

Malle, B. F., & Tate, C. (2006). Explaining the past, predicting the future. In L. J. Sanna & E. C. Chang (Eds.), *Judgments over time: The interplay of thoughts, feelings, and behaviors* (pp. 182-209). New York: Oxford University Press.

# 11

## Explaining the Past, Predicting the Future

---

BERTRAM F. MALLE  
CHUCK TATE

### The Cognitive Apparatus of Explanation

#### *Classic Attribution Theory*

The psychological literature on explanation was born in the 1950s and 1960s, with seminal contributions by Heider (1958), Jones and Davis (1965), and Kelley (1967). Putting important differences aside (see Malle, 2004, chapter 1), all these authors identified cognitive inference as the underlying mechanism by which explanation is achieved. Specifically, explanations are based on inference from observation, past experience, and knowledge structures (Abelson & Lalljee, 1988). According to traditional attribution theories, social perceivers wonder primarily whether a person cause or a situation cause explains the event in question. For example, did she invite him to dinner because of something about her (e.g., her friendly personality) or because of something about him (e.g., his irresistible charm)? Researchers tried to identify the rules of inference people follow when making these person-situation attributions, and Kelley's (1967) covariation model was revolutionary in this respect, specifying the inferential integration of multiple pieces of information in a process akin to analysis of variance.

Traditional attribution theories suffered from a number of shortcomings. First, the theories narrowed the kinds of explanations people seek to only two types (person vs. situation attributions), and even though some theorists expanded this set (Abramson, Seligman, & Teasdale, 1978; Weiner, 1985), the assumption remained that people select "types of causes" in terms of their

abstract features (internal-external, stable-unstable, global-specific). Evidence was accumulating, however, that people explain behavior with a variety of conceptual entities, such as goals, reasons, and preconditions (Buss, 1978; McClure, 2002; McClure & Hilton, 1997; Read, 1987).

A second and related shortcoming was that traditional attribution theories considered explanations of generic events, thereby failing to distinguish, as people do, between explanations for intentional and other types of events (Malle & Knobe, 1997b). Actually, Heider (1958) did make this distinction under the labels "personal vs. impersonal causality," but this concept pair was mistakenly interpreted as the distinction between person and situation causes and folded into Kelley's scheme (Malle, 2004; Malle & Ickes, 2000). The first and second shortcomings are related because the genuine variety of explanation types (goals, reasons, causes, etc.) suddenly makes sense when we separate out intentional from unintentional events.

A third shortcoming was that covariation inference was considered the only process that generates explanations. Even though such inference may play an important role in certain contexts, other processes for generating explanations clearly play a role as well (Abelson & Lalljee, 1988; Ahn, Kalish, Medin, & Gelman, 1995; Andrews, 2003; Malle, 2004, chapter 5). For example, explainers sometimes directly recall a particular goal or purpose and at other times simulate the agent's reasoning or experience in order to arrive at a plausible explanation.

Over the past decade, our research group has attempted to mend the shortcomings of attribution theory, integrate the important criticisms voiced over the years, and offer a comprehensive model of explanation that takes into account conceptual, cognitive, linguistic, and social aspects of explanation (Knobe & Malle, 2002; Malle, 1999, 2001, 2004, in press; Malle, Knobe, O'Laughlin, Pearce, & Nelson, 2000; O'Laughlin & Malle, 2002). Here we summarize the major features of our folk-conceptual theory of explanation in order to develop new ideas about the temporal dimension of explanation.

### *The Folk-Conceptual Theory of Explanation*

The name *folk-conceptual theory* was chosen because the theory highlights the folk concepts by which people make sense of mind and behavior—concepts such as intention and intentionality, reasons, enabling factors, and the like. Besides this conceptual layer, the theory has two other layers: psychological processes by which people generate explanations and the linguistic tools by which they express explanations. Finally, the model is embedded in considerations of social functions that explanations serve, including impression management, conversational clarification, persuasion, and propaganda (Malle, 2004).

We now introduce the three layers of the theory and then apply them to an analysis of temporal aspects of explanation.

#### First Layer: Conceptual Framework and Modes of Explanation

The theory's first and fundamental layer describes the *conceptual framework* that underlies behavior explanations. As Heider (1958) argued, people sharply distinguish between intentional and unintentional behavior. Moreover, a close inspection of the folk concept of intentionality (Malle & Knobe, 1997a) suggested that people employ at least three different modes of explanation for intentional actions (Malle, 1999, 2001): reason explanations, causal history of reason explanations, and enabling-factor explanations.

Most often, people explain an action by the *reasons* for which the agent performed the action. Reasons are typically beliefs and desires in light of which and on the grounds of which the agent decided to act. That is, reasons have two defining features: agents are *aware* of their reasons, and agents consider their reasons to be *rational grounds* for acting. For example:

I am going to read some proverbs *because they are supposed to be extremely wise, and I need some wisdom in my life.*

Second, people explain some actions by the *causal history* of those reasons—if, for example, the reasons themselves are not known or are of little interest. Causal history or reason explanations refer to context, personality, culture, or unconscious processes that led up to (lay in the history of) an agent's consideration of certain beliefs or desires as reasons for acting. However, causal history factors are not themselves reasons. For a causal history explanation to be appropriate, the agent need not be aware of the causal history factor, nor does the factor have to provide any rational grounds for acting. Like reasons, causal history explanations account for intentional action; but, unlike reasons, they do not conform to the constraints of awareness and rationality. To sharpen the contrast between reason explanations and causal history explanations, Table 11.1 juxtaposes both explanation modes for the same actions.

Third, occasionally people are not interested in what motivated an agent's action but what made it possible that the action was successfully performed (Malle et al., 2000; McClure & Hilton, 1997). These *enabling factor* explanations refer to processes inside or outside the agent that facilitated or permitted (i.e., enabled) the successful action performance.

Whereas people's concept of intentional action provides three distinct modes of explanation, their concept of unintentional behavior is far simpler. Unintentional behaviors are treated like any other events (e.g., physical or biological)—“mechanically” brought about without the necessary involvement of intention,

Table 11.1. Examples of reason explanations and CHR explanations that were distinguished reliably by social perceivers

Behavior	Reasons	CHR explanations
Nancy chose not to vote in the last election.	None of the candidates appealed to her.	Her mom died that week.
	She wasn't interested in the issues.	She doesn't realize that every vote counts.
Ian worked 14 hours a day last month.	To make more money	He is driven to achieve.
	He wants to get ahead.	That's the cultural norm.

Note: Adapted from O'Laughlin and Malle (2002). Behaviors and explanations were selected from Malle (1999) and Malle et al. (2000).

awareness, or rationality. People explain such events with *causes* as the single mode of explanation.

The proposed folk-conceptual model of behavior explanation thus includes four modes of explanation, each of which has a distinct conceptual basis and distinct functions. The theory also identifies specific types of explanation within each mode—for example, belief reasons and desire reasons that are linguistically marked (as mental states) or not. We return shortly to these more specific types.

#### Second Layer: Psychological Processes in Generating Explanations

The second layer of the folk-conceptual theory concerns the *psychological processes* that govern the actual construction of explanations (for more details, see Malle, 2004, chapter 5). In constructing explanations, people have to solve two problems (not necessarily in a fixed order). The first is to choose among the various explanatory tools they have available (i.e., four modes and various types of explanation within each). Three factors appear to determine these choices: (1) features of the behavior to be explained (e.g., intentionality, difficulty); (2) pragmatic goals (e.g., impression management, audience design); and (3) information resources (e.g., knowledge structures, perceived action context). These determinants of explanation choice are systematically related to the kinds of explanations people offer. Studies have shown, for example, that the choice between reasons and causal history explanations is primarily a function of knowledge and impression management goals; the choice of enabling factors (rather than reasons or causal histories) is a function of behavior difficulty and audience design; and the choice between causes and any of the other three explanation modes is a function of judgments about the behavior's intentionality, as well as additional

impression management goals, e.g., averting blame by portraying an ambiguous behavior as unintentional. (For evidence, see Malle, 1999; Malle, Knobe, & Nelson, 2005; Malle et al., 2000; O’Laughlin & Malle, 2002).

The second problem in constructing explanations is that people must select *specific* reasons, causes, and so on (not just a generic “belief reason” or “situation cause”), and they do so by relying on a variety of cognitive strategies. These strategies include retrieving information from knowledge structures, projection of one’s own beliefs and preferences, simulation of another person’s reasoning or experience, and, occasionally, covariation analysis. Research has just begun to examine which processes are used for which kinds of explanation modes and in response to which psychological determinants of explanation, such as pragmatic goals and information resources. (See Malle, 2004, chapter 5, for an initial, qualitative exploration.)

### Third Layer: Linguistic Expression

The third layer of the theory identifies the specific *linguistic forms* speakers have available in their language to express behavior explanations. People can exploit these linguistic forms when using explanations as a tool of social influence. For example, to distance themselves from an agent’s reasons, observers explicitly mark those reasons as the other person’s mental states (e.g., Why did she refuse dessert?—“Because *she thinks* she’s been gaining weight” vs. “Because she’s been gaining weight”: Malle et al., 2000). Similarly, when people watch a video displaying them from a third-person perspective, they seem to distance themselves from their own reasons by increasing the use of mental state markers (Malle, Heim, & Knorek, 2005.).

We now apply this model of explanations to a consideration of temporal aspects of explanation. An analysis of such temporal factors must take all three layers of explanation into account. The conceptual level enumerates the different modes and types of explanation that can vary with changing temporal perspective or passage of time. The psychological process level surveys the cognitive routes people take to construct explanations, permitting us to examine in detail how explanations of past, recent, and future events may be differentially constructed. Finally, the linguistic level alerts us to the variations in language that might reflect either underlying psychological processes or social functions and practices that govern explanations.

## Temporal Aspects of Explanations

We can distinguish at least two temporal aspects of explanation. First, how do explanations of a given event change as time passes after the event? For example, does a person’s immediate explanation of a transgression differ from an

explanation offered years later? Second, do people offer only explanations of past and present events (that actually occurred), or do they ever explain future events? If so, do such future explanations vary either in the modes of explanation they employ or the cognitive processes from which they are constructed?

Traditional research is available only on the first question, so we begin with a brief review of this research.

### *Do Explanations Change Over Time?*

#### Classic Research

About a dozen studies in the social-psychological literature discuss effects of time on attribution patterns. We set aside four studies that concern not behavior explanations but trait inferences (i.e., the fundamental attribution error; Burger, 1991; Truchot, Maure, & Patte, 2003; Wright & Wells, 1985) or estimations of contributions to joint outcomes (Burger & Rodman, 1983). Of the remaining studies, five show a pattern of actors increasing person attributions for successes but increasing situation attributions for failures, which the authors consider evidence for a time-dependent strengthening of the self-serving bias (Burger, 1985, 1986; Burger & Huntzinger, 1985; Moore, Sherrod, Liu, & Underwood, 1979; see also Burger & Pavelich, 1994). Two additional studies showed an increase of situational attributions for interpersonal interactions (Funder & Van Ness, 1983; Miller & Porter, 1980), and one showed no change (Frank & Gilovich, 1989), but it is unclear whether the events explained (interpersonal interactions) were perceived as undesirable, akin to failures. Finally, one study documented an increase of stable attributions for actors, a finding that is compatible with the previous ones in that the stable factors were typically positive (e.g., desirable personality traits).

We cannot draw compelling conclusions from these findings. One difficulty in generalizing the results is that in most studies the specific outcomes (success or failure) were manipulated by the experimenter, and participants essentially had to explain a fabricated reality (Burger, 1986; Burger & Huntzinger, 1985; Sanna & Swim, 1992). It may not be surprising that some number of participants in the failure condition attributed this surprising outcome to situational circumstances, because the failure feedback may have been discordant both with their experience in the experiment and their personal performance history. More important, studies that showed a reasonably consistent time effect focused on achievement outcomes, which are not the dominant events people explain (Malle & Knobe, 1997b). Indeed, the pattern of self-serving change over time disappears when we expand the sample of events explained. In a meta-analysis of the actor-observer asymmetry in attribution (Malle, 2005; see also Malle, in press), 102 studies assessed attributions immediately after the event, and 70 studies assessed delayed attributions. If a self-serving bias increases over time,

we would expect that for positive events actors should increase person attributions (weakening the actor-observer asymmetry), and for negative events they should increase situation attributions (strengthening the actor-observer asymmetry). There was a trend showing that actors increase their person attributions for positive events, but the same trend held true for negative events, which does not support the hypothesis that the self-serving bias increases over time.

Perhaps the most serious limitation of traditional studies on time and attribution is that there is no theoretical model that predicts the findings (or its exceptions). Why would people increase situational attributions, why would they decrease them, or why would they do both, depending on the outcome? Why the person-situation dimension rather than the stability dimension or the other way around? As in other instances (e.g., Knobe & Malle, 2002; Malle et al., 2000), there are no readily available theoretical principles that predict variations in person-situation attributions as a function of critical variables. The only candidate theory, Kelley's covariation model, doesn't actually predict person-situation attributions very well (Ahn et al., 1995; Malle, 2002), but that aside, the model has little to say about self-serving time changes in attributions. Covariation considerations predict changes in attributions only when new information becomes available (e.g., a systematic increase in consistency information, leading to a situational shift; Lau, 1984), but the whole notion of time effects in attribution presumes a *lack* of new information; otherwise, time itself wouldn't be the critical force of change.

A possibly more fruitful alternative approach is to investigate time effects on behavior explanations across the full range of unintentional and intentional events, taking into account the entire set of explanatory tools people bring to the task and considering the specific psychological determinants that govern such explanations. The following section develops this approach in more detail and, in particular, offers testable hypotheses about time effects on explanation. Our major claim is that there is no single psychological process that mediates between passing time and changes in explanation. Rather, time can affect explanations in different ways, depending on the specific process that changes over time.

### Folk-Conceptual Analysis of Explanations Across Time

Our approach relies on the folk-conceptual theory of explanations, which identifies the modes and types of explanations people use and the psychological processes that determine this use. If we can identify the psychological processes that actually change over time, then we can rely on known relationships between these processes and modes of explanation to predict how explanations change over time. There is no perfect, and certainly no one-to-one, relationship between psychological processes and explanatory tools, but there is growing evidence that systematic relationships exist (Malle, 2004, chapter 5). To make our

point, we focus on explanations of *intentional* behavior, both because they demand more complex explanatory tools and because the strength of the folk-conceptual approach is most salient in this domain. (For considerations of unintentional behaviors, see Malle, 2004.)

The three determinants of choosing modes and types of explanation introduced earlier are *behavior features*, *pragmatic goals*, and *information resources*, and we can explore time effects for each of them.

1. *Behavior features* that are critical for explanatory choices include at least two. Intentionality determines whether explainers will choose cause explanations (for unintentional events) or other explanation modes (for intentional behavior); behavior difficulty determines whether explainers will choose enabling factor explanations or other explanations of intentional behavior. These behavior features can be set aside for our analysis, however, because they must remain constant across time when we examine temporal change in explanations. Unless people explain the *same* behavior at two different times (with the same intentionality and difficulty), we cannot assess time effects; rather, we would assess effects of new information.

2. *Pragmatic goals* partially determine the choice between reasons and causal history of reason (CHR) explanations for intentional actions. One important pragmatic goal is impression management. When trying to make the agent (self or other) look good and rational for having performed a positive action, explainers typically increase their use of reasons relative to CHRs (Malle et al., 2000); by contrast, when limiting blame for negative actions, explainers typically increase their use of CHRs relative to reasons (Nelson & Malle, 2005). Another important pragmatic goal is parsimony in conversation (Grice, 1975). For example, when explaining a whole series of actions to an audience, explainers tend to decrease their use of reasons because there are too many distinct reasons to mention (potentially a different one for each action in the series). As a result, explainers tend to increase their use of CHRs, because a single CHR can precede a variety of different reasons and thereby provide a succinct explanation for the complex series of events (O’Laughlin & Malle, 2002). For example, a mother who was asked why she goes shopping several times a week replied: “Because I have three children.” The series of shopping acts is parsimoniously explained by offering this causal history of having three children because it underlies the variety of specific reasons she has for shopping each individual time (e.g., buying more milk, a new supply of diapers, or a special carpet cleaner for crayon stains).

The question is whether these processes of impression management and conversational parsimony change over time. Impression management, we suggest, may not be rigidly connected with passing time, but more often than not, impression management pressures for a given action will decrease over time because the action was performed by a “past self,” whereas impression management typically applies to the “present self.”<sup>1</sup> (This point applies to observer



explanations just as to actor explanations but far more frequently to the latter.) This process would therefore predict a decrease in reasons for positive actions and a decrease of CHR for negative actions.

Conversational parsimony per se is unlikely to change over time, but there is one condition under which passing time can elicit parsimony considerations: when the initial action has been repeated and therefore, at a later time, is explained as part of a whole series of actions. Perhaps we don't want to consider this an explanation of truly the *same* action, but if we do, time will increase the demand for parsimony and thereby the likelihood of CHR explanations.

3. *Information resources* represent the broadest determinant of explanatory choice. They are anchored by stable knowledge structures at one end and context-immediate information access at the other, with additional resources in between.

Stable knowledge structures are unlikely to change over time, so to the extent that explanations are based on these structures, we predict no notable change of explanations over time.

Context-immediate information access, at the other extreme, refers to information that is available right there at the time of acting. To analyze its impact profile over time, we need to distinguish the actor perspective from the observer perspective. From the actor perspective, this kind of access will yield information about facts, affordances, or constraints of the immediate context, which will translate into reasons and especially *unmarked belief reasons*—beliefs that are not explicitly marked by a mental state verb (e.g., “Why are you watering the plants?”—“Well, [I see that] their soil is dry!”) To the extent that the actor's initial explanation is guided by such context-immediate information, a later explanation (after time has passed) may well retrieve the same reason content from memory but will now mark the reason with a mental state verb (“*I saw* that their soil was dry”).<sup>2</sup> This predicted increase in markers is supported by the finding that episodic memories over time shift toward a third-person visual perspective (Nigro & Neisser, 1983) and that actors in third-person perspective increase their use of mental state inferences (Malle, Heim, & Knorek, 2005; see Moore et al., 1979, for a similar argument).

From the perspective of an observer who is present at the time of action, immediate information may similarly refer to contextual facts and constraints, but these have to be translated into ascriptions of belief reasons to the agent. In addition, immediate information also includes the perception of the “moving body,” which is likely to elicit ascriptions of desires and intentions (Baird & Baldwin, 2001). Ascriptions of both belief and desire reasons (which may be based at least in part on “simulation” processes—e.g., Gordon, 1986, 1992; Goldman, 1989, 2001) will typically be marked with mental state verbs (“She thought”; “He wanted”) to identify states of “that mind over there.” Over time observers will retrieve far less of this immediate (often perceptual) information, which may reduce their ability to ascribe mental states to the agent (perhaps be-

cause of limited simulation opportunities), and in that case observers may resort increasingly to causal history of reason explanations. Thus, for actors, the fading of context-immediate information should primarily affect the formulation of their explanations (increasing the use of mental state markers), but for observers, this fading should decrease their overall ascriptions of reasons.

A final type of information resource is the actor's direct recall of her reasons for acting. In some number of intentional actions, actors deliberate about their decision to act (i.e., they consider reasons for or against courses of action), and in those cases an actor who explains her action can easily recruit the very reasons (beliefs, desires) for which she acted (Brewer, 1994; Russell & D'Hollosy, 1992). Such direct recall may fade, at first weakly, but more noticeably with passing time, especially for insignificant actions. An example: One of us bikes home up a hill on a route that has a fork in the road such that a decision has to be made to take the shorter and steeper path or the longer and more gently sloped path. It is virtually impossible to recall the reasons for taking the longer, flatter route this day a year ago, even though on that day, especially within minutes of the decision, the reasons were almost certainly directly recalled. This fading of direct recall will undermine the actor's offering of reasons, especially context-sensitive reasons, which are typically beliefs.

Even without a parallel analysis of explanations for unintentional behavior, we feel confident in our conclusion that time has no uniform effect on explanatory choices. One must take into account the various determinants of explanations and the perspectives of actor and observer; only then can mere time passing be translated into actual explanatory changes.

### *Do People Explain Future Events (and How)?*

The second important temporal aspect of behavior explanations is not oriented toward the past but the future. The question we would like to explore is to what extent people explain future behaviors, and if they do, how.

It may appear somewhat paradoxical to wonder whether people *explain* future behavior. After all, explanation is often defined as a clarification of events that already happened (e.g., Heider, 1958; Hempel & Oppenheim, 1948; Salmon, 1989). This is a reasonable position, but we still would like to examine whether the same *processes* that give rise to explanations of past behaviors can sometimes generate what look like explanations of future behaviors. In the end, these explanations may come close to *predictions* of future behaviors, which will make the tenor of our subsequent section yet more plausible: that predictions of future events are closely tied to explanatory processes.

The best case for explanations of future events can be made in the domain of *reason explanations* for intentional behavior. The agent's deliberation process

(rudimentary as it may be in some cases) consists of the consideration of reasons, primarily desires for certain outcomes and beliefs about ways to achieve those outcomes. This reasoning process weighs the motivational and rational grounds for forming various intentions, and the reasoning process ends when the agent settles on an intention to act for particular reasons (Malle & Knobe, 2001; Mele, 1992, 2003). Schematically:

Reasons [beliefs, desires] → intention to *A* → action *A*.

The folk conception of intentional action and its underlying reasoning process is such that an explanation of the *completed action*, were it performed, would refer to the agent's reasons for which she chose that action; and so would an explanation of the *intention* to act, even if the action has not yet been performed (Malle, 1999). Nothing guarantees that actions are performed. Lara may deliberate and then decide to go to the movies tonight, but something comes up and she doesn't go. Nonetheless, at the time of the decision, and even later on, we can explain why she decided to go to the movies by pointing to her reasons—for example, because it was Thursday night and the film was about to leave the theater.

Every intention (e.g., "I am going to *A*") is future directed. This is nicely driven home by the fact that in many languages the future tense either stems from or is closely related to intention verbs, as in the English "I will . . ." (Bybee, Perkins, & Pagliuca, 1994). And because intentions are explained by the reasons for which the intention was formed, people explain at least some future events—those that are intended (but not yet performed) actions. Consider the following examples (from students' diaries and audiotaped conversations; Malle & Knobe, 1997b; Malle et al., 2005):

*Since I'll be gone* she's going to take Lara out instead.

I'm not going to apply for "live-out" 'cause I don't want to step on anybody's toes.

Reason explanations dominate these explanations of future events, though under rare circumstances, an intended future action might be explained by reference to a causal history of reason factor. We could not find a real example in our database of explanations, but the following constructed example appears plausible:

She intends to buy, not rent, *because she is from a rich family*.

Perhaps more likely would be a reformulation as a pure prediction: "She *will* buy, not rent, because she grew up in a rich family." CHR explanations of intended future actions are rare because, when explicitly announcing an agent's intention, the speaker is normally expected to clarify the agent's mo-

tives for intending to act, which requires reason explanations. This demand for reason explanations should be particularly true for actors explaining their own intended future actions. Announcing one's own intention comes with a strong conversational expectation to justify and clarify that intention with one's reasons for adopting it; CHR explanations do not meet this expectation. Rarely would an actor say, "I intend to buy, not rent, because I grew up in a rich family."

Enabling factor explanations cannot be used to explain intended future events, because these explanations account for the successful performance of actions, not for their motives and intentions. We can, however, *predict* that somebody will likely succeed in clearing 7 feet in high jumping because he or she has trained for 6 months or that a team will succeed in finishing a project because it started early enough.

Similarly, neither actors nor observers can explain future *unintentional* events; they can only predict them. Consider the following examples:

When I started taking those math classes I thought, "This is going to be hard  
*'cause I really hate math.*"

I will hopefully get alumni status, *'cause when you affiliate to another chapter,  
that's what you get.*

In each of these cases, the *because* clause offers a cause that makes it likely for the future event to obtain, and in that sense most would call this a prediction, not an explanation. In addition, the *because* clause has a second function: It provides *backing* or *evidence* for the claim that the future event will obtain. Claim backings are most transparent when the *because* clause does not cite any causal factor at all, as in the following example:

Joe is going to be angry *because I have seen him react to conflict before.*

Clearly, Joe's anger will not be caused by the speaker's having seen Joe react that way before. What the speaker does in the *because* clause is back up, provide evidence for, the prediction that Joe will be angry. Davies (1979) proposed a criterion that distinguishes clear cases of claim backing from explanation: Claim backings can be formulated only in the canonical order of "event + because," whereas explanations can also be formulated in the inverse order, "because + event." For example, it makes no sense to say, "Because I have seen him react to conflict before, Joe is going to be angry." The infelicitous use of the inverse order exposes the *because* clause as a pure claim backing.<sup>3</sup>

In sum, we suggest that *because* clauses paired with future event statements can be explanations of intended future events, predictions made on the basis of identified causes, or predictions with pure claim backing. This range already suggests a close relationship between explanations and predictions. We now ex-

plore more thoroughly the topic of prediction and inspect once more the particular role of explanatory processes in predictions.

## The Cognitive Apparatus of Prediction

### *Existing Prediction Literature*

To date, psychological studies of prediction have largely used likelihood estimates as a method of shedding light on this phenomenon. Likelihood estimates are studied under different names (e.g., subjective probability) and have been related to a wide variety of judgment and decision-making topics such as choice (see Dawes, 1998, for a review), alternative outcomes (Windschitl & Wells, 1998), anticipated regret (e.g., Zeelenberg, 1999), risk (e.g., Slovic, 2000), contrast effects in social judgments (e.g., Schwarz & Bless, 1992), and affective forecasting (e.g., Gilbert, Lieberman, Morewedge, & Wilson, 2004). In addition, virtually all prediction studies have some standard of accuracy against which predictions are examined. This accuracy focus has allowed researchers to uncover heuristics and biases people use when making predictions (e.g., Kahneman & Tversky, 1973; MacDonald & Ross, 1999; Peterson & Pitz, 1986) and to compare human thinking to regression models or other statistical methods of prediction (e.g., Kahneman & Tversky, 1982a; Snook, Canter, & Bennell, 2002).

Compared to work on accuracy of predictions, far less research has focused on *how* humans are able to predict in the first place. Studying the accuracy of predictions is only an indirect method of examining the processes underlying prediction because accuracy is a result of comparing a predicted outcome to what actually happened, not a focus on how the person generated or evaluated that prediction in the first place. In fact, the underlying processes of prediction have remained virtually unstudied since the work by Kahneman and Tversky (1973), who provided an initial answer by arguing (using accuracy data) that humans make predictions using representativeness (i.e., stereotypical associations and/or expectations).

### *Types of Predictions*

#### The Simulation Heuristic

Some years after their proposal that prediction relies on representativeness, Kahneman and Tversky (1982b) developed five “scenarios” (or kinds) of mental simulations. Three of these focused on mentally simulating possible futures, and the remaining two pertained to mental simulation of the past. The scenar-

ios about the past are currently studied under the heading of “counterfactual thinking” in social and cognitive psychology, but less research has examined the first three scenarios concerning future thinking. The future-oriented scenarios were called (1) *prediction*, (2) *assessing the probability of a specified event*, and (3) *assessing conditioned probabilities* (Kahneman & Tversky, 1982b, p. 202). Briefly, *prediction* concerns how individuals generate predictions, such as that two strangers will “get on famously” (Kahneman & Tversky, 1982b, p. 202). *Assessing the probability of a specified event* concerns a specified target event and “the ‘ease’ with which this target state can be produced” (Kahneman & Tversky, 1982b, p. 202). An example of this scenario would be: What is the likelihood of an American invasion of Iran in the next decade? Here, the invasion is the specified target state, and, according to Kahneman and Tversky, a person evaluates the ease with which the target state can be imagined or simulated. Most studies of prediction focus on this kind of scenario because the operationalization of prediction as likelihood estimates or subjective probabilities is amenable to studying the ease with which specified target outcomes can be imagined. The specification of the outcome also usually allows for testing of these estimates against some standard of accuracy. Finally, *assessing conditioned probabilities* refers to simulating future events under initial conditions. For example, a person might be asked: “If civil war breaks out in Saudi Arabia, what are the likely consequences?” (Kahneman & Tversky, 1982b, p. 202). This task requires a person to assume a state of affairs that may or may not be consistent with reality (e.g., civil war in Saudi Arabia might or might not be happening currently). After accepting the conditional, it then appears that a person can focus on the “ease” with which the specified target event can be imagined.

#### A New Fourfold Typology of Prediction

We believe that the simulation heuristic provides a good starting point for distinguishing among types of predictions, but by itself it does not cover the full range of prediction types or processes. When examining the structure of predictions, it appears that virtually all predictions have three elements: (1) the outcome or event that is predicted, (2) the starting conditions in light of which the prediction is made, and (3) some sort of connection between outcome and starting conditions. This tripartite structure of prediction can be depicted as  $S \rightarrow O$ , wherein  $S$  represents the starting conditions,  $O$  represents the outcome, and  $\rightarrow$  represents some connection that “leads to” the outcome given the starting conditions. This conceptual structure serves as the cornerstone for the new typology of prediction that we propose and develop here. The typology focuses on whether the starting conditions  $S$  and outcome  $O$  are explicitly specified or left unspecified. These considerations result in a  $2 \times 2$  model encompassing four types of prediction (see Table 11.2).

Table 11.2. Fourfold Typology of Prediction.

Starting Conditions	Outcome	
	Specified	Unspecified
Specified	Fully Specified <b>SO</b>	Starting Conditions Specified <b>S</b>
Unspecified	Outcome Specified <b>O</b>	Unspecified <b>U</b>

*Fully Specified Predictions (SO).* Type SO predictions have specified starting conditions and a specified outcome. In language, this kind of prediction might be stated as "Given that [S], how likely is it that [O]?"

*Outcome-Specified Predictions (O).* Type O predictions have unspecified starting conditions and a specified outcome. In language, this kind of prediction might be stated as "How likely is it that [O] will happen?"

*Starting-Conditions-Specified Predictions (S).* Type S predictions have specified starting conditions and an unspecified outcome. In language, this kind of prediction might be stated as "Given [S], what will happen?"

*Unspecified Predictions (U).* Type U predictions contain both unspecified starting conditions and an unspecified outcome. In language, this kind of prediction might be stated as "What will happen?" Such formulations of predictions, however, may be inconsistent with conversational norms (Grice, 1975), and listeners might therefore move to specify at least one of the elements, thereby shifting to a Type SO, Type O, or Type S prediction.

#### A Comparison of the Two Models

The fourfold typology of prediction subsumes and clarifies the three scenarios of prediction offered by Kahneman and Tversky's (1982b) analysis. The *prediction* scenario is analyzed as having specified starting conditions but unspecified outcomes (Type S). *Assessing the probability of a specified event* can be either a Type SO or a Type O prediction, as both these types involve specified outcomes. Our new typology distinguishes, however, between the two contexts in which a specified event can be evaluated: with or without a specified set of starting conditions. This distinction enables the investigation of cognitive strategies for identifying starting conditions when relevant (i.e., for Type O predictions). Finally, *assessing conditioned probabilities* is analyzed as a Type SO prediction, which is fully specified.

## *From Simulation to Multiple Cognitive Processes*

We propose one more addition to Kahneman and Tversky's (1982b) analysis of prediction. Kahneman and Tversky subsumed all forms of prediction under the simulation heuristic, relying on the intuition that to predict the future people have to simulate it. But much work in cognitive science over the past 20 years strongly suggests that simulation is a unique cognitive process, and not all prediction relies on this process. Simulation is best considered a process of running an analogous "model" of the to-be-predicted domain and, akin to a computer simulation, arriving at a likely outcome that is then formulated as a prediction about the domain (e.g., Gordon, 1986; Goldman, 1989). Simulation understood this way is very different from theoretical or rule-driven inferences, which are often the basis of prediction (Gopnik & Wellman, 1994; Rips, 1994). One can easily make a prediction about a future event without literally simulating the steps that lead up to this event; one might, instead, use past experience to take a best guess or rely on a general rule to predict the event. These distinct processes underlying prediction must somehow be incorporated into a theory of prediction.

Regardless of whether starting conditions and outcomes are explicitly stated or left implicit, we assume that all cases of prediction can be modeled by the general  $S \rightarrow O$  structure, which is the basis for the preceding fourfold typology. But a theory of prediction must specify in more detail the nature of the connection operator (the "leads to" connections). We propose that the connections between  $S$  and  $O$  can be characterized by two generic strategies—simulation and inference—and we identify three different forms that the inference strategy can take. Moreover, we propose that all the connections are, to varying degrees, explanatory in nature. That is, they provide some explanation of how the starting conditions lead to the outcome. As a result, we propose a total of four *explanatory strategies* (simulation and three forms of inference), which all specify, to varying degrees, the connection or relationship between starting conditions and a predicted outcome.

### *Four Explanatory Strategies for Prediction*

The proposed explanatory strategies characterize subjective human experience and can be construed as cognitive tools that people have at their disposal. The explanatory strategies described here were developed from examining various literatures in philosophy (see Reschler, 1997, for a review) and cognitive, social, and developmental psychology, such as behavior explanation (e.g., Malle, 1999, 2004), inductive inference and rule-based inference (see Sloman, 1996, for a review), causal inference (e.g., Cohen, Rundell, Spellman, & Cashon,



1999), and theory of mind (e.g., German & Leslie, 2000; Gordon, 1986). The four explanatory strategies gleaned from these reviews are: (1) noncontradiction, (2) association (induction), (3) derivability (deduction), and (4) scenario generation (simulation).

*Noncontradiction* is a minimal connection between *S* and *O* such that there is no apparent contradiction between the outcome and both the starting conditions and other relevant knowledge. For instance, when predicting that a party will be wild and crazy, a participant in one of our studies provided the following explanation: "There'll be some crazy happenings because, I don't know, it's a party and stuff like that can happen." That is, the outcome does not seem to contradict the assumed starting conditions and other relevant knowledge, lending some plausibility to the predicted event. In addition to providing a minimal requirement for predicting events at above-zero probabilities, noncontradiction might be useful for situations in which a person has little knowledge about the domain. In this case, the person can at least predict that an outcome is "somewhat likely" because it does not appear to contradict what else is known. To be sure, noncontradiction is not what one might characterize as a full-fledged prediction, because the mechanism through which the outcome is believed to result remains unspecified; yet this strategy provides the necessary foundation for developing full-fledged predictions, as any predicted outcome must be possible given the starting conditions and other relevant knowledge.

*Association* draws a correlational or stereotypical connection between the starting conditions and the outcome. This strategy includes representativeness as Kahneman and Tversky (1973) defined it and was inspired by literature on inductive inference (see Sloman, 1996, for a review). This strategy is qualitatively different from noncontradiction in that association at least hints at a causal mechanism (or process) through which the outcome ensues given the starting conditions. In this manner, association begins to resemble full-fledged prediction. Interestingly, association-based predictions are comparable to causal history of reason explanations (see the second section of this chapter) in that the association strategy, too, mentions factors that led up to or lay in the history of an event without offering details about the mechanism by which the event came about. Nevertheless, association-based predictions are broader than CHR explanations because they are believed to be domain-general, not merely applicable to intentional actions. Examples of associative explanations include "The Los Angeles Lakers always beat the Chicago Bulls, so this time will not be any different" (empirical association) or "good students don't cheat, so she won't" (stereotypical association). It should be noted that one can use the same strategy for nonhuman objects, such as physical or natural events: "Old Faithful spouts everyday, so today will be no different."

*Derivability* provides a deduction of the outcome from the starting conditions together with a rule, law, or psychological principle. This explanatory strategy mirrors rule-based inference (Sloman, 1996) and the *theory theory* position in

theory of mind research (Gopnik & Wellman, 1994). This strategy is different from noncontradiction and association in that derivability actually identifies a causal mechanism or process that connects the starting conditions and outcome. Reason explanations of intentional action resemble derivability-based predictions in that they posit a particular mechanism (e.g., belief-desire reasoning leading to intentions and actions) and thus connect specific starting conditions (particular belief or desire reasons) with the outcome (the action). In the realm of physical events, derivability-based prediction uses laws or rules such as gravity or force to provide a connection between starting conditions and outcome, thus going beyond past co-occurrences of two events or stereotypical associations. An example of a derivability explanation for human behavior would be: "Christine will invite Diego to the party because she is interested in him." Here, Christine's interest in Diego serves as the deductive premise from which it is very likely that she will ask him to the party, in light of the general principle that people pursue that in which they are interested. An example of derivability in a physical domain would be: "Given that the structural supports for the bridge are weak, it will probably collapse."

*Scenario generation* provides a story or simulation of the events unfolding over time, as in a film or narrative. This strategy was inspired by Pennington and Hastie's (1993) story model of decision making, simulation theory in the theory of mind literature (e.g., Gordon, 1986), and the underlying theme of Kahneman and Tversky's simulation heuristic. This strategy is distinguishable from the inference strategies described earlier because scenario generation fills in concrete steps between the starting conditions and the outcome, typically relying on a wealth of domain-relevant knowledge in order to create complex causal chains. To simulate a predicted event, the person creates an explanatory structure that resembles a story, often containing a plot (the starting condition-outcome connection) and subplots (starting condition-starting condition connections). Take the example of a person predicting the victory of a sports team by envisioning a scenario in which the team trails in the first half, rallies in the second, and beats the opponent at the buzzer.

We now provide initial empirical support for the distinction among these four strategies. In particular, we examined whether these strategies (1) can be identified in people's introspections about how they generate predictions and (2) have systematic relationships to the four types of prediction we proposed earlier and to different content domains of prediction.

### *Explanatory Strategies and Prediction Types: Initial Support for the Model*

We first conducted a study that examined the efficacy of the proposed model of explanatory strategies and prediction typology in characterizing actual predictions that people encounter, generate, and evaluate in daily life. We used a diary

methodology in which nine participants recorded predictions over a 4-day period that either they or someone else had generated. Participants were required to record 12 predictions during this interval but were allowed to record as many as 20. For each recorded prediction (their own or another person's), participants were asked to speculate on what was going on in their own minds (or what they believed was going on in the other person's mind) just before making the prediction.

Of the 180 records across participants, only 130 were valid predictions. The remaining entries were either desires (e.g., "I want to go the game"), wonderings (e.g., "am I hungry?"), or statements (e.g., "I don't like work"). Of the 130 valid records, 106 (81.5%) were coded as specified-outcome (Type O) predictions, 16 (12.3%) were coded as fully specified (Type SO) predictions, and 8 (6.2%) were coded as specified-starting-condition (Type S) predictions. None of the valid records were codable as unspecified (Type U) predictions.

A total of 116 records provided valid responses to the question of what was going on in the mind of the predictor just before making the prediction. Of these, 82 (70.6%) were coded as *derivability* predictions (e.g., "DeShaun Foster isn't playing, so their offense will be weak [and, therefore, their team will lose]"; "I've done all the right things on time [so I will get a good grade]"). Furthermore, 20 (17.2%) were coded as *association* predictions (e.g., "I beat him by 22 yesterday"; "I've saved a person on Sunday for the past 3 weeks [so I will save a person this week too]"), 11 (9.5%) were coded as *noncontradiction* predictions (e.g., restating the predictive thought or simply agreeing with it; e.g., "[there will be some crazy happenings at a party] because stuff like that happens"), and 3 (2.5%) were coded as scenario generation (e.g., "I can see us getting down by a few points early but then coming back strong in the second half").

The relative proportion of the types of predictions and explanation strategies should be interpreted with caution. With the relatively small number of predictions chosen (20 each across nine people), these proportions might not be representative. Additionally, the proportions may be influenced by the diary methodology. For instance, specified-outcome (Type O) predictions might be the most frequent in this sample because starting conditions for familiar outcomes are often left implicit. Relative proportions of prediction types and explanatory strategies notwithstanding, these results do suggest that the proposed prediction typology characterizes people's naturally occurring predictions quite well<sup>4</sup> and that people rely on all four proposed explanatory strategies when constructing or analyzing predictions.

Our next step was to examine more systematically whether explanatory strategies vary by type of prediction and by content domain. To this end, we interviewed 11 participants and asked them to generate or evaluate a total of nine predictions. We created the prediction question so that they fell into a 3 (domain)  $\times$  3 (prediction type) design. Predictions were of Type SO, Type O, and Type S. Domains included (1) natural or physical events (e.g., occurrence of tor-

nadoes in a region, the spouting of geysers), (2) the actions of strangers (e.g., the United States Congress, the university basketball team), and (3) the actions of intimates (i.e., close others—friends, family—and the self). An example of a prediction question is: “How likely is it that the United States government will find Osama bin Laden?” (Type O prediction in the strangers domain). Participants’ responses to each prediction question were coded into one of the four explanatory strategies (noncontradiction, association, derivability, or scenario generation).

The data were analyzed using a repeated measures analysis of variance on the number of times participants provided a specific explanatory strategy by domain and prediction. Of note, there was a main effect for explanatory strategies such that scenario generation-based predictions ( $M = .13$ ) were listed least frequently and association-based predictions ( $M = .50$ ) were listed most frequently. Noncontradiction-based predictions ( $M = .14$ ) and derivability-based predictions ( $M = .22$ ) showed intermediate frequency ( $M = .22$ ).

There was also a significant interaction between explanatory strategies and domain. The results, displayed in Figure 11.1, are somewhat complex but show distinct patterns for the four explanatory strategies. Noncontradiction-based predictions tended to be provided significantly more in the nature domain ( $M =$

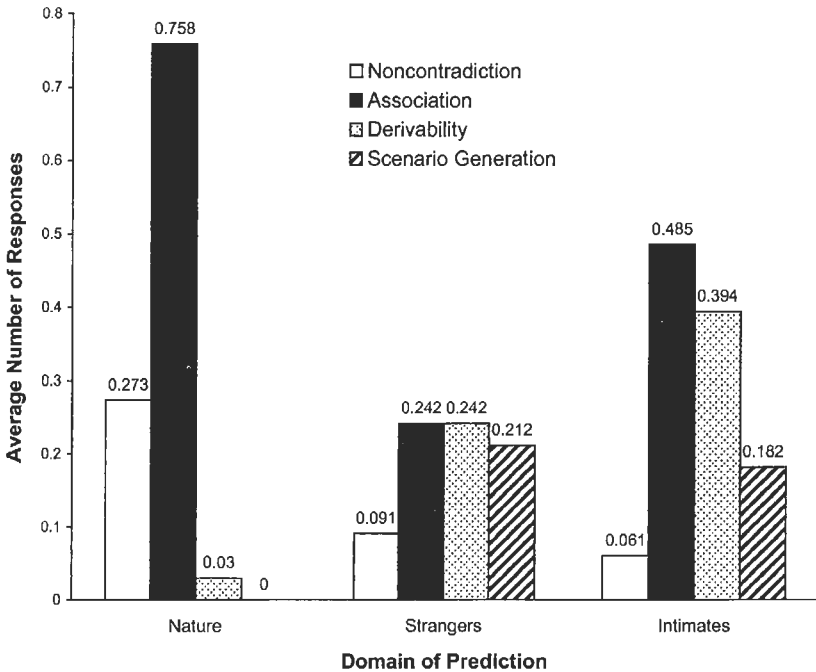


Figure 11.1. Predictions based on different explanatory strategies function as a domain of prediction.

.27) than in the human domain ( $M = .08$ ). Association-based predictions were also listed significantly more often in the nature domain ( $M = .76$ ) than in the human domain ( $M = .36$ ) and significantly more often for intimates ( $M = .49$ ) than for strangers ( $M = .24$ ). Derivability-based predictions were provided significantly more often in the human domain ( $M = .32$ ) than in the nature domain ( $M = .03$ ). Finally, scenario generation predictions occurred significantly more often in the human domain ( $M = .20$ ) than in the nature domain ( $M = 0$ ).

These results suggest that people use explanatory strategies differently depending on the domain of prediction. Noncontradiction and association strategies appear to be preferred over derivability and scenario generation strategies in the nature domain, presumably because it is easier to predict natural events by either associating past occurrences or simply finding no good reason why it should not happen again in the future. Derivations and scenarios, in contrast, might be harder to generate in the nature domain because undergraduate students normally do not have extensive knowledge about such topics as weather patterns, behavior of sharks, or the functioning of geysers.

In the human domain, noncontradiction-based predictions are close to zero, but, for strangers, the three other explanatory strategies occur equally often. This result suggests that when predicting human action for those with whom the predictor has little familiarity, elaborated inference strategies and simulation are equally preferred. Yet association explanations show a further familiarity effect such that they are used even more to predict the behavior of intimates. This finding makes sense to the extent that people have more past history information about close friends, family, and the self than about sports teams or congressional bodies. The lack of a difference between strangers and intimates in using derivability and scenario generation suggest that these two strategies work well for predicting the behaviors for people in general—well known or not.

In sum, these initial studies provide some basis for adopting a model of prediction that distinguishes between multiple types of prediction and multiple explanatory strategies in service of prediction. Moreover, the use of these strategies appears to vary with task, domain, and knowledge. In the concluding section, we relate the folk-conceptual theory of explanation to this new model of prediction and ask more generally to what extent prediction and explanation are served by the same psychological processes.

## Summary and Conclusion

We have introduced a model of prediction that postulates two kinds of cognitive processes with which people construct predictions: simulation and inference. We elaborated the category of inference by distinguishing assessment of noncontradiction, reliance on association, and derivation from principles. In the section on explanations, we introduced a number of psychological pro-

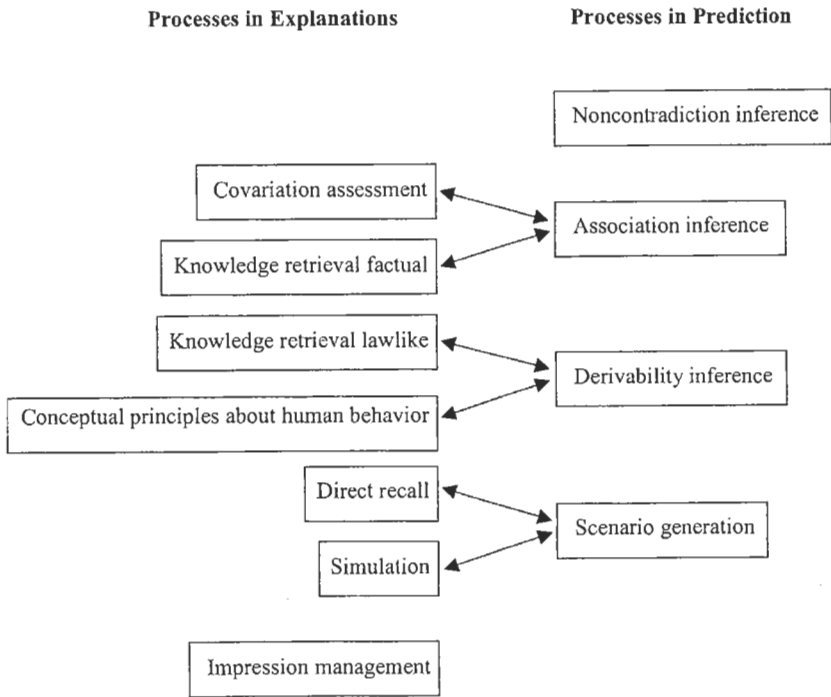


Figure 11.2. Unique and shared cognitive processes in explanation and prediction.

cesses that underlie explanations, in particular impression management, retrieval from knowledge structures, direct recall, simulation, and covariation assessment. It is now time to relate these two sets of processes. Our claim is that explanations and predictions overlap to a considerable extent because they share a number of cognitive processes—processes that support the agent’s attempt to construct the judgment at issue, whether it be an explanation or a prediction.

Figure 11.2 shows the psychological processes that we identified as underlying explanations and predictions, respectively. Some processes are unique to explanations, and at least one process is unique to predictions; but the two kinds of judgments have at least three processes in common. Their labels are different, but the cognitive operations they point to are the same.

First, covariation assessment and factual knowledge retrieval in explanations correspond to association-based prediction. In all of these cases, stored knowledge about past events and their co-occurrence supports the goal of explaining or predicting. Second, explanations based on retrieval of lawlike knowledge or reliance on conceptual principles (e.g., that reasons engender intentions) correspond to derivability-based predictions. In these cases, given information is connected with unknown information by means of laws or

principles. Third, simulation and direct recall of explanations correspond to scenario generation for predictions. Here, the person mentally goes through the actual steps (e.g., of belief-desire reasoning) that explain or predict the outcome in question.

More generally, both explanations and predictions are underdetermined by the data—that is, the available information does not favor a single explanation or prediction. For explanations we know the outcome but not the starting conditions, and for predictions we know the starting conditions but not the outcome. Events in closed physical systems may be the extreme case in which explanation and prediction are, as Hempel and Oppenheim (1948) proposed, symmetric. However, in everyday situations, in which people explain human action and experience, institutional events, and complex natural patterns, predictions are more underdetermined because the unknown (the outcome) has not yet happened. Predictions' more serious underdetermination also results in their being cognitively more demanding (Bartsch, 1998; Robinson & Mitchell, 1995; Moses & Flavell, 1990) and less accurate (Clarke & Blake, 1997).

In addition to this difference in underdetermination, the functions of explanations and predictions differ as well. The thirst for meaning inherent in explanations (Malle, 2004) is powerful and cannot be quenched by predictions alone. Predictions may provide hope for reaching an understanding (as when scientific theories make competing predictions and wait for reality to select one as correct), but it is the explanation following the outcome that provides meaning. Explanations can also settle social tension and uncertainty, whereas predictions normally cannot. Finally, predictions when paired with *intentions* can support social coordination, whereas explanations cannot by themselves solve coordination problems because these problems necessarily reach into the future.

These differences notwithstanding, this chapter has emphasized the similarities between explanation and prediction—similarities that exist primarily at the cognitive level. In attempting to master the social and physical world, the human mind relies on a toolbox of cognitive strategies and operations of which some uniquely support explanations, some uniquely support predictions, but a good number support both.

## Acknowledgments

Thanks to Lisa A. Johnson and Meghan Bean for coding the explanatory strategies in the described studies.

## Notes

1. This prediction seemingly contradicts the studies in the literature that found increasing self-servingness for explanations of achievement outcomes over time. But as we argued earlier, unintentional outcomes and intentional actions may be differentially affected by passing time. Relatedly, a self-serving bias is not a result of impression management but an opportunity to exploit ambiguity and vague memories for self-protective purposes. In fact, if participants in the achievement outcome studies were managing their impressions in front of a real audience, they might not present themselves as self-serving but as modest.

2. This situation may be comparable to a visual perspective switch in which an actor sees herself on video as if from an observer's viewpoint (famously manipulated by Storms, 1973). In a recent study we found that switching actors' visual perspective this way increases their use of mental state markers for their own belief reasons ("I thought"; "I assumed"), more than doubling their normally quite low frequency of marker use (Malle, Heim, & Knorek, 2005).

3. One can say, however, "Because I have seen him react to conflict before. *I believe that* Joe is going to be scared," making the claim and its backing transparent.

4. Although we did not observe any unspecified (Type U) predictions, recall that our discussion of the typology suggests that people may turn such predictions into one of the other three types, consequently making the expected frequency of Type U predictions very low.

## References

- Abelson, R. P., & Lalljee, M. (1988). Knowledge structures and causal explanations. In D. J. Hilton (Ed.), *Contemporary science and natural explanation: Commonsense conceptions of causality* (pp. 175–203). Brighton, UK: Harvester.
- Abramson, L. Y., Seligman, M. E. P., & Teasdale, J. D. (1978). Learned helplessness in humans: Critique and reformulation. *Journal of Abnormal Psychology, 87*, 49–74.
- Ahn, W., Kalish, C. W., Medin, D. L., & Gelman, S. A. (1995). The role of covariation versus mechanism information in causal attribution. *Cognition, 54*, 299–352.
- Andrews, K. (2003). Knowing mental states: The asymmetry of psychological prediction and explanation. In Q. Smith & A. Jokic (Eds.), *Consciousness: New philosophical perspectives* (pp. 201–219). Oxford: Oxford University Press.
- Baird, J. A., & Baldwin, D. A. (2001). Making sense of human behavior: Action parsing and intentional inference. In B. F. Malle, L. J. Moses, & D. A. Baldwin (Eds.), *Intentions and intentionality: Foundations of social cognition* (pp. 193–206). Cambridge, MA: MIT Press.
- Bartsch, K. (1998). False belief prediction and explanation: Which develops first and why it matters. *International Journal of Behavioral Development, 22*, 423–428.
- Brewer, W. F. (1994). Autobiographical memory and survey research. In N. Schwarz



- & S. Sudman (Ed.), *Autobiographical memory and the validity of retrospective reports* (pp. 11–20). New York: Springer.
- Burger, J.M. (1985). Temporal effects on attributions for academic performances and reflected-glory basking. *Social Psychology Quarterly*, 48, 330–336.
- Burger, J.M. (1986). Temporal effects on attributions: Actor and observer differences. *Social Cognition*, 4, 377–387.
- Burger, J.M. (1991). Changes in attributions over time: The ephemeral fundamental attribution error. *Social Cognition*, 9, 182–193.
- Burger, J.M., & Huntzinger, R.M. (1985). Temporal effects on attributions for one's own behavior: The role of task outcome. *Journal of Experimental Social Psychology*, 21, 247–261.
- Burger, J.M., & Pavelich, J.L. (1994). Attributions for presidential elections: The situational shift over time. *Basic and Applied Social Psychology*, 15, 359–371.
- Burger, J.M., & Rodman, J.L. (1983). Attributions of responsibility for group tasks: The egocentric bias and the actor-observer difference. *Journal of Personality and Social Psychology*, 45, 1232–1242.
- Buss, A.R. (1978). Causes and reasons in attribution theory: A conceptual critique. *Journal of Personality and Social Psychology*, 36, 1311–1321.
- Bybee, J., Perkins, R., & Pagliuca, W. (1994). *The evolution of grammar: Tense, aspect, and modality in the languages of the world*. Chicago: University of Chicago Press.
- Clarke, D.D., & Blake, H. (1997). The inverse forecast effect. *Journal of Social Behavior and Personality*, 12, 999–1018.
- Cohen, L.B., Rundell, L.J., Spellman, B.A., & Cashon, C.H. (1999). Infants' perception of causal chains. *Psychological Science*, 10, 412–418.
- Davies, E.C. (1979). *On the semantics of syntax: Mood and condition in English*. London: Croom Helm.
- Dawes, R.M. (1998). Behavioral decision making and judgment. In D.T. Gilbert, S.T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (4th ed., vol. 1, pp. 497–548). New York: McGraw-Hill.
- Frank, M.G., & Gilovich, T. (1989). Effect of memory perspective on retrospective causal attributions. *Journal of Personality and Social Psychology*, 57, 399–403.
- Funder, D.C., & Van Ness, M.J. (1983). On the nature and accuracy of attributions that change over time. *Journal of Personality*, 51, 17–33.
- German, T.P., & Leslie, A.M. (2000). Attending to and learning about mental states. In P. Mitchell & K.J. Riggs (Eds.), *Children's reasoning and the mind* (pp. 229–252). London: Psychological Press.
- Gilbert, D.T., Lieberman, M.D., Morewedge, C.K., & Wilson, T.D. (2004). The peculiar longevity of things not so bad. *Psychological Science*, 15, 14–19.
- Goldman, A.I. (1989). Interpretation psychologized. *Mind and Language*, 4, 161–185.
- Goldman, A.I. (2001). Desire, intention, and the simulation theory. In B.F. Malle, L.J. Moses, & D.A. Baldwin (Eds.), *Intentions and intentionality: Foundations of social cognition* (pp. 207–225). Cambridge, MA: MIT Press.
- Gopnik, A., & Wellman, H.M. (1994). The theory theory. In L.A. Hirschfeld and S.A. Gelman (Eds.), *Mapping the mind: Domain specificity in cognition and culture* (pp. 257–293). New York: Cambridge University Press.

- Gordon, R. M. (1986). Folk psychology as simulation. *Mind and Language*, 1, 158–171.
- Gordon, R. M. (1992). The simulation theory: Objections and misconceptions. *Mind and Language*, 7, 11–34.
- Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics: Vol. 3. Speech acts* (pp. 41–58). New York: Academic Press.
- Heider, F. (1958). *The psychology of interpersonal relations*. New York: Wiley.
- Hempel, C. G., & Oppenheim, P. (1948). Studies in the logic of explanation. *Philosophy of Science*, 15, 135–175.
- Jones, E. E., & Davis, K. E. (1965). From acts to dispositions: The attribution process in person perception. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 2, pp. 219–266). New York: Academic Press.
- Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review*, 80, 237–251.
- Kahneman, D., & Tversky, A. (1982a). Intuitive prediction: Biases and corrective procedures. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 414–421). New York: Cambridge University Press.
- Kahneman, D., & Tversky, A. (1982b). The simulation heuristic. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 201–210). New York: Cambridge University Press.
- 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.
- Knobe, J., & Malle, B. F. (2002). Self and other in the explanation of behavior: 30 years later. *Psychologica Belgica*, 42, 113–130.
- Lau, R. R. (1984). Dynamics of the attribution process. *Journal of Personality and Social Psychology*, 46, 1017–1028.
- MacDonald, T. K., & Ross, M. (1999). Assessing the accuracy of predictions about dating relationships: How and why do lovers' predictions differ from those made by observers? *Personality and Social Psychology Bulletin*, 25, 1417–1429.
- Malle, B. F. (1999). How people explain behavior: A new theoretical framework. *Personality and Social Psychology Review*, 3, 23–48.
- Malle, B. F. (2001). Folk explanations of intentional action. In B. F. Malle, L. J. Moses, & D. A. Baldwin (Eds.), *Intentions and intentionality: Foundations of social cognition* (pp. 265–286). Cambridge, MA: MIT Press.
- Malle, B. F. (2002). Verbs of interpersonal causality and the folk theory of mind and behavior. In M. Shibatani (Ed.), *The grammar of causation and interpersonal manipulation* (pp. 57–83). Amsterdam: Benjamins.
- Malle, B. F. (2004). *How the mind explains behavior: Folk explanations, meaning, and social interaction*. Cambridge, MA: MIT Press.
- Malle, B. F. (2005). *The actor-observer asymmetry in causal attribution: A (surprising) meta-analysis*. Manuscript in preparation.
- Malle, B. F. (in press). Self-other asymmetries in behavior explanations: Myth and reality. In M. D. Alicke, D. Dunning, & J. I. Krueger (Eds.), *The self in social perception* (pp. 155–178). New York: Psychology Press.

- Malle, B. F., Heim, K., Knorek, J. (2005). *Storms revisited: Visual perspective and actor-observer asymmetries in explanation*. Manuscript in preparation.
- Malle, B. F., & Ickes, W. (2000). Fritz Heider: Philosopher and psychologist. In G. A. Kimble & M. Wertheimer (Eds.), *Portraits of pioneers in psychology* (Vol. 4, pp. 193–214). Washington, DC: American Psychological Association.
- Malle, B. F., & Knobe, J. (1997a). The folk concept of intentionality. *Journal of Experimental Social Psychology*, *33*, 101–121.
- Malle, B. F., & Knobe, J. (1997b). Which behaviors do people explain? A basic actor-observer asymmetry. *Journal of Personality and Social Psychology*, *72*, 288–304.
- Malle, B. F., & Knobe, J. (2001). The distinction between desire and intention: A folk-conceptual analysis. In B. F. Malle, L. J. Moses, & D. A. Baldwin (Eds.), *Intentions and intentionality: Foundations of social cognition* (pp. 45–67). Cambridge, MA: MIT Press.
- Malle, B. F., Knobe, J., & Nelson, S. (2005). *Actor-observer asymmetries in folk explanations of behavior: New answers to an old question*. Manuscript in preparation.
- Malle, B. F., Knobe, J., O’Laughlin, M., Pearce, G. E., & Nelson, S. E. (2000). Conceptual structure and social functions of behavior explanations: Beyond person—situation attributions. *Journal of Personality and Social Psychology*, *79*, 309–326.
- McClure, J. (2002). Goal-based explanations of actions and outcomes. In W. Stroebe & M. Hewstone (Eds.), *European review of social psychology* (Vol. 12, pp. 201–235). Wiley.
- McClure, J., & Hilton, D. (1997). For you can’t always get what you want: When preconditions are better explanations than goals. *British Journal of Social Psychology*, *36*, 223–240.
- Mele, A. R. (1992). *Springs of action: Understanding intentional behavior*. New York: Oxford University Press.
- Mele, A. R. (2003). *Motivation and agency*. New York: Oxford University Press.
- Miller, D. T., & Porter, C. A. (1980). Effects of temporal perspective on the attribution process. *Journal of Personality and Social Psychology*, *34*, 532–541.
- Moore, B. S., Sherrod, D. R., Liu, T. J., & Underwood, B. (1979). The dispositional shift in attribution over time. *Journal of Personality and Social Psychology*, *15*, 553–569.
- Moses, L. J., & Flavell, J. H. (1990). Inferring false beliefs from actions and reactions. *Child Development*, *61*, 929–945.
- Nelson, S., & Malle, B. F. (2005). *Self-serving biases in explanations of intentional behavior*: Manuscript in preparation.
- Nigro, G., & Neisser, U. (1983). Point of view in personal memories. *Cognitive Psychology*, *15*, 467–482.
- O’Laughlin, M. J., & Malle, B. F. (2002). How people explain actions performed by groups and individuals. *Journal of Personality and Social Psychology*, *82*, 33–48.
- Pennington, N., & Hastie, R. (1993). Reasoning in explanation-based decision making. *Cognition*, *49*, 123–163.
- Peterson, D. K., & Pitz, G. F. (1986). Effects of amount of information on predictions of uncertain quantities. *Acta Psychologica*, *61*, 299–241.

- Read, S. J. (1987). Constructing causal scenarios: A knowledge structure approach to causal reasoning. *Journal of Personality and Social Psychology*, 52, 288–302.
- Reschler, N. (1997). *Predicting the future: An introduction to the theory of forecasting*. New York: State University of New York Press.
- Rips, L. J. (1994). Deduction and its cognitive basis. In R. J. Sternberg (Ed.), *Thinking and problem solving* (pp. 149–178). San Diego, CA: Academic Press.
- Robinson, E. J., & Mitchell, P. (1995). Masking of children's early understanding of the representational mind: Backwards explanation versus prediction. *Child Development*, 66, 1022–1039.
- Russell, E. W., & D'Hollosy, M. E. (1992). Memory and attention. *Journal of Clinical Psychology*, 48, 530–538.
- Salmon, W. C. (1989). *Four decades of scientific explanation*. Minneapolis, MN: University of Minnesota Press.
- Sanna, L. J., & Swim, J. K. (1992). Temporal perspective and attributions: The role of causal stability and certainty. *Basic and Applied Social Psychology*, 13, 371–387.
- Schwarz, N., & Bless, H. (1992). Constructing reality and its alternatives: An inclusion/exclusion model of assimilation and contrast effects in social judgments. In L. L. Martin & A. Tesser (Eds.), *The construction of social judgments* (pp. 217–245). Hillsdale, NJ: Erlbaum.
- Sloman, S. (1996). The empirical case for two reasoning systems. *Psychological Bulletin*, 119, 3–22.
- Slovic, P. (Ed.). (2000). *The perception of risk*. Sterling, VA: Earthscan.
- Snook, B., Canter, D., & Bennell, C. (2002). Predicting the home location of serial offenders: A preliminary comparison of the accuracy of human judges with a geographic profiling system. *Behavioral Sciences and the Law*, 20, 109–118.
- Storms, M. D. (1973). Videotape and the attribution process: Reversing actors' and observers' points of view. *Journal of Personality and Social Psychology*, 27, 165–175.
- Truchot, D., Maure, G., & Patte, S. (2003). Do attributions change over time when the actor's behavior is hedonically relevant to the perceiver? *Journal of Social Psychology*, 143, 202–208.
- Weiner, B. (1985). An attributional theory of achievement-related emotion and motivation. *Psychological Review*, 29, 548–573.
- Windschitl, P. D., & Wells, G. L. (1998). The alternative-outcomes effect. *Journal of Personality and Social Psychology*, 75, 1411–1423.
- Wright, E. E., & Wells, G. L. (1985). Does group discussion attenuate the dispositional bias? *Journal of Applied Social Psychology*, 15, 531–546.
- Zeelenberg, M. (1999). Anticipated regret, expected feedback and behavioral decision making. *Journal of Behavioral Decision Making*, 12, 93–106.