened and implicitly denying the deterministic character of the situation. We examine these beliefs in the following sections.

Objective Reference

By definition, counterfactual statements refer to events that did not, in fact, occur. However, there is a compelling intuition that some counterfactuals are treated as having an objective character, not as mere mental constructions. Refuting the possible-worlds analysis of counterfactuals, Goodman (1983) put the point strongly: "We have come to think of the actual as one among many possible worlds. We need to repaint that picture. All possible worlds lie within the actual one" (p. 57). In ascribing objective status to counterfactuals, we intend to contrast the attitude toward these objects of thought from the attitude toward imaginings, fantasies, and desires, which are normally tagged as subjective (M. K. Johnson, 1988). The discrimination of what actually happened from what almost did is, of course, essential in the monitoring of reality—the counterfactual event is not perceived as real, but it is not treated as subjective. The "fact" is that the outcome truly is close, or is not close, independently of anyone's beliefs.

The distinction between beliefs that have objective or subjective status recalls an earlier discussion of alternative cognitive representations of probability (Kahneman & Tversky, 1982b). Two main interpretations of that notion were identified, which respectively assign it subjective or objective status. In the subjective interpretation, which is standard in Bayesian philosophy, a probability judgment describes the subject's degree of belief in a proposition. "The probability that the Nile is longer than the Amazon is *p*" is usually understood as describing the speaker's beliefs, not as a fact about these rivers. The attribution of subjective status to a belief is often marked by the possessive pronoun: "My probability that the Nile is longer than the Amazon . . ." is more natural than "the probability that the Nile. . . ." In contrast, the probability that a thumbtack will land on its point if tossed and the probability that Team A will defeat Team B are normally understood as descriptions of the causal dispositions of the thumbtack or of the competing teams. A speaker who wishes to indicate a subjective interpretation of probability will use the possessive pronoun: "My probability that Team A will win is . . ." acknowledges the possibility of valid alternatives, a stance that is not usually adopted in factual statements.

The two types of representation of uncertainty are most clearly distinguished when the uncertainty is removed. The possessive pronoun is then obligatory if the probability has subjective status. "The probability that the Nile is longer than the Amazon was. . ." is simply anomalous. In contrast, it is reasonable to maintain that the probability that the thumbtack would land on its point was .55 even when it is known that it did not do so on a particular instance. More interestingly, the statement that "the probability that Team A would win was high . . ." is acceptable even if that team is known to have lost. The statement of past probability need not refer to anyone's beliefs at the time of the episode. The statement could be made, for example, by a speaker who learned, after the game had ended, that a player of Team B had undertaken to throw the game if he had an opportunity to do so. As this example illustrates, current knowledge of the relevant causal factors may allow a speaker to say retrospectively that an event that did not take place had high probability—and as a special and rather extreme case to assert that the event almost occurred.

Although in this article we often appeal to the reader's intuitions in the expectation that they match ours, we also tested some of our conclusions by collecting judgments of appropriateness from native speakers of English. Subjects were recruited on the Berkeley campus by a poster offering students a small payment in return for a completed questionnaire. Respondents were given instructions and several questions as illustrated by the examples below.

> In the following questions you are asked to rate statements on a scale from "appropriate" to "very peculiar." One or more statements are presented for each question. You are to rate whether the statement in italics is appropriate, given the information in the rest of the question.
>
> 1. *Tom almost died* but in fact he was never in real danger.
>
> Appropriate 7% Somewhat peculiar 27% Very peculiar 66%
>
> (n = 29)[2]

[2] The number in parentheses refers to the number of respondents answering the question. In later examples, results will only be reported for the two extreme categories of response.

2. *Everyone thought Phil almost died* but in fact he was never in real danger.

Appropriate 69% Very peculiar 10% (n = 29)

3. The autopsy showed that when he was a child, Sid had suffered from a rare childhood disease. The pathologist said that if the disease had lasted a few days longer, it would have killed him. No one knew about it at the time; they thought he had a mild case of measles. *Sid almost died as a child from that rare disease.*

Appropriate 61% Very peculiar 0% (n = 18)

These examples illustrate that the close counterfactual has the status of a historical fact. As is generally true when such facts are asserted, everything known to the speaker at the time of the utterance can be relevant, but the beliefs of observers of the actual event are not. As shown by Example 2, the objective status of close counterfactuals allows them to be believed erroneously. Indeed, counterfactuals can be faked. Professional wrestlers on television have perfected the art of appearing almost to kill one another, but they avoided regulation by demonstrating that their occupation is actually quite safe.

Like counterfactuals and (some) probabilities, causal attributions are also treated as objective facts about the world. This is true of causality directly observed, as in the perception of a collision and in Michotte's (1946) demonstrations of launching. It is also true of the more abstract causes that are judged to raise the (objective) probabilities of events or, in some contexts, render them inevitable (Mackie, 1974). The counterfactual assertion that an effect would not have occurred in the absence of the cause, the sine qua non condition of necessity, has the same objective character.

Inherent Uncertainty

The frequent mentions of counterfactual possibilities in everyday discourse demonstrate a prevailing intuition that things could have been different, and in some cases almost were. This intuition commits the speaker to a particular set of beliefs about causality. Specifically, X is neither necessary nor inevitable if it can properly be said that Y almost happened instead of it. Naive intuitions are evidently not dominated by a pervasive belief in strict determinism. Kvart (1986) reached a similar conclusion in his discussion of counterfactual conditionals.

There is an intriguing tension between the intuition that things could have been otherwise and the well-known hindsight effect, in which the inevitability of events that actually took place tends to be exaggerated. The evidence is compelling that retrospective assessments of the probability of events are affected by knowledge of whether or not these events have taken place (Fischhoff, 1975, 1982). The term *creeping determinism* has been used in this context. Two distinct forms of hindsight effects are associated, respectively, with subjective and objective interpretations of probability. The most common test of hindsight effects requires the retrieval of a past state of belief: "What was your probability at the end of 1988 that the Berlin Wall would be opened within a year?" A hindsight bias is revealed in such questions by a tendency to exaggerate the past subjective probabilities of whatever is now known to be true. An example of an objective hindsight question could be "In the light of current knowledge, what was the probability in 1988 that the Berlin Wall would be opened within a year?" A discrepancy between prospective and retrospective probabilities is typically observed in tests of objective as well as subjective hindsight (Fischhoff, 1975).

Unlike the subjective case, hindsight with an objective interpretation of probability is not necessarily a mistake. It is entirely reasonable for an observer to make inferences about a causal system from the knowledge that it produced a particular outcome. Indeed, what is most puzzling in this context is the limited extent of creeping determinism in retrospective evaluations of outcomes.

Of course, not all causal systems are uncertain. As illustrated by most people's attitudes toward the mechanical and electronic devices that surround them, a belief in strict determinism does not require much understanding of how the system works; it cannot be ignorance about the causal system that precludes determinism about close counterfactuals. It is an important fact about causal reasoning that a sense of the necessity of consequences is often absent. In particular, there is no sense of necessity or inevitability in considering games of chance, many contests and competitions, some physical systems (e.g., weather and chance devices), or intentional actions.

The Representation of Causal Episodes

The idea that perceived goals serve to organize the representation of action and imbue events with meaning was articulated by Heider (1958), and is at the core of the more recent treatments of scripts and story grammars (Black & Bower, 1979; Kintsch & van Dijk, 1978; Rumelhart, 1977; Schank, 1975; Trabasso, Secco, & van den Broek, 1984) and treatments of decision making and causal reasoning that rely on a story-based account (Pennington & Hastie, 1988; Read, 1987). Our conception of causal episodes generalizes this idea to achievements. The class of achievements contains the outcomes of intentional action but is much richer: Dying from a disease, a river overflowing its bank, and the Dow–Jones index rising 1,000 points in a year are all achievements. Many achievements are associated with particular causal scripts. The representation of an episode as an instantiation of a causal script is therefore organized in terms of its possible *focal outcomes,* and attention to different achievements will alter the representation. The storm that could fill the reservoirs could also ruin the cherry crop, and its representation will be different if attention is directed to one of these outcomes rather than the other.

To illustrate the function of focal outcomes, we introduce a thought experiment to which we shall repeatedly return. Imagine observing a sequence of red and blue balls as they are drawn from an urn, or the representation of such a process on a computer screen. Note the potent effects of an intention to watch for a particular outcome, such as the color that is most frequent after 11 draws, or an excess of four red balls or six blue balls—whichever happens first. Although these focal outcomes are not goals, they serve the same function in organizing the impression of the sequence. Most important, watching the same sequence with different outcomes in mind alters the experience. Although our thought experiment involves real-time observation and uncertainty, neither of these elements is essential: A designated outcome will affect the interpretation of a story,

and the effect is not reduced when thinking of an episode whose outcome is already known.

The probability of the focal outcome may fluctuate in the course of a causal episode. Changes of probability are always involved in close counterfactuals: Perhaps the most compelling intuition about the statement "*X* almost happened" is that the probability of *X* must have been quite high at some point before it dropped—all the way to zero if another outcome eventually terminated the episode.[3]

There are several reasons for probability changing in the course of an episode. We turn again to the urn example to illustrate two types of probability change. If there was initial uncertainty about the composition of the urn, beliefs about the urn will change to accommodate observed events—by Bayes's rule for an ideal observer—and the probability of the focal outcome will change accordingly. In addition, the actual probability of the focal outcome also changes *because* of the intervening events. Every red ball drawn makes it more probable that the aggregate outcome will be an excess of red over blue balls. The probability of the focal outcomes will change, more or less regularly and perhaps with large fluctuations, until a decisive event brings about an outcome that terminates the episode. Note that this situation can support a close counterfactual: It is easy to imagine a sequence of draws of which it can appropriately be said that the focal outcome almost occurred (red almost won), or almost did not.

The same types of changes of probability will also be found in observing (or hearing about) a storm that could cause a flood or a couple deciding on a joint future. The events that constitute the episode reveal the strength of an underlying causal process, and also contribute to bring about or retard the outcome. They also indicate possible changes in the causal system—changes that could be modeled by an urn whose composition is modified after each draw, perhaps in response to the draw.

Propensities and Dispositions

The discussion so far has been in terms of "objective" probabilities—in the chance example these are probabilities that could be computed precisely, given some initial beliefs about the composition of the urn. It is evident from this example that an account of *almost* in terms of probability has some appeal: The probability of the counterfactual outcome must have been high at some point. It turns out, however, that an account that relies exclusively on objective probabilities will not work. Some aspects of the puzzle to be solved are illustrated by the following examples:

> 4. Mark tried to register for the chess tournament. Because of a problem in mailing the form he missed the registration deadline by one day. Mark is a much stronger player than all the participants in the tournament. *Mark almost won the tournament.*
>
> Appropriate 0% Very peculiar 97% (n = 33)
>
> 5. At the end of a long game of chance, John could have won the whole pot if a die that he rolled showed a six. The die that he rolled was loaded to show six 80% of the time. John rolled it and it showed a two. *The die almost showed six.*
>
> Appropriate 0% Very peculiar 77% (n = 31)
>
> 6. At the end of a long game of chance, John could have won the whole pot if a die that he rolled showed a six. The die that he rolled was loaded to show six 80% of the time. John rolled it and it showed a two. *John almost won the whole pot.*
>
> Appropriate 43% Very peculiar 20% (n = 31)

The close counterfactual is decisively rejected in Examples 4 and 5 but not in Example 6, although the prior probability of the focal outcome was high in all cases. Before it was rolled, the probability of the die showing six was .80 in Examples 5 and 6, and the prior probability of Mark winning the tournament was also high, though unspecified. Despite this, the intuition that *almost* is inappropriate in the first two examples is so strong that they seem almost absurd. A strong belief in the counterfactual conditional "Mark would have won if he had played" is not sufficient to support the close counterfactual "Mark almost won," even if it is also accepted that he almost played in the tournament. Why is this the case? And what else is required for the close counterfactual to be appropriate?

The answer to the first question is that the close counterfactual is never appropriate if it is only supported by indications of likelihood or causal force that were available before the onset of the relevant causal episode. Achievements, such as winning a tournament, getting married, or a die showing six, are associated with causal scripts that usually have a definite starting point: when play begins, when the couple start dating, when the die is rolled. Probabilities can be assigned to possible outcomes of a causal process before it is initiated: Mark may be a rated player, the couple could appear severely mismatched, the die could be loaded. We shall refer to the cognitive representation of such prior probabilities as the (perceived) *disposition* of a causal system to yield particular outcomes. Examples 4 and 5 show that dispositions, however strong, do not suffice to support the assertion of a close counterfactual.

A close counterfactual must be supported by the evidence of event cues, as these accumulate in the course of the causal episode. We use the term *propensity* for what is learned about the probability of an outcome from observing event cues or from hearing about them. Mark had a disposition to win his chess tournament and probably would have won it if he had registered, but the causal episode for his victory never began, and there was therefore no opportunity to establish a propensity for that outcome. The standard example of propensity in a chance event is the cinematic cliché of the roulette wheel that slows down as it approaches a critical number, slows down even more, leans against the spring, then finally trips it and stops on a neighboring number. To be described as almost showing six, a die must display a propensity to stop its roll in that position.

The contrasting responses to Examples 5 and 6 illustrate the need to distinguish propensity from probability. We suppose that our respondents would have assigned a probability of .8 both to the die showing six and to John winning the whole pot. However, the propensities of the two outcomes clearly differ.

---

[3] Kvart (1986) has offered a treatment for a broad class of counterfactual conditionals in which causality is explicated by probabilities that change over time.

Example 6 illustrates a common structure in which one achievement (the die showing six) is nested inside another (John winning the pot). Because the focal outcome of winning the pot invokes a more inclusive causal episode that had begun long before the critical play, John can be said to have had a propensity to win, even if it is not established that the die had a propensity to show six. Thus, although the two statements have the same probability before the throw, the differential effects of propensity and disposition allow "John almost won" to be appropriate although "the die almost showed six" is not.

We should now review the rather subtle relations among the concepts of disposition, propensity, and probability, as they are used in this article. Disposition has been defined as the cognitive representation of the probability of a focal outcome, before the beginning of the relevant causal episode. A disposition can be assessed either prospectively or in hindsight, depending on whether or not the outcome is known. Disposition is a psychological construct, not a logical or mathematical one, and in view of what is known about intuitive judgment there is little reason to expect dispositions to obey the standard axioms of probability (Kahneman, Slovic, & Tversky, 1982). Dispositions represent knowledge about the particular causal system that will (or will not) produce the focal outcome of current concern. Dispositions are inferred from the base rates of outcomes previously produced by that system (Mark has won most of his tournaments) or from structural knowledge that supports causal inferences (the die was loaded in a particular fashion). Thus, the concept of disposition has causal as well as statistical implications.

Our concept of propensity is even more imbued with causal content. Event cues reveal the causal system in action. They indicate advance toward the focal outcome, or regression away from it. They suggest changes in the momentary state of the causal system—changes that may be real or illusory, as when a player is seen to have a "hot hand" (Gilovich, Vallone, & Tversky, 1985). Perhaps most important, propensities depend on the proximity of the outcome, on the possibility of quickly achieving a decisive advance to it. In sharp contrast to probability, the propensities for all competing outcomes of a process may be low early in a causal episode, and more than one propensity can be high at once when the end is close. These ideas are elaborated in subsequent sections.

Our main interest in the remainder of this article is to use close counterfactuals to learn about propensity. We consider propensity to be a dimension of the experience and cognitive representation of events, just as pitch is a dimension of auditory experience. There should be no presuppositions about the determinants of propensity; in particular, propensity could reflect causality as well as probability, just as pitch depends on both the frequency and the intensity of sound. To anchor this speculative analysis in observables, we assume that the appropriateness of *almost,* in its literal meaning, provides a usable indication of high propensity.

Disposition Neglect

Dispositions and propensities are differentially susceptible to revision in hindsight. Consider two cases in which the observer of the last lap of a footrace might assign a high probability of victory to a particular runner: (a) a runner who is in contention and is known to have a strong finish, or (b) a runner who has been catching up rapidly with the leader. The real-time expectations are equally strong in both cases, we assume, but they are based on different cues—dispositional knowledge in (a) and event cues in (b). Now imagine that the two runners both fail to win, by the same amount: The first did not show a strong finish and the second never quite caught up. Note that it will not do to say of the runner who usually has a strong finish that he or she almost won the race with a strong finish, if in fact he or she showed no evidence of talent on that particular occasion. The close counterfactual that the loser almost won is more applicable to (a) than to (b), although a counterfactual conditional could be appropriate in (a). The general hypothesis is that dispositional expectations that are not confirmed by event cues become irrelevant in hindsight.

The differential weighting of event cues and dispositional expectations in retrospective judgments will be called *disposition neglect;* the effect bears an intriguing resemblance to the relative neglect of base-rate information that has been observed in some prospective judgments. For example, the judged probability that a short personality sketch describes a lawyer rather than an engineer is not much affected by the proportion of engineers and lawyers in the sample from which it was drawn. The information about the individual case largely supersedes the information about the base rate instead of combining with it according to Bayes's rule (Kahneman & Tversky, 1973). Similarly, Ajzen (1977) found that people predicting exam success for a student based their predictions on a descriptive sketch and gave little weight to the information that the student was drawn at random from a set selected by a researcher to include 75% failures. As Ajzen observed, however, the neglected base rate in these examples is merely statistical. There is no causal connection between the composition of the student sample and the factors that would make a particular student succeed or fail. The situation changes when such a causal connection is provided: The information that 75% of students taking the test failed it leads readily to the inference that the test was a difficult one, and the information has much more impact on the judgment of the probable success of an individual (Ajzen, 1977). There have been other demonstrations of the general principle that causally relevant base-rate information will not be neglected (Gigerenzer, Hell, & Blank, 1988; Tversky & Kahneman, 1980, 1982; see also Bar-Hillel, 1990, for a discussion of these issues).

Ajzen's (1977) experiment demonstrated that dispositional information tends to dominate statistical base rates, and that dispositional information from two sources (the difficulty of the exam and the student's ability) tends to be integrated. A variation of this experiment would demonstrate disposition neglect: Evidence that a student is extremely able does not support the inference that the student almost passed an exam that he failed, nor does the knowledge that a test was very hard support the conclusion that a student who passed almost failed it. In assessing close counterfactuals, event cues dominate causal base rates and other dispositional information.

The neglect of statistical base rates leads to violations of

Bayes's rule in prospective judgments. The neglect of dispositional expectations in hindsight is not necessarily an error, but the psychology of the two effects may well reflect a single general principle. In both cases the data that bear most directly on the causal forces at work in the individual case have the greatest impact.

Correlates of Propensity

In this section we develop the concept of propensity by examining two of its close correlates: shrinking distance and increasing impact. The role of distance and motion in the close counterfactual is evident in the near synonymy of "*X* almost happened," "*X* nearly happened," and "*X* was close to happening." These expressions invoke a rich metaphor in which an extended causal process is represented as movement in space (e.g., see the "source–path–goal" kinesthetic image schema analyzed in Lakoff, 1987; the various "journey" metaphors in Lakoff & Johnson, 1980; and the force and space images in Talmy, 1981, 1983). This metaphor imposes a metric of causal distance between situations and suggests the closest approach to an outcome as a measure of its propensity. The second correlate of increasing propensity is an escalation in the apparent causal significance of events as the outcome is approached.

Causal Proximity

The present analysis has emphasized causal processes that extend over time, but close counterfactuals can be asserted on the basis of a measure of proximity or similarity even when the process is instantaneous. For example, the statement "The house was almost struck by lightning" is appropriate when lightning struck nearby. The actual outcome is the only event cue in such cases, and it induces a gradient of propensity in its spatial and temporal vicinity. From the fact that lightning struck in a particular place at a particular time, a propensity is inferred to strike in neighboring places, and at about the same time. Similarly, it is appropriate to say that Tom almost got six sixes in rolling dice if he got five sixes and a two. Indeed, it would be even more appropriate to say that Tom almost got six sixes if he rolled five sixes and a five.

Scripts for achievements often specify a series of landmarks that provide a provisional metric of proximity to the outcome. Getting a wedding license, for example, is one of the last landmarks in the script for marriage. It will usually be appropriate to say of a couple that came that far but did not marry that they almost got married. However, although high propensity for an outcome can be inferred from the near completion of the script for that outcome, such inferences are tentative and dependent on default assumptions about the causal system. Thus, it is not correct to say of a tethered mountain climber who falls that he or she "almost fell to the bottom of the cliff," or even was close to doing so, although the script for a fall to the bottom was almost completely satisfied. Nor will it be correct to say that Tom almost rolled six sixes if one of the dice has been altered to make that outcome impossible. The propensity for a counterfactual outcome cannot be reduced to a superficial assessment of the similarity of the actual episode to the completed script for that outcome.

Intentions can contribute to an impression of propensity. For example, it is more appropriate to say that the escaping murderer was almost killed by a shot that went six inches above his or her head if the shot was intended to kill than if it was intended to warn. Intentions do not suffice, however, when there are significant obstacles to be overcome. For an individual to "consider doing *X*" is sometimes sufficient to support the inference that the individual "almost did *X*," but not always. Selected examples follow:

7. Martin considered getting married to Meg. *Martin almost married Meg.*

Appropriate 14% Very peculiar 34% (n = 29)

8. Neil considered not getting married to Amanda. *Neil almost didn't marry Amanda.*

Appropriate 62% Very peculiar 19% (n = 32)

9. Fred considered stealing his child's savings. *Fred almost stole his child's savings.*

Appropriate 30% Very peculiar 15% (n = 75)

10. Ned considered breaking into a bank vault. *Ned almost broke into a bank vault.*

Appropriate 18% Very peculiar 44% (n = 75)

Mere consideration of a marriage is not sufficient (at least in this culture) to support the assertion that the marriage almost took place. The situation is somewhat different in Example 8, because either party (again in this culture) has the power singlehandedly to put a stop to plans to marry. Responses to Examples 9 and 10 show that subjects are sensitive to the fact that much more remains to be done, beyond mere consideration, for the project of breaking into a bank vault than for stealing one's child's savings.

Decisiveness

Many outcomes are produced by a conjunction of events, all contributing to making the outcome necessary. It is useful to distinguish two privileged roles of events in multiple causation: *Critical events* are those that initiate a causal episode, potentiate subsequent causal events, or both; *decisive events* are those that rule out all alternatives, and ensure (or almost ensure) a particular outcome. The special role of critical events that initiate coherent causal episodes has been confirmed in studies of blame (J. T. Johnson, Ogawa, Delforge, & Early, 1989) and studies of mental simulations that "undo" outcomes (Wells et al., 1987). The person who starts a quarrel will get much of the blame for its consequences. However, the decisive and irreversible events that terminate causal episodes are also important, especially when the events in the causal sequence are not themselves causally related (D. T. Miller & Gunasegaram, 1990). Hart and Honore (1959) proposed that a cause is found by "tracing back" from the effect to the nearest plausible candidate in the causal chain. They also discussed the legal doctrine of the last clear chance: The last person who had a good chance to avoid harm is alone held responsible (see also Wells & Gavanski, 1989). The responsible individual is the one whose actions cannot be reversed by anyone else. The same intuition shows up in the context of blackjack; many players believe that the player on the

seventh box, who receives cards immediately prior to the dealer's draw that all players are trying to beat, determines the outcomes for all players (Keren & Wagenaar, 1985). By the time the cards are dealt, the sequence of cards is fixed, though unknown, and the seventh player, by refusal or acceptance of a card, decisively determines its allocation.

It is instructive to analyze decisiveness in terms of probability. Consider an urn game that ends whenever the excess of balls of one color reaches a critical value. Suppose the prior probability of red being the "winning" color is high, because there are more red than blue balls in the urn. Now imagine another scenario, which involves a balanced urn and a majority of red balls in early draws. When the objective probabilities of a red victory are matched in these two scenarios, the probabilities of two more specific events will be higher in the case favored by event cues: (a) the probability of the outcome occurring *soon*, and (b) the probability that the current lead will be preserved until the end of the game. We suggest that impressions of propensity are related to the probability of the next favorable event being decisive, and of current progress not being reversed before the outcome is reached.

The intuition that causal impact increases in the course of the episode is especially compelling when the episode terminates at a fixed time. Obviously, the probability that a team that leads by a touchdown will win the ball game must increase as time remaining to play diminishes. A score that changes the lead is accordingly perceived as more likely to be decisive if it comes late rather than early in the game. Correspondingly, the close counterfactual is most compelling if the propensity for the unrealized outcome peaked late in the causal episode. An early event may support a counterfactual such as "Team A could have won if Fred had not missed that touchdown in the first quarter," but the description "Team A almost won" is much more convincing if the missed touchdown happened in the closing minutes of the game.

Propensities for all outcomes will be weak in the early phases of a causal episode, if no decisive advantage can be gained at that time. Early in a football game, neither team has a strong propensity to win, although one of them may have a strong disposition to do so. Later on, propensities to win will be attributed to a team to the extent that it already has, or appears on its way to achieving, a lead that is likely to be maintained to the end. Toward the end of the game, a team with a large lead has an overwhelming propensity to win, and both teams have a significant propensity if the game is close. On the usual interpretation of probability, of course, the sums of the probabilities of victory for the two teams (barring ties) should add to one at all stages of the game. A formal representation of propensities should incorporate the attribute of noncomplementarity, which is admissible in some nonstandard models of probability (Shafer, 1976).

Competitive Causation

The psychological concept of propensity that was introduced in the preceding section has a dual meaning as a probabilistic and as a causal notion. We have interpreted propensity as an intuitive assessment of the current probability of the focal outcome based on event cues, and also as an assessment of the current probability of particular cases of the focal outcome—for example, the event of this outcome occurring soon. But the term *propensity* was chosen because it also denotes a direct expression of causal force—*Webster's Dictionary* defines propensity as "an urgent and often intense natural inclination." Urgency and intensity are not part of the meaning of probability in theoretical discourse. We suggest, however, that these dynamic features are important aspects of the cognitive representation of many causal processes, including, in particular, the processes that have achievements as outcomes.

The probabilistic and the causal aspects of propensity suggest different representations of the relation between the alternative outcomes of a causal process. In the language of probability, this relation is expressed by complementarity: Changes in the probability of the focal outcome are mirrored by compensating changes in the aggregate probability of other outcomes. In the language of causal dynamics, the relation between alternative outcomes is best described as competition and conflict. The competition metaphor is evident in many phrases chosen to describe episodes and their outcomes (e.g., "They had to admit defeat and gave up hope of beating the deadline" or "The Harvard job offer won out"). A competitive model of causation is particularly appealing for close counterfactuals, where the strongest propensity is associated first with one outcome, then with another—suggesting a shifting balance between variable opposing forces.

A schema of competing and interacting propensities is most obviously applicable to athletic contests, from which several of our examples have been drawn, but is not restricted to these situations. Displays of the chance games that we have discussed invite a competitive interpretation, much as the figures in the famous Heider and Simmel (1944) animation evoke impressions of intentionality and meaningful interaction. We propose the general hypothesis that the competitive schema is commonly evoked by situations in which the focal outcome is an achievement. These include such varied cases as the making of a difficult individual decision, the vicissitudes of a couple that may or may not break up or get married, the struggle of a firm threatened with bankruptcy, the story of a life-threatening illness, the construction of a building under time constraints, and the wrecking of a building by a tropical storm. Each of these situations is defined by one or more focal achievements. Causal episodes that produce such achievements, or fail to produce them, are naturally described as a struggle of conflicting and variable forces favoring alternative outcomes, or in some cases as a struggle between a single variable force and a series of obstacles.

The notion of conflict between opposing forces is not new to psychological analyses of causality, at least in the context of explaining action. Lewin (1936) introduced motion in a force field as a model of action under conflict. His theory influenced Heider's subsequent analysis of the naive theory of action, in which action is the resultant of the effective personal force and the effective environmental force (Heider, 1958). Both models explain action as a vectorial combination of forces. Lewin's famous theory of conflict also incorporated a dynamic element: The forces acting on the individual change predictably as the individual moves toward sources of attraction or away from aversive states. In general, however, applications of force field

analysis have been static. There has been little emphasis on time or on the possibility of causal forces interacting and changing in the course of an event.

A model of competing propensities would extend Lewinian force field analysis in several ways. First, the concept of focal achievement applies to outcomes that are not goals, such as someone dying from a disease, and to situations that do not involve intentions at all, such as a storm destroying a building. Second, the focus of the competitive model is on extended causal episodes. Third, the competitive model attributes inherent uncertainty to causal systems and describes causal episodes in terms of propensities that may change and interact.

Psychology—or Philosophy?

The present study occupies a somewhat uncomfortable middle ground between psychology and philosophy. We have attempted to identify the conditions under which a particular class of counterfactual assertions would be considered true, or appropriate, and we have introduced a notion of propensity to account for these observations. The questions we addressed are similar to those modern philosophers often raise: Philosophical analyses of counterfactuals, for example, focus on the truth conditions or assertability conditions for counterfactual conditionals. Furthermore, some elements of the method are similar: Persuasive philosophical arguments commonly draw on compelling examples that evoke strong shared intuitions. Although the final product of philosophical analysis often has the form of a formal deductive system, induction from intuitions about particular examples is clearly an important part of philosophical endeavor. However, there are important differences between the aims and assumptions of the two disciplines. Philosophers try to understand causality, probability, or counterfactual conditionals, whereas psychologists try to understand how people think about these topics. These different aims have important consequences in the attitude toward logical consistency: Understanding a matter involves imposing a consistent logical structure on it, but the study of human thinking should neither assume nor impose consistency on its subject matter.

There is a large and interesting philosophical literature on counterfactuals (e.g., Adams, 1976; Goodman, 1954; Lewis, 1973, 1979; Nute, 1980; Pollock, 1976; Skyrms, 1980; Stalnaker, 1968).[4] After developing our notions of causal episodes and changing propensities, we encountered similar ideas in Kvart's treatment of counterfactuals and in his later work on causality (Kvart, 1986, 1989). Kvart (1986) described the truth conditions for counterfactual conditionals on the basis of causal processes diverging from actual historical processes at a particular point in time. He introduced a notion of causal paths, explicated by reference to conditional probabilities changing over time. Kvart also emphasized that the commonsense view of the world is nondeterministic, involving a concept of an open future. As might be expected in a philosophical analysis, Kvart treated counterfactuals as objects of thought, not as constructions of the mind. He also had recourse to formal notions of probability and to formal constraints on causal paths, which we have avoided.

Psychologists have drawn most heavily on the tools and concepts of logical and philosophical analysis in studies of deductive reasoning (Braine, 1978; Johnson-Laird, 1983; Rips, 1990). The costs of such borrowing could be high in studies of causality, probability, and counterfactuals. The intensity of current philosophical debate regarding these topics suggests the existence of compelling but mutually inconsistent intuitions. The concepts that have been developed in attempts to resolve these inconsistencies are sometimes quite remote from the naive categories of thought with which psychologists are concerned. Just as an understanding of naive physics may benefit more from acquaintance with Aristotelian physics than with the modern variety, psychological studies of causality, probability, and counterfactuals may do well to avoid exaggerated dependence on the categories of modern philosophical thought.

The dominant approach to causality in psychology, perhaps reflecting a similar dominance in philosophy, treats causation as a particular relationship of dependency between events—expressed by necessary or sufficient conditions or by increased conditional probabilities (Einhorn & Hogarth, 1986; Kelley, 1967; Mackie, 1974). There is another view, however, which treats causality as a directly perceived link between events or as an emergent property of a patterned sequence of events. The main sources of this approach to causality in psychology are still the classic works by Michotte (1946) and Heider and Simmel (1944), which, respectively, explored variations on the themes of spatiotemporal contiguity and of schemas of intentional action. Ducasse (1969) has developed a philosophical analysis that draws on similar intuitions. The notion of propensity that has been presented here belongs to this tradition of research in causality.

In our use of the term, the representation of propensity is inherently causal, and inherently predictive, much like the perception of an object in motion (Freyd & Finke, 1984). Our emphasis on event cues to propensity deliberately straddled the standard distinction between causes and effects as well as the distinction between causal force and probability. Is there a justification for a concept that blurs accepted distinctions between important categories of thought? There may be. We have described propensity as a perceived attribute with objective reference, much like the perceived length of a line or the perceived distance of an object. Even in the case of lengths and distances, the crude correspondence of the dimensions of percepts to the dimensions of physical description of the world does not guarantee correspondence of the geometries that describe the space people perceive and the space in which they move. The more general point is that the mental representations of events and their relations may not correspond to any logical analysis of causality or probability, and that intuitions about these matters may not be internally consistent. The student of lay intuitions faces a problem that is familiar to cultural anthropologists: How does one make sense of a system of thought without imposing alien categories on it?

---

[4] Skyrms (1980) has a treatment of counterfactual conditionals that relies on what he calls "prior propensities." However, the meaning of his term more closely resembles our usage of dispositions.

References

Adams, E. W. (1976). Prior probabilities and counterfactual conditionals. In W. L. Harper & C. A. Hooker (Eds.), *Foundations of probability*

theory, statistical inference, and statistical theories of science (Vol. 1, pp. 1–21). Dordrecht, Holland: Reidel.

Ajzen, I. (1977). Intuitive theories of events and the effects of base-rate information on prediction. *Journal of Personality and Social Psychology, 35,* 303–314.

Bar-Hillel, M. (1990). Back to base rates. In R. M. Hogarth (Ed.), *Insights in decision making: A tribute to Hillel J. Einhorn* (pp. 200–216). Chicago: University of Chicago Press.

Black, J. B., & Bower, G. H. (1979). Episodes as chunks in narrative memory. *Journal of Verbal Learning and Verbal Behavior, 18,* 187–198.

Braine, M. D. S. (1978). On the relation between the natural logic of reasoning and standard logic. *Psychological Review, 85,* 1–21.

Ducasse, C. J. (1969). *Causation and the types of necessity.* New York: Dover.

Einhorn, H., & Hogarth, R. (1986). Judging probable cause. *Psychological Bulletin, 99,* 3–19.

Fielding, H. (1975). *The History of Tom Jones, a foundling* (Vol. 2). Oxford, England: Wesleyan University Press/Oxford University Press. (Original work published 1749)

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. (1982). For those condemned to study the past. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 422–444). Cambridge, England: Cambridge University Press.

Freyd, J., & Finke, R. (1984). Representational momentum. *Journal of Experimental Psychology: Learning, Memory and Cognition, 10,* 126–132.

Gigerenzer, G., Hell, W., & Blank, H. (1988). Presentation and content: The use of base rates as a continuous variable. *Journal of Experimental Psychology: Human Perception and Performance, 14,* 513–525.

Gilovich, T., Vallone, R., & Tversky, A. (1985). The hot hand in basketball: On the misperception of random sequences. *Cognitive Psychology, 17,* 295–314.

Goodman, N. (1954). *Fact, fiction and forecast* (1st ed.). London: Athlone Press.

Goodman, N. (1983). *Fact, fiction and forecast* (4th ed.). Cambridge, MA: Harvard University Press.

Hart, H. L. A., & Honore, A. M. (1959). *Causation in the law.* London: Oxford University Press.

Heider, F. (1958). *The psychology of interpersonal relations.* New York: Wiley.

Heider, F., & Simmel, M. (1944). An experimental study of apparent behavior. *American Journal of Psychology, 57,* 243–259.

Hofstadter, D. (1979). *Godel, Escher, Bach: An eternal golden braid.* New York: Basic Books.

Johnson, J. T. (1986). The knowledge of what might have been: Affective and attributional consequences of near outcomes. *Personality and Social Psychology Bulletin, 12,* 51–62.

Johnson, J. T., Ogawa, K. H., Delforge, A., & Early, D. (1989). Causal primacy and comparative fault: The effect of position in a causal chain on judgments of legal responsibility. *Personality and Social Psychology, 15,* 161–174.

Johnson, M. K. (1988). Discriminating the origin of information. In T. F. Oltmanns & B. A. Maher (Eds.), *Delusional beliefs: Interdisciplinary perspectives* (pp. 34–65). New York: Wiley.

Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness.* Cambridge, MA: Harvard University Press.

Kahneman, D., & Miller, D. T. (1986). Norm theory: Comparing reality to its alternatives. *Psychological Review, 93,* 136–153.

Kahneman, D., Slovic, P., & Tversky, A. (Eds.). (1982). *Judgment under uncertainty: Heuristics and biases.* New York: Cambridge University Press.

Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review, 80,* 237–251.

Kahneman, D., & Tversky, A. (1982a). The simulation heuristic. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 201–208). New York: Cambridge University Press.

Kahneman, D., & Tversky, A. (1982b). Variants of uncertainty. *Cognition, 11,* 143–157.

Kelley, H. H. (1967). Attribution theory in social psychology. In D. Levine (Ed.), *Nebraska Symposium on Motivation* (Vol. 15, pp. 192–240). Lincoln: University of Nebraska Press.

Keren, G., & Wagenaar, W. A. (1985). On the psychology of playing blackjack: Normative and descriptive considerations with implications for decision theory. *Journal of Experimental Psychology: General, 114,* 133–158.

Kintsch, W., & van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review, 85,* 363–384.

Kvart, I. (1986). *A theory of counterfactuals.* Indianapolis: Hackett.

Kvart, I. (1989). *Causal independence.* Unpublished manuscript.

Lakoff, G. (1987). *Women, fire, and dangerous things: What categories reveal about the mind.* Chicago: University of Chicago Press.

Lakoff, G., & Johnson, M. (1980). *Metaphors we live by.* Chicago: University of Chicago Press.

Landman, J. (1987). Regret and elation following action and inaction. *Personality and Social Psychology Bulletin, 13,* 524–536.

Lewin, K. (1936). *Principles of topological psychology* (F. Heider & G. M. Heider, Trans.). New York: McGraw-Hill.

Lewis, D. (1973). *Counterfactuals.* Cambridge, MA: Harvard University Press.

Lewis, D. (1979). Counterfactual dependence and time's arrow. *Noûs, 13,* 455–476.

Lyons, J. (1977). *Semantics, Volume 2.* Cambridge, England: Cambridge University Press.

Mackie, J. L. (1974). *The cement of the universe: A study of causation.* Oxford, England: Clarendon Press (Oxford University Press).

Michotte, A. E. (1946). *La perception de la causalité* [The perception of causality]. Louvain, Belgium: University of Louvain Publications.

Miller, D. T., & Gunasegaram, S. (1990). Temporal order and the perceived mutability of events: Implications for blame assignment. *Journal of Personality and Social Psychology, 59,* 1111–1118.

Miller, D. T., Turnbull, W., & McFarland, C. (1990). Counterfactual thinking and social perception: Thinking about what might have been. In M. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 23, pp. 305–331). Orlando, FL: Academic Press.

Miller, G. A., & Johnson-Laird, P. L. (1976). *Language and perception.* Cambridge, England: Cambridge University Press.

Nute, D. (1980). *Topics in conditional logic.* Boston: Reidel.

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.

Pollock, J. L. (1976). *Subjunctive reasoning.* Dordrecht, Holland: Reidel.

Read, S. J. (1987). Constructing causal scenarios: A knowledge structure approach to causal reasoning. *Journal of Personality and Social Psychology, 52,* 288–302.

Rips, L. (1990). Reasoning. *Annual Review of Psychology, 41,* 321–353.

Rumelhart, D. E. (1977). Understanding and summarizing brief stories. In D. Laberge & J. Samuels (Eds.), *Basic processes in reading, perception and comprehension* (pp. 265–303). Hillsdale, NJ: Erlbaum.

Schank, R. C. (1975). The structure of episodes in memory. In D. G.

Bobrow & A. Collins (Eds.), *Representation and understanding: Studies in cognitive science* (pp. 237–272). New York: Academic Press.

Schank, R. C., & Abelson, R. P. (1977). *Scripts, plans, goals, and understanding.* Hillsdale, NJ: Erlbaum.

Shafer, G. (1976). *A mathematical theory of evidence.* Princeton, NJ: Princeton University Press.

Skyrms, B. (1980). The prior propensity account of subjunctive conditionals. In W. L. Harper, R. Stalnaker, & G. Pearce (Eds.), *Ifs: Conditionals, belief, decision, chance and time* (Vol. 1, pp. 125–162). Dordrecht, Holland: Reidel.

Stalnaker, R. C. (1968). A theory of conditionals. In N. Rescher (Ed.), *Studies in logical theory: American Philosophical Quarterly monograph series 2* (pp. 98–112). Oxford, England: Blackwell.

Talmy, L. (1981, May). *Force images.* Paper presented at the Conference on Language and Mental Imagery, University of California, Berkeley.

Talmy, L. (1983). How language structures space. In H. Pick & L. Acredolo (Eds.), *Spatial orientation: Theory, research and application* (pp. 219–282). New York: Plenum Press.

Trabasso, T., Secco, T., & van den Broek, P. (1984). Causal cohesion and story coherence. In H. Mandl, N. L. Stein, & T. Trabasso (Eds.), *Learning and comprehension of text* (pp. 83–111). Hillsdale, NJ: Erlbaum.

Tversky, A., & Kahneman, D. (1980). Causal schemas in judgments under uncertainty. In M. Fishbein (Ed.), *Progress in social psychology* (Vol. 1, pp. 153–160). Hillsdale, NJ: Erlbaum.

Tversky, A., & Kahneman, D. (1982). Evidential impact of base rates. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 201–208). New York: Cambridge University Press.

Vendler, Z. (1967). *Linguistics in philosophy.* Ithaca, NY: Cornell University Press.

Wells, G. L., & Gavanski, I. (1989). Mental simulation of causality. *Journal of Personality and Social Psychology, 56,* 161–169.

Wells, G. L., Taylor, B. R., & Turtle, J. W. (1987). The undoing of scenarios. *Journal of Personality and Social Psychology, 53,* 421–430.

Received December 30, 1989
Revision received July 5, 1990
Accepted July 9, 1990 ■