Norm Theory: Comparing Reality to Its Alternatives

Daniel Kahneman
University of British Columbia
Vancouver, British Columbia, Canada

Dale T. Miller
Simon Fraser University
Burnaby, British Columbia, Canada

A theory of norms and normality is presented and applied to some phenomena of emotional responses, social judgment, and conversations about causes. Norms are assumed to be constructed ad hoc by recruiting specific representations. Category norms are derived by recruiting exemplars. Specific objects or events generate their own norms by retrieval of similar experiences stored in memory or by construction of counterfactual alternatives. The normality of a stimulus is evaluated by comparing it to the norms that it evokes after the fact, rather than to precomputed expectations. Norm theory is applied in analyses of the enhanced emotional response to events that have abnormal causes, of the generation of predictions and inferences from observations of behavior, and of the role of norms in causal questions and answers.

This article is concerned with category norms that represent knowledge of concepts and with stimulus norms that govern comparative judgments and designate experiences as surprising. In the tradition of adaptation level theory (Appley, 1971; Helson, 1964), the concept of *norm* is applied to events that range in complexity from single visual displays to social interactions. We first propose a model of an activation process that produces norms, then explore the role of norms in social cognition.

The central idea of the present treatment is that norms are computed after the event rather than in advance. We sketch a supplement to the generally accepted idea that events in the stream of experience are interpreted and evaluated by consulting precomputed schemas and frames of reference. The view developed here is that each stimulus selectively recruits its own alternatives (Garner, 1962, 1970) and is interpreted in a rich context of remembered and constructed representations of what it could have been, might have been, or should have been. Thus, each event brings its own frame of reference into being. We also explore the idea that knowledge of categories (e.g., "encounters with Jim") can be derived on-line by selectively evoking stored representations of discrete episodes and exemplars.

The present model assumes that a number of representations can be recruited in parallel, by either a stimulus event or an abstract probe such as a category name, and that a norm is produced by aggregating the set of recruited representations. The assumptions of distributed activation and rapid aggregation are not unique to this treatment. Related ideas have been advanced in theories of adaptation level (Helson, 1964; Restle, 1978a, 1978b) and other theories of context effects in judgment (N. H. Anderson, 1981; Birnbaum, 1982; Parducci, 1965, 1974); in connectionist models of distributed processing (Hinton & Anderson, 1981; McClelland, 1985; McClelland & Rumelhart, 1985); and in holographic models of memory (Eich, 1982; Metcalfe Eich, 1985; Murdock, 1982). The present analysis relates most closely to exemplar models of concepts (Brooks, 1978, in press; Hintzman, in press; Hintzman & Ludlam, 1980; Jacoby & Brooks, 1984; Medin & Schaffer, 1978; Smith & Medin, 1981). We were drawn to exemplar models in large part because they provide the only satisfactory account of the norms evoked by questions about arbitrary categories, such as "Is this person friendlier than most other people on your block?"

Exemplar models assume that several representations are evoked at once and that activation varies in degree. They do not require the representations of exemplars to be accessible to conscious and explicit retrieval, and they allow representations to be fragmentary. The present model of norms adopts all of these assumptions. In addition, we propose that events are sometimes compared to counterfactual alternatives that are constructed ad hoc rather than retrieved from past experience. These ideas extend previous work on the availability and simulation heuristics (Kahneman & Tversky, 1982; Tversky & Kahneman, 1973).

A constructive process must be invoked to explain some cases of surprise. Thus, an observer who knows Marty's affection for his aunt and his propensity for emotional displays may be surprised *if Marty does not cry at her funeral*—even if Marty rarely cries and if no one else cries at that funeral. Surprise is produced in such cases by the contrast between a stimulus and a counterfactual alternative that is constructed, not retrieved. Constructed elements also play a crucial role in counterfactual emotions such as frustration or regret, in which reality is compared to an imagined view of what might have been (Kahneman & Tversky, 1982).

At the core of the present analysis are the rules and constraints that govern the spontaneous retrieval or construction of alter-

natives to experience. A question of particular interest concerns the mutability of different attributes of an evoking stimulus: Which of its features will be retained in the norm elements that it evokes? Which attributes will have highly variable norms? The differential mutability of attributes restricts both the retrieval and the construction of norm elements.

An abnormal event is one that has highly available alternatives, whether retrieved or constructed; a normal event mainly evokes representations that resemble it. The treatment of normality in this article is guided by the phenomenology of surprise rather than by formal or informal conceptions of probability. The main difference between the two notions is that probability is always construed as an aspect of anticipation, whereas surprise is the outcome of what we shall call backward processing—evaluation after the fact. Probability reflects expectations. Surprise (or its absence) reflects the failure or success of an attempt to make sense of an experience, rather than an evaluation of the validity of prior beliefs. In his critique of standard notions of probability, the statistician Shafer (1976) developed the related idea that events do not merely alter the strength of belief in existing possibilities; they also evoke and shape the set of relevant possibilities.

Of course, specific anticipations that exist in advance of an event will be included in the norm to which it is compared. The event will then appear normal if it confirms expectations, abnormal or surprising if it violates them. However, an unanticipated event will also be judged normal if it simply fails to evoke strong alternatives. This formulation distinguishes two ways in which an occurrence may affect the normality of subsequent events: (a) by eliciting hypotheses and expectations, which later events confirm or disconfirm, or (b) by laying down a trace that is activated when a subsequent event provides an appropriate reminder (Schank, 1982).

Consider an observer, casually watching the patrons at a neighboring table of a fashionable restaurant, who notices that the first guest to taste the soup winces, as if in pain. The normality of a multitude of events will be altered by this incident. For example, it is now unsurprising for the guest who first tasted the soup to startle violently when touched by a waiter; it is also unsurprising for another guest to stifle a cry when tasting soup from the same tureen. These events and many others appear more normal than they would have done otherwise, but not necessarily because they confirm advance expectations. Rather, they appear normal because they recruit the original episode and are interpreted in conjunction with it. In general, selective retrieval of pertinent episodes tends to reduce surprise and to favor hindsight about both the recent event and its predecessor (Fischhoff, 1975, 1982).

Reasoning flows not only forward, from anticipation and hypothesis to confirmation or revision, but also backward, from the experience to what it reminds us of or makes us think about. This article is largely dedicated to the power of backward thinking. Its aim is not to deny the existence of anticipation and expectation but to encourage the consideration of alternative accounts for some of the observations that are routinely explained in terms of forward processing.

We first introduce a model of norms and illustrate some of its applications. The rules that govern the generation of norm elements are then introduced, as well as some consequences of these rules in the domain of emotion. The remainder of the article explores the function of norms as representations of stored knowledge of persons and the role of norms in causal reasoning.

A Model of Norms

A model of norms is sketched in Figure 1. A probe (which may be the experience of an object or event, or a reference to a concept) recruits an *evoked set* that consists of *elements* (A, B, and C in the example). The elements of the evoked set are activated (made available) to different degrees, as indicated in the figure by the thickness of the arrows. The elements are representations of objects, episodes, or classes of elements. Representations of the neighbor's dog Fido or of the category "poodle" could be recruited as elements of the set evoked by the probe "dog." Each element is internally described by *features*, which are specific values of *attributes* (X and Y in the example). The evoked set is characterized by norms for each of the attributes that describe its elements. Norms for the attributes X and Y are shown in the bottom panels.

Elements are internally described in terms of physical attributes (e.g., size), more abstract ones (e.g., friendliness), and some conjunctions of elementary attributes (e.g., size and strength). For simplicity, the attributes X and Y are presented in Figure 1 as ordered dimensions, but the treatment extends readily to attributes that have other similarity structures. As indicated in the figure, each feature of an element is described by a profile or distribution of activation over a range of attribute values. When the element represents an individual object or event, the shape of the profile can be interpreted as a gradient of generalization. The profile is also flattened by any uncertainty or imprecision in the assignment of a feature to an element. When the element stands for a class, the profile represents the internal variability of the class. The degree of activation of an element determines the size of the profiles for its attributes.

The entire evoked set is described by summing, for each attribute, the profiles associated with all activated elements. We shall say that the aggregate profile assigns a measure of *availability* to values of the attribute. A measure of *normality* is obtained by rescaling the availability profile, assigning a normality of 1.0 to the most available value. The normality measure therefore ranges between 0 and 1, and the normality of any attribute value is the ratio of the availability of that value to the modal (maximal) availability. A norm is a function that assigns a normality measure to values of an attribute.

In summary, a probe recruits an evoked set of individual elements, each of which is described by several features. The evoked set is described by an aggregate of individual features. The norm for an attribute is the envelope of this aggregate profile, scaled to assign unit normality to the modal value. In the text that follows, "norm element" is often used as shorthand for "element of the evoked set."

As is evident in Figure 1, the principal independent variables of the model are the factors that control the activation of potential elements of an evoked set. Direct and indirect manifestations of normality are the dependent variables. The direct measures include expressions of surprise and judgments of normality or typicality. Indirect measures include intuitive predictions, emotional responses to abnormal events, and various aspects of causal reasoning. The schema for applying the theory is the following:
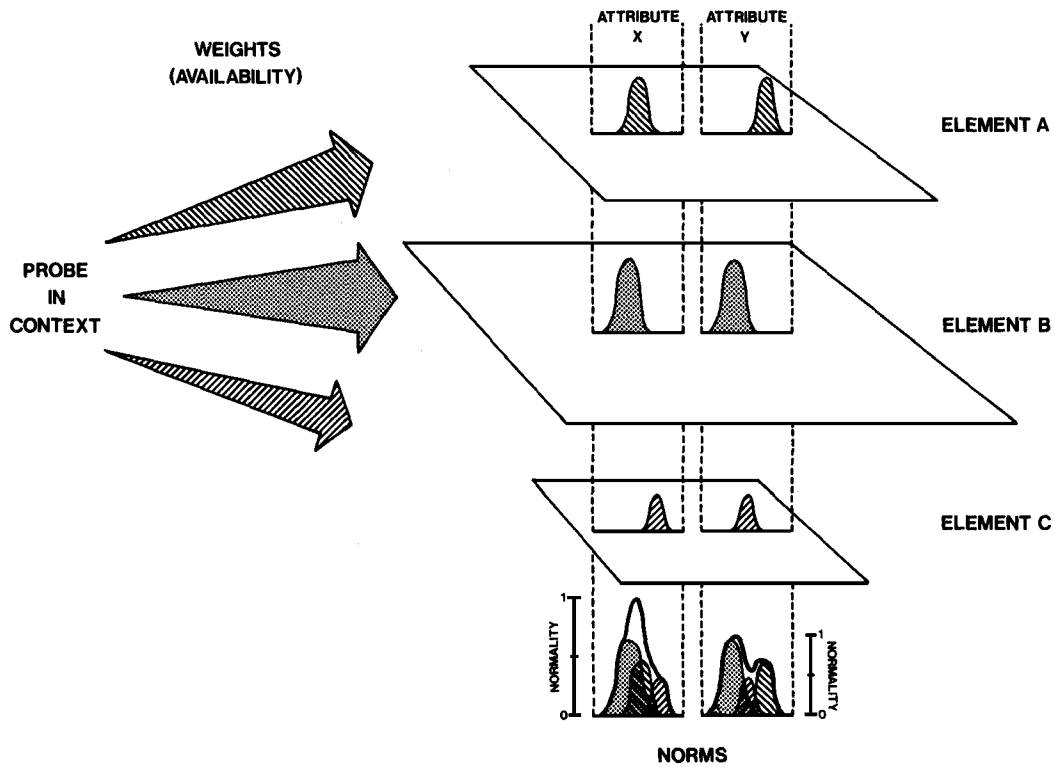
Figure 1. Model of norms. (A stimulus or category label, and its context, activates elements A, B, and C to different degrees. The values of each element on two attributes, X and Y, are represented by profiles, which are summed to establish a norm for each attribute. The normality of a value is defined as the ratio of its aggregate availability [the height of the norm at that value] to the modal [maximum] availability in the norm.)

Keeping all else constant, a manipulation that increases the probability of a given element being recruited by a probe will also increase the weight of that element in the norms evoked by that probe.

In the next sections we elaborate on three central aspects of the present analysis: the recruitment of norm elements, the role of norms in the representation of knowledge, and the concept of normality.

Modes of Recruitment

The model of Figure 1 applies to two types of norms that are distinguished by their evoking probe: (a) *stimulus norms,* which are evoked by experiences of objects and events, and (b) *category norms,* which are evoked by references to categories. Two modes of recruitment are distinguished: (a) retrieval of memory representations of individual objects and events or of subordinate categories and (b) construction of counterfactual alternatives to experience. As these concepts are used here, the scope of "recruitment" is broader than "retrieval," and "element" is broader than "memory trace."[1]

A process that retrieves or generates specific exemplars appears necessary to account for people's ability to deal with arbitrary collections or functionally defined ad hoc categories, such as "reasons for firing an employee" or "things to take on a camping trip" (Barsalou, 1983). Instances of ad hoc categories can be evaluated on such attributes as reasonableness for a dismissal or

bulk for a camping implement. Furthermore, these comparative judgments appear neither especially difficult nor especially slow. A dismissal can be judged outrageous or a sleeping bag bulky without conscious examination of many (or any?) instances of the category. Multiple representations appear to be activated in parallel without entering consciousness or working memory. Summary statistics for the activated elements are used for comparative judgments, and perhaps for other purposes as well.

The recruitment of norm elements has many of the characteristics attributed to spreading associative activation (J. R. Anderson, 1983; Ratcliff & McKoon, 1981). However, not all of the representations that are associatively connected to the probe will be included in its norm. Inclusion in a category norm, in particular, must be restricted to members of the designated category. The norm for horses should not include carriages, and the category label "married graduate students" should not strongly activate representations of married nonstudents or unmarried students (Osherson & Smith, 1982). The required selectivity of recruitment is achieved more economically by precisely controlled activation than by inhibition of irrelevant activated elements. A

---

[1] An analysis of perceptual norms requires a third process of recruitment, in which a focal stimulus selectively interacts with representations of some of the objects present in the current perceptual field. Context effects in perception are often analyzed in such terms (N. H. Anderson, 1981; Coren & Girgus, 1978; Coren & Miller, 1974; Restle, 1978a, 1978b).

plausible hypothesis is that the activating effects of distinct constituents of the probe, or of the probe and its context, are mutually reinforcing (Medin & Schaffer, 1978). The specification of a category by a conjunction of features serves to restrict the spread of activation to elements that possess most or all of them. There may also be a limit on the number of elements that can be simultaneously activated by a single probe—perhaps half a dozen or less (J. R. Anderson, 1983; Mandler, 1967, 1975).

The same probe can elicit both retrieval and construction. For example, consider a person who is involved in an accident. The occurrence is a memory probe, which acts as a reminder of similar experiences in the past. The current occasion will appear more normal if traces of similar experiences are activated than otherwise (Schank, 1982). Any serious accident will also provoke an examination of the sequence of events that led to it, and this examination in turn involves the generation of counterfactual alternatives. The occurrence will appear especially abnormal if some scenarios that yield a different outcome are highly available. The outcome will appear inevitable if no such alternatives come readily to mind.

The generation of alternatives to reality appears to be quite disciplined. Inclusion in a stimulus norm is restricted to objects and events that share the immutable features of the evoking stimulus. For example, the alternative scenarios that are produced in mentally "undoing" an accident tend to alter some features of the real sequence of events, leaving others constant (Kahneman & Tversky, 1982). We return later to a discussion of mutable and immutable features, which also confronts an obvious difficulty for the present model: If stimuli only evoke norm elements that are similar to them, how can a stimulus ever appear abnormal? In responding to this question we shall propose that recruited elements are only constrained to share the immutable features of the evoking stimulus. Norm elements are allowed to differ from the evoking stimulus and from each other in other attributes, and any mutable feature of the evoking stimulus can therefore appear abnormal.

Norms as Representations of Knowledge

The model of Figure 1 invokes the same exemplar representation for category norms and for stimulus norms and the same diffuse description for individuals and for sets. We next sketch the reasoning that led to the apparent neglect of these distinctions.

The decision to describe both category norms and stimulus norms as temporary patterns of memory activation is a response to a failure. Although our concern was mainly with stimulus norms, which are temporary by definition, we were unable to deal with these norms without invoking category knowledge, and equally unable to draw a clean boundary between temporary and durable representations of such knowledge. At issue is the representation of knowledge that is not at the moment in use. One common view is that knowledge of categories (e.g., "dogs" or "encounters with Jim") is contained in a single compact representation in semantic memory, like a book in a library, which is consulted when information is needed about dogs or about Jim. This notion of a schema is most useful in dealing with information that is needed often and that can be applied time after time with little variation (Hastie, 1981; Mandler, 1984; Rumelhart, 1980). However, the need to represent context dependence in the use of knowledge has prompted several theorists to search for a mode of representation that is more molecular and more flexible than a schema (Alba & Hasher, 1983; Barsalou, 1985; Johnson, 1983; Lakoff, in press; Schank, 1982). The possibility considered here is an assemblage of exemplars (Brooks, 1978, in press; Medin & Schaffer, 1978). Memory is assumed to store content-addressable files, each containing data about a specific episode of experience. When information is required some of the files are selectively retrieved, computations are performed as needed—and perhaps a book about the topic is printed and bound for conscious inspection, all within 200 milliseconds or so. This mode of representation by exemplars appears necessary for stimulus norms. It is also very plausible for ad hoc categories such as "encounters with Jim in the elevator," which in turn shade into the domain of familiar categories and natural kinds.

It is important to stress that no exemplar model can account for all variants of category knowledge (see also Murphy & Medin, 1985). The inheritance of properties, in particular, cannot be explained by such a model. Thus, it is surely normal for birds to have stomachs, but this norm is derived deductively from knowledge about animals, not induced from observed exemplars. The only reasonable claim for exemplar models is that some knowledge of categories can be represented by norms that are computed on the fly, not that all category knowledge is achieved in this manner. Furthermore, it appears likely that norms that are evoked repeatedly are eventually stored as summary statistics rather than as raw data (Fried & Holyoak, 1984). In spite of these qualifications, the assumption that category knowledge is often represented by exemplars appears necessary. It also appears especially fruitful in the representation of knowledge about other people.

The properties of on-line computation and context sensitivity are common to exemplar models and to connectionist and holographic representations (e.g., Eich, 1982; McClelland & Rumelhart, 1985; Murdock, 1982). However, the current versions of these models represent a category by a composite pattern of memory activation, which obliterates information about both individual elements and higher order statistics. The easy access of observers to variances and covariances of attributes in collections of instances is perhaps the strongest reason to reserve a role for individual exemplars in a model of category knowledge (Medin, 1983; Medin, Altom, Edelson, & Freko, 1982; Medin & Schaffer, 1978).

The model of Figure 1 assigns a diffuse description to the features of individual objects and events. As in the classic composite photograph, the representation of the evoked set is an aggregate, but in this case each of the pictures in the composite is fuzzy on its own. The view of norms as aggregates of diffuse descriptions accommodates the basic finding that an unpresented prototypical member of a category is likely to be erroneously recognized as familiar (Posner & Keele, 1968). Because activation spreads to the vicinity of each presented value, a feature that has not been presented can be highly normal (typical) if neighboring values of the attribute have frequently appeared.

An important implication of diffuse description is that the model does not distinguish individuals from sets. As a consequence, the elements of a norm can be either tokens (episodes of experience) or types (concepts of lower order categories). For example, it is standard to view the category "bird" as consisting

of kinds of birds, such as robin and chicken (Smith & Medin, 1981). In applying an exemplar model to such classes, the subordinate categories contribute their norms to the norm for the inclusive class. Here again, some elements will contribute more than others: a reference to "bird" is likely to activate the representations of "robin" and "sparrow" more strongly than those of "ostrich" or "chicken" (Rosch & Mervis, 1975; Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976). The present treatment offers the mode of the category norm as a counterpart to the notion of prototypical value. It also incorporates a notion of differential activation of elements, akin to the idea of graded membership in a category (Barsalou, in press; Rosch, 1978; Rosch & Mervis, 1975; Zadeh, 1965, 1976).

Normality and Probability

A norm and a probability distribution can both be displayed as a function of attribute values, but normality and probability differ in their formal characteristics. Normality values are represented by heights (scaled to a maximum of 1.0), whereas probabilities are represented by areas (scaled so that the total area under the profile is 1.0). One immediate consequence is that, unlike probability, the normality of a value of an attribute can increase without a corresponding reduction of the normality of any other. It can be seen in Figure 1, for example, that the addition of element C to elements A and B made high values of X more normal without reducing the normality of low values of that attribute.

Another contrast concerns disjunctions and conjunctions of values. The probability of a disjunction of mutually exclusive events is the sum of their probabilities, but the most reasonable rule for the normality of a disjunction is that it equals the maximal normality of its constituents (Zadeh, 1965). In particular, extending the range of a category does not necessarily increase normality. It is no more normal for a man to be 5 ft 10 in. tall (to the nearest inch) than to be 175 cm tall (to the nearest cm), although the former event is substantially more probable.

A message is compared to alternatives at the same level of detail. Thus, the information that the favorite lost the first set of a tennis match will only evoke the alternative possibility of the favorite winning that set. A more detailed message, stating that the favorite lost the first set but won the match will evoke alternative conjunctions of outcomes for the first set and the match. Because normality is relative, it is quite possible for a single event (e.g., that the favorite lost the first set) to be more surprising than the conjunction of that event with another (e.g., that the favorite lost the set and won the match). The rule that the probability of a conjunction cannot exceed that of any of its constituents extends neither to normality nor to surprise. Confusion of probability with normality or surprise may be a contributing factor in some—not all—violations of the conjunction rule in naive judgments of probability (Tversky & Kahneman, 1983).

The concept of normality is closely related to both typicality (Rosch, 1973) and representativeness (Tversky & Kahneman, 1983). A significant theoretical difference is that representativeness is modeled as a similarity relation, whereas normality has been traced to the availability of exemplars. However, availability and representativeness are intertwined in the generation of norms.
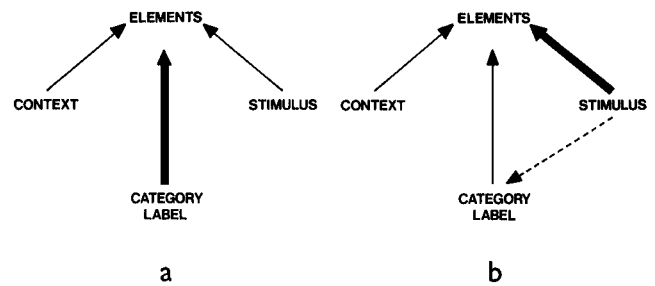

Figure 2. Two models of recruitment of norm elements.

The name of a category will surely evoke (make available) its representative exemplars more strongly than others (Barsalou, in press; Rosch & Mervis, 1975; Rosch et al., 1976). The characteristics of the most available instances, in turn, are likely to be judged most representative of the category. The present analysis focuses on availability as the immediate determinant of norms.

Comparative Judgments

We have distinguished two types of probes that evoke norms: experiences of objects and events and references to categories. An application of this distinction to comparative judgment is illustrated in Figure 2. In *category-centered* comparisons, the object of judgment is compared to the norm for a specified category. In *stimulus-centered* comparisons, the elements of the norm tend to be recruited directly by the stimulus itself.

"Jane owns a small dog" is an example of a category-centered judgment. To make and to interpret such judgments, a norm of size for a particular category must be invoked. However, the set of recruited exemplars can be biased by context or background. The same statement will yield different norms of size and different ideas of the size of Jane's dog if she is known to live in a New York apartment or on a farm in Maine (Barsalou, in press; Roth & Shoben, 1983).

Stimulus-centered judgments are more elusive. Consider the following information: "Ms. Z is 26 years old, with a Master of Science degree in geography. She earns $33,000 a year." Most readers will probably recognize that, although not instructed to do so, they have already evaluated Ms. Z's salary as high or low. The norm for this judgment is not precomputed: Few people will have access to stored statistics for the income of 26-year-old women with a master's degree in geography. Furthermore, we suspect that the norm that yields the spontaneous judgment of Ms. Z's salary is not quite the same as would be elicited by the category-centered instruction "Compare Ms. Z's income to 26-year-old. . . ." In particular, some friends have joined us in confessing that thoughts of their own past and present income were not irrelevant to their evaluation of Ms. Z's salary. Stimulus-centered norms are not restricted to members of a particular category and are likely to be biased toward highly available examples.

The distinction between stimulus-centered and category-centered judgments identifies a continuum rather than a dichotomy. In the pure case of category-centered judgment, only instances of the category will be evoked, and the stimulus and context further narrow the selection of category members. At the other

extreme, a stimulus-centered norm can be evoked without any explicit categorization of the evoking stimulus. As indicated by the broken arrow in Figure 2, classification is optional in stimulus-centered judgments. If an explicit category label is activated, elements attached to it will be evoked, but the norm will not be restricted to these elements.

Probes and task demands are an important aspect of the context that affects the recruitment of elements in both stimulus-centered and category-centered judgments. In particular, the designation of an attribute as focal tends to increase the mutability of that attribute, that is, its variability among members of the evoked set. For example, questions about Ms. Z's income or educational level will increase the respective variability of each of these attributes in the norms to which Ms. Z is compared.

Control of Recruitment

Explicit reference to a category permits a high degree of control over the selection of norm elements. In particular, the use of a category label is very effective in restricting the evoked set to category members (Rips & Turnbull, 1980). Consider the sentence "The large fly climbed up the trunk of the small elephant." The sentence is understood without difficulty, although its interpretation almost simultaneously invokes two different norms of size. The work of Barsalou (1983, in press) shows that norms may be generated even for ad hoc categories, and also that people can take the point of view of others in considering exemplars of a category. The ability to simulate another person's norms is often essential to successful communication. The task is presumably carried out by selecting a subset of one's own relevant experiences. Parents seem to have no difficulty in determining that a particular animal is the largest that their 2-year-old has ever seen.

Voluntary control of invoked categories is of course not perfect. The idea that recruitment is controlled by selective activation rather than by inhibition suggests that it might be difficult to exclude designated instances from a category norm—just as it is difficult to obey the instruction not to think of elephants. This reasoning entails an interesting asymmetry: An observer might be able to include selected elements in a norm by deliberately thinking about them but fail to exclude specified elements from a norm if they have been associatively activated. When Parducci (1956) informed subjects of a change in the composition of a series that they were judging, he observed rapid adjustment to the news that additional stimuli would extend the range previously included, but he observed almost no adjustment in response to an announced restriction of the range. This finding suggests that size judgments assigned to common mammals will change abruptly when respondents are informed that whales and elephants may occur in the list, but the judgments will not adjust with equal ease to the information that the larger mammals will not, after all, be included. In a weight judgment task, Brown and Reich (1971) found that the instruction to make the judgment of each weight relative only to weights of the same color was largely ineffective if it was given after a common frame of reference was established.

Effects of Availability and Mutability

The factors that govern the weighting of norm elements in comparative judgments have been studied most extensively in the context of adaptation level theory (Appley, 1971; Helson, 1964). The instructions given to subjects in adaptation level studies do not usually specify a reference category. The judgments are therefore stimulus centered rather than category centered. As a consequence, stimuli encountered outside the immediate task context often have substantial effects on observed adaptation levels. A 5-g weight is rarely judged heavy, regardless of the experimental context. Stimuli presented for judgment also vary in their influence on the adaptation level. Two sets of factors determine the weight of previous stimuli in a norm: (a) factors that affect the salience or memorability of specific experiences and (b) the strength of the connection between norm elements and the evoking stimulus, which is mainly determined by the similarity between them.

Avant and Helson (1973, p. 440) offered the following summary list of the factors that control the impact of stimuli on adaptation level: "recency, frequency, intensity, area, duration, and higher-order attributes such as meaningfulness, familiarity and ego-involvement." This list could serve equally well as a list of determinants of availability for retrieval. There is evidence that the weight of stimuli in the norm is higher for recent stimuli than for those presented earlier (Lockhead & King, 1983; Ward, 1979). The weight of a stimulus is also increased by making it distinctive or salient on some irrelevant attribute such as duration of exposure (Helson & Kozaki, 1968). On the other hand, an anchor stimulus that is redundantly presented before each trial has much less weight in the norm than the aggregate of other stimuli, although the total frequency of presentation is the same for the anchor and for all other stimuli combined (Helson, 1964).

The notion of postcomputed norms implies that recruitment favors elements that resemble the evoking stimulus. As a consequence, different objects of judgment may be compared to somewhat different norms even in the context of a single task. In support of this idea, Restle (1978a) highlighted the finding by Sarris (1976) that a repeated anchor does not have the same effect on the judgments of all stimuli. The anchor is assigned the largest weight in judgments of its nearest neighbors in the test series.

The first trial in a judgment experiment provides the best example of the process in which a stimulus evokes its own context. The present model suggests that some features of the evoking stimulus are treated as immutable in that process: The recruitment of the evoked set tends to be restricted to elements that share these features. A plausible hypothesis is that the essential features that define the identity of the stimulus are most likely to be maintained as immutable. This hypothesis has surprising consequences: It entails that judgments of a stimulus evaluated in isolation will tend to be dominated by features that are not the most central.

A striking result observed by Slovic (1985) provides an illustration. Slovic asked a group of subjects to evaluate on a 20-point scale the attractiveness of the following bet: "a 7/36 chance to win $9." The mean evaluation was 9.4. A different group of subjects evaluated the bet "a 7/36 chance to win $9 and a 29/36 chance to lose 5 cents." Although the second bet is strictly inferior to the first, its average rating on the attractiveness scale was 14.4. The two bets appear to be compared to different norms. The second bet appears very favorable among bets that involve a risk of loss, but a modest chance to win $9 is mediocre in a

context of purely positive prospects. The attributes of certain objects, such as risky prospects, are hierarchically related. Dominant attributes—in this case the presence or absence of a risk of loss—are relatively immutable and therefore control the recruitment of the norm. The attributes that are allowed to vary in the norm account for most of the variance in judgments of such objects.

A similar interpretation applies to another curious phenomenon in the evaluation of positive bets: The probability of winning and the amount of the prize have different weights in choices and in judgments of attractiveness (Goldstein, 1982). Consider, for example, the two bets "a 31/36 chance to win \$3" and "a 7/36 chance to win \$13." In an experiment conducted at the University of British Columbia, the two bets were preferred about equally often in a direct choice. When they were rated for attractiveness, however, the bet with the higher probability of winning was rated 12.1 on a 20-point scale, and the bet with the lower probability of winning was rated only 6.5. As this example illustrates, the probability attribute accounts for most of the variance in ratings of attractiveness. This is the pattern of judgments that would be expected if a bet that offers a probability P to win \$$X$ is compared mainly to other bets in which the same amount X can be won with different probabilities. The dominant role of probability in these judgments fits easily within the present framework, with the added assumption that the amount to be won is the central attribute of a positive bet—and therefore the attribute most likely to be adopted as immutable in the recruitment of a norm. Indeed, most people state that they would rather know the prize than the probability of winning in assessing the attractiveness of a bet.

Evaluations of the attractiveness of bets illustrate what appears to be a more general effect. Consider two factors that may determine the impression of the career success of a civil servant: rank and performance ratings. The hierarchical ordering of the two attributes is clear: Rank and the tasks associated with it are commonly presupposed in evaluating performance, and the very meaning of successful performance is altered by varying rank. Now imagine two civil servants of different ranks who are judged equally successful in direct comparison, because the one at the lower rank has a higher performance rating. The present analysis entails that when the success of the two individuals is evaluated in isolation, the judgments will be mainly determined by the performance ratings. The general rule is that, other things being equal, the more mutable and less important of two attributes will have a disproportionately large effect in single-stimulus judgments.

The application of this analysis of mutability is straightforward when a single object is presented for judgment, with no prior experimental context. The situation is ambiguous in a within-subject design, where several objects are evaluated in immediate succession. The early items in the series are likely to play a significant role in the evaluation of later ones, and some of the paradoxical effects of the hierarchy of attributes are weakened or even reversed in such cases. When positive and mixed bets are evaluated in the same series, for example, the positive ones tend to be rated higher. Birnbaum (1982) described another paradoxical result that is eliminated in a within-subject design. When a single case is evaluated, judgments of a rape victim's responsibility are higher for a virgin than for a divorcee, perhaps because the raped virgin is more discrepant in the norm that she evokes than the raped divorcee is in hers. The effect is strongly reversed when the same subjects judge both cases. Birnbaum (1982) commented that in a between-subjects design, the stimulus is completely confounded with its context—which is another way of saying that the stimulus brings its context into being.

The role of the immutable features of a stimulus in recruiting its norms is the same as the role of the category label in recruiting category norms. A category label—whether it is a single name or a complex specification—constrains the retrieval process to category members. Other factors, such as typicality or recency, control further selection among representations of exemplars and ultimately determine how strongly individual instances are activated. We propose that the immutable features of a stimulus similarly guide and constrain the spontaneous recruitment of alternatives to it.

The processes that have been sketched in this discussion of comparative judgment yield norms for the various attributes of a stimulus. The availability profile (see Figure 1) defines a range of possible values and provides a proxy for a frequency distribution. In particular, the rank of the stimulus in its norm is readily computable from this information. Thus, a norm has the characteristics necessary to support the scaling operations envisaged in range–frequency theory (Parducci, 1965, 1974) and in related analyses of comparative judgment (Birnbaum, 1982; Mellers & Birnbaum, 1983). More complex judgments involve matching values of different attributes by their position in their respective norms. For example, nonregressive predictions are made by choosing a criterion value that matches the position of the predictor feature in its norm (Kahneman & Tversky, 1973). Similarly, judgments of equity appear to be based on a comparison of the relative positions of an individual in norms of salary and merit (Mellers, 1982).

Mutability and the Availability of Counterfactuals

One theme of this article is that the experienced facts of reality evoke counterfactual alternatives and are compared to these alternatives. The development of this theme takes us to regions more often traveled by philosophers than by psychologists. Philosophical treatments of counterfactuals and possible worlds have explored the compelling intuition that some alternatives are closer to reality than others and that some changes of reality are smaller than others (see Lewis, 1973, for a particularly engaging treatment). As Hofstadter (1985) noted, the word "almost" provides a key to some of these intuitions. For example, the statement "I almost caught the flight" is appropriate for an individual who reached the departure gate when the plane had just left but not for a traveler who arrived half an hour late. The world in which the passenger arrives five minutes earlier than she did is closer to reality than a world in which she arrives half an hour earlier. The present analysis links these intuitions to mutability: A counterfactual possibility should appear "close" if it can be reached by altering some mutable features of reality.

Our notion of mutability is similar to the concept of slippability introduced by Hofstadter (1979, 1985). The shared ideas are that the mental representation of a state of affairs can always be modified in many ways, that some modifications are much more

natural than others, and that some attributes are particularly resistant to change.

Another cognate of mutability is the distinction between the information presupposed and the information asserted in a verbal message (Clark & Haviland, 1977). The cleft sentence "It was Tom who set fire to the hotel" designates an immutable aspect of the situation (the hotel was set on fire) and a mutable one (the identity of the arsonist). Note that either aspect of the sentence could be presupposed: "It was a hotel that Tom set on fire" has the same basic structure, with the two components interchanging their roles. The cleft sentence invites the listener to consider alternatives to the asserted content, even as it denies these alternatives. The presupposition is shared by all the alternatives.

As this example illustrates, presuppositions are highly flexible, and the relative mutability of attributes can be controlled almost at will. In the absence of deliberate intent or conversational guidance, however, differences in the mutability of attributes will affect the spontaneous recruitment of norm elements.

In the following sections we examine several hypotheses about factors that determine the relative mutability of different aspects of an event. We also illustrate some of the ways in which the elusive concept of "availability of counterfactual alternatives" can be operationalized.

Exception and Routine

A complex situation may combine some routine and some exceptional features. Kahneman and Tversky (1982) tested the hypothesis that exceptional features are more mutable than routine ones by eliciting alternatives to a stipulated reality. Subjects were given a story describing a fatal road accident, in which a truck driven by a drug-crazed teenager ran a red light and crashed into a passing car, killing Mr. Jones, its occupant. The following instructions were given:

> As commonly happens in such situations, the Jones family and their friends often thought and often said "If only . . ." during the days that followed the accident. How did they continue that thought? Please write one or more likely completions.

Two versions of the story were constructed, labeled *route* and *time,* which were identical except for one paragraph. In the route version the critical paragraph read as follows:

> On the day of the accident, Mr. Jones left his office at the regular time. He sometimes left early to take care of home chores at his wife's request, but this was not necessary on that day. Mr. Jones did not drive home by his regular route. The day was exceptionally clear and Mr. Jones told his friends at the office that he would drive along the shore to enjoy the view.

The time version of this paragraph was as follows:

> On the day of the accident, Mr. Jones left the office earlier than usual, to attend to some household chores at his wife's request. He drove home along his regular route. Mr. Jones occasionally chose to drive along the shore, to enjoy the view on exceptionally clear days, but that day was just average.

Both versions suggest route and time as possible attributes that might be changed to undo the accident, but the change introduces an exception in one case, and restores the routine in the other. As predicted, over 80% of the responses that mentioned either

time or route altered the exceptional value and made it normal. The results support two related propositions about the availability of counterfactual alternatives: (a) Exceptions tend to evoke contrasting normal alternatives, but not vice versa, and (b) an event is more likely to be undone by altering exceptional than routine aspects of the causal chain that led to it.

Ideals and Violations

Barsalou (1985) and Lakoff (in press) have emphasized the role of distance from an ideal or paragon as a determinant of typicality. For example, zero-calorie foods are judged to be highly typical members of the category "things to eat on a diet," although they are neither the most common nor the most similar to the prototypical diet food (Barsalou, 1985). In the terms of the present model, elements that have ideal values on significant attributes appear to be highly available. A hypothesis about differential mutability follows: When an alternative to an event could be produced either by introducing an improvement in some antecedent or by introducing a deterioration, the former will be more available.

Evidence for this proposition was obtained in unpublished research by D. Read (1985). Subjects were taught the rules of a simple two-person card game. They were then shown pictures of the players' hands and were asked to complete the blanks in the following statement by changing one card: "The outcome would have been different if the _____ had been a _____." The question of interest was whether the subjects would choose to weaken the winning hand or to strengthen the losing one. The rule discussed in the preceding section suggests that the winning hand might be more readily altered, since the strongest combinations (e.g., four of a kind) are more exceptional than weaker ones (e.g., three of a kind). However, the tendency to eliminate exceptions was overcome in the data by a tendency to approach an ideal value. In a significant majority of cases, subjects chose to modify the outcome by strengthening the losing hand rather than by weakening the stronger one. Informal observations of spectators at sports events suggest that the outcome of a contest is more commonly undone by improving the losing performance (e.g., imagining the successful completion of a long pass in the last seconds) than by imagining a poorer performance of the winning team.

Differential availability of changes that improve or degrade a performance could be one of the factors that explain the answers to the following question:

> Tom and Jim both were eliminated from a tennis tournament, both on a tie-breaker. Tom lost when his opponent served an ace. Jim lost on his own unforced error. Who will spend more time thinking about the match that night?

> Jim   85%   Tom   15% (N = 92)

Note that Tom and Jim could both imagine themselves winning the game, but the judgment of our subjects is that these thoughts are likely to be more available when they involve an imagined improvement of one's own performance than an imagined deterioration of the opponent's game.

Reliable and Unreliable Knowledge

Tversky and Kahneman (1982) noted an asymmetry in the confidence with which people made inferences and predictions

from one attribute to another. Inferences from a reliable measure to an unreliable one are made with greater confidence than inferences in the opposite direction, although correlation is actually symmetric. For example, people were more confident in predicting the score on a short IQ test from a long form of the test than vice versa. They also believed, erroneously, that the statement "The individual who won the decathlon won the first event" is more probable than "The individual who won the first event won the decathlon."

Similar asymmetries are observed when subjects are presented with a statement that includes an apparent discrepancy and are asked to choose how to eliminate the discrepancy. Most subjects do so by altering the less reliable item (e.g., the performance on the short test or in the first event of the decathlon) to fit the more reliable one. Thus, attributes about which little is known appear to be relatively mutable.

Causes and Effects

We propose the hypothesis that, when people consider a cause-effect pair, alternatives to the effect will be more available than alternatives to the cause. The tendency to presuppose causes is reflected in everyday conversation. When an observation departs from the normal covariation of cause and effect, the discrepancy is usually attributed to the effect rather than to the cause. Thus, a child may be described as "big for her age" but not as "young for her size," and students may be described as overachievers, not as undertalented.

When a particular conjunction of effect and causal attribute is observed, the alternatives that are recruited should mainly consist of cases in which the same cause is followed by variable effects. A spectator at a weight lifting event, for example, will find it easier to imagine the same athlete lifting a different weight than to keep the achievement constant and vary the athlete's physique. We turned this hunch into a small experiment. The participants were given information that was described as a form sheet for the members of a club of weight lifters. The data were presented in two columns, stating the body weight of each athlete and his best achievement (the order of columns was varied in alternate forms). The numbers in both columns were in strict ascending order. Data for 10 athletes were given, with the 10th observation deviating markedly from the trend established in the first 9 cases. Half of the participants received a form in which the 10th athlete was only heavier than the 9th by 3 kg but lifted 30 kg more. In the other forms the 10th athlete was 30 kg heavier than the 9th and lifted only 3 kg more. All forms included the following question:

> Do you find the relationship between body weight and lifted weight in the last entry surprising? (Yes/No) If you do, please change it to make it conform better to what you would expect. Mark your change on the sheet and return it.

Most subjects found the entry to be surprising and changed only one of the two items of information on the critical line. As predicted, a substantial majority (86%) of those who changed one item altered the weight lifted by the athlete rather than his body weight. The order of the two columns on the sheet did not matter.

The differential mutability of effects and causes suggests an apparently untested hypothesis concerning the social comparison process (Suls & Miller, 1977). People should prefer to compare themselves to others who resemble them on causal factors rather than to others who resemble them on outcome variables. For example, a student should be more interested in discovering how well a peer of comparable industry did on an exam than in discovering the industriousness of a student who achieved a comparable grade.

Focal and Background Actors

We propose that the mutability of any aspect of a situation increases when attention is directed to it and that unattended aspects tend to become part of the presupposed background. The hypothesis that the attributes of a focal object or agent are more mutable than those of nonfocal ones was explored by D. Read (1985). In the card-game study described earlier, he showed subjects the hands of two players, A and B, and asked them to "complete stems such as "A would have won if . . ." or "A would have lost if . . . ." The large majority of completions involved changes in the hand held by A, although the same outcome could have been generated just as well by altering B's hand. Other tests of the hypothesis used vignettes such as the following:

> Helen was driving to work along a three-lane road, where the middle lane is used for passing by traffic from both directions. She changed lanes to pass a slow-moving truck, and quickly realized that she was headed directly for another car coming in the opposite direction. For a moment it looked as if a collision was inevitable. However, this did not occur. Please indicate in one sentence how you think the accident was avoided.

The situation of the two cars is symmetric, or perhaps biased against Helen's being able to do much to prevent the accident, given that the circumstances of the other car are not described. Nevertheless, a substantial majority of subjects completed the story by ascribing the critical action to Helen.

The idea that the actions of a focal individual are mutable may help explain the well-documented tendency for victims of violence to be assigned an unreasonable degree of responsibility for their fate (Lerner & Miller, 1978). Information about a harmful act often presents the actions of the perpetrator in a way that makes them part of the presupposed background of the story, and therefore relatively immutable. Alternatives to the victim's actions are likely to be more mutable, and counterfactual scenarios in which the harm is avoided are therefore likely to be ones that change the victim's actions but keep the aggressor's behavior essentially constant. The high availability of such counterfactual scenarios can induce an impression that the victim is responsible for her fate—at least in the sense that she could easily have altered it. Any factor that increases the attention focused on the victim increases the availability of alternatives to the victim's reactions and the blame attached to the victim. The finding that emotional involvement with victims can increase the blame attributed to them (Lerner, 1980) is consistent with this speculation.

Our analysis is based on the idea that features of a situation that have highly available alternatives are attributed greater causal effectiveness than equally potent but less mutable factors. This analysis does not imply, of course, that causal responsibility is never assigned where it belongs. Probes that draw attention to the perpetrator, such as those concerning the punishment to be

given, can be expected to evoke appropriate degrees of blame. The point made here is simply that the actions of the perpetrator, unless made salient, are likely to be presupposed, with the result that constructed scenarios of how the victimization might have been avoided will focus on the actions of the victim. Although seemingly paradoxical, it is not a new discovery that apparent immutability reduces attributions of responsibility. This fact has been known to bullies since time immemorial.

The hypotheses that we have discussed illustrate, but do not exhaust, the factors that control the differential mutability of attributes and the differential availability of alternatives to reality. In other exploratory work we have found suggestive support for several additional hypotheses. One of these hypotheses is that temporal order affects mutability. Consider the pair of consonants XF. Now quickly change it by replacing one of the letters by another consonant. Which letter did you change? A robust majority of respondents replace the second letter rather than the first. This example illustrates a general rule: The second member of an ordered pair of events is likely to be more mutable than the first. Another hypothesis is that a change that can be visualized (e.g., undoing an accident) is more available than a change that cannot be visualized (e.g., undoing a heart attack).

In concluding this section, we note that the differential availability of counterfactual alternatives defines the grain of experience—the fault lines along which reality is likely to be undone in mental transformations (Hofstadter, 1985). In particular, differences of mutability determine which among the many alternatives to a particular occurrence will be viewed as normal. Mutability also determines which aspects of reality are likely to be accepted with relatively little questioning. In the next section we consider some effects of the differential availability of counterfactuals on the intensity of emotional responses.

Affective Role of Counterfactuals

This section develops some implications of the analysis of counterfactual thought for the domain of affect. We examine this question in conjunction with a hypothesis of *emotional amplification*, which states that the affective response to an event is enhanced if its causes are abnormal. In each of the following examples, the same misfortune is produced by two sequences of events, which differ in normality. The respondents assess the intensity of the affective responses that are likely to arise in the two situations. The first demonstration (Kahneman & Tversky, 1982) tested the prediction that outcomes that are easily undone by constructing an alternative scenario tend to elicit strong affective reactions.

> Mr. C and Mr. D were scheduled to leave the airport on different flights, at the same time. They traveled from town in the same limousine, were caught in a traffic jam, and arrived at the airport 30 minutes after the scheduled departure time of their flights. Mr. D is told that his flight left on time. Mr. C is told that his flight was delayed, and only left 5 minutes ago. Who is more upset?
>
> Mr. D   4%   Mr. C   96% (N = 138)

There is essentially unanimous agreement that Mr. C. is more upset than Mr. D., although their objective situations are identical—both have missed their planes. Furthermore, their expectations were also identical, since both had expected to miss their planes. The difference in the affective state of the two men appears to arise from the availability of a counterfactual construction. Mr. C. and Mr. D. differ in the ease with which they can imagine themselves—contrary to fact—catching up with their flights. This is easier for Mr. C., who needs only to imagine making up 5 minutes, than for Mr. D., who must construct a scenario in which he makes up half an hour. A preferred alternative is thus more available (normal) for Mr. C. than for Mr. D., which makes his experience more upsetting.

The next demonstration tests the prediction that outcomes that follow exceptional actions—and therefore seem abnormal—will elicit stronger affective reactions than outcomes of routine actions.

> Mr. Adams was involved in an accident when driving home after work on his regular route. Mr. White was involved in a similar accident when driving on a route that he only takes when he wants a change of scenery. Who is more upset over the accident?
>
> Mr. Adams   18%   Mr. White   82% (N = 92)

As predicted, the same undesirable outcome is judged to be more upsetting when the action that led to it was exceptional than when it was routine.

The next example tests the prediction that people are most apt to regret actions that are out of character:

> Mr. Jones almost never takes hitch-hikers in his car. Yesterday he gave a man a ride and was robbed. Mr. Smith frequently takes hitch-hikers in his car. Yesterday he gave a man a ride and was robbed. Who do you expect to experience greater regret over the episode?
>
> Mr. Jones   88%   Mr. Smith   12%
>
> Who will be criticized most severely by others?
>
> Mr. Jones   23%   Mr. Smith   77% (N = 138)

The results confirm the hypothesis and incidentally indicate that regret cannot be identified with an internalization of others' blame.

The final example (Kahneman & Tversky, 1982) tests the hypothesis that consequences of actions evoke stronger emotional responses than consequences of failures to act. The intuition to which this example appeals is that it is usually easier to imagine oneself abstaining from actions that one has carried out than carrying out actions that were not in fact performed.

> Mr. Paul owns shares in company A. During the past year he considered switching to stock in company B, but he decided against it. He now finds out that he would have been better off by $1,200 if he had switched to the stock of company B. Mr. George owned shares in company B. During the past year he switched to stock in company A. He now finds that he would have been better off by $1,200 if he had kept his stock in company B. Who feels greater regret?
>
> Mr. Paul   8%   Mr. George   92% (N = 138)

The finding that acts of commission produce greater regret than acts of omission was replicated by Landman (1984) and is in accord with formulations that distinguish omission from commission in attributions of causality and responsibility (Hart & Honoré, 1959; Heider, 1958).

Miller and McFarland (1986) conducted a series of studies to test the hypothesis that the abnormality of a victim's fate affects the sympathy that the victim receives from others. Subjects were told that the purpose of the studies was to provide victim compensation boards with information about the public's view of

various types of victims. Subjects were presented with a brief description of an incident and then were asked to indicate on an 11-point scale how much compensation they thought the victim should receive. Normality was manipulated in one study by varying the mutability of an action that led to a bad outcome. The victim was a man who had been severely injured during a robbery. In one condition, the robbery took place in the store at which the victim shopped most frequently. In a second condition, the robbery took place in a store to which the victim had decided to go only after finding his regular store closed for renovations. It was predicted that subjects would view the fate that befell the victim in the "unusual" store to be more abnormal, and hence more unjust, than the fate that befell the victim in the "usual" store. Consistent with this hypothesis, subjects recommended significantly more compensation (over $100,000 more) for the same injury in the exceptional context than they did in the routine context.

This study demonstrates that even morally charged judgments such as those involving compensation can be influenced by the normality of the outcome. It is as though a negative fate for which a more positive contrast is highly available is worse or more unfair than one for which there is no highly available positive alternative. It is important to note that the different reaction to the two victims is not due to the perceived probability of their fate. The probabilities of being shot in the two stores were both judged very low and indistinguishable from one another. The subjects apparently presupposed both the robbery and the store at which it took place. Given this presupposition, it is relatively more normal for an individual to be shot where it is normal for that individual to be—in the store that is regularly frequented.

A second compensation study, paralleling the missed-flight script, varied the distance between the negative outcome and a more positive alternative. The victim in this study had died from exposure after surviving a plane crash in a remote area. He had made it to within 75 miles of safety in one condition and to within ¼ mile in the second condition. Assuming that it is easier to imagine an individual continuing another ¼ mile than another 75 miles, it was predicted that the fate of the "close" victim would be perceived to be more abnormal, and hence more unfair, than the fate of the "distant" victim. The results supported the prediction, inasmuch as subjects once again recommended significantly more compensation for the family of the victim whose fate was more easily undone.

These results confirm the correlation between the perception of abnormality of an event and the intensity of the affective reaction to it, whether the affective reaction be one of regret, horror, or outrage. This correlation can have consequences that violate other rules of justice. An example that attracted international attention a few years ago was the bombing of a synagogue in Paris, in which some people who happened to be walking their dogs near the building were killed in the blast. Condemning the incident, a government official singled out the tragedy of the "innocent passers-by." The official's embarrassing comment, with its apparent (surely unintended) implication that the other victims were not innocent, merely reflects a general intuition: The death of a person who was not an intended target *is* more poignant than the death of a target. Unfortunately, there is only a small step from this intuition to the sense that the persons who are chosen as targets thereby lose some of their innocence.

Codes and Category Norms in Person Perception

The present approach, like several others (N. H. Anderson, 1981; Hastie & Kumar, 1979; Higgins & Lurie, 1983; Wyer & Gordon, 1984; Wyer & Srull, 1981), assumes the possibility of dual memory representations—raw memories of episodes and stored codes. Many features of a person can be stored (a) as comparative trait labels, which assign the individual to a position in an interpersonal norm, (b) as a set of episodes that define an intrapersonal norm of behavior for that individual, or (c) in both forms at once. In this section we pursue the implications of norm theory for two aspects of knowledge about persons: the ambiguity of codes and the formation and retrieval of intrapersonal norms. We discuss these issues in turn.

Norms and Ambiguous Codes

As was seen earlier in the discussion of comparative judgment, the interpretation of a comparative code is necessarily dependent on the norm to which the object of judgment is related. For example, hearing that "Jim has been given a long jail term" will suggest different jail terms to listeners, if they differ in the norm for jail terms that they attribute to the speaker. Communication will fail if speakers and listeners do not share, or at least coordinate, their norms. Coordination of norms is also involved when an individual uses a remembered code to reconstruct the literal detail of an experience—for example, the length of a jail sentence that is only remembered as long. Accurate performance depends on the match between the norm that is applied when the code is interpreted and the norm that supported the initial judgment. As Higgins and Lurie (1983) demonstrated in an impressive experiment, the reconstruction of the initial episode will be systematically biased if the norm changes in the interval. This effect, which Higgins and Lurie termed *change of standard*, can yield a range of cognitive and affective responses, including the disappointment that people often experience when they meet a former teacher whom they had always remembered as brilliant (Higgins & King, 1981).

Trait labels and expressions fall into two categories: (a) relative predicates, which specify the individual's position on an interpersonal norm, and (b) absolute predicates, which summarize an intrapersonal norm of actions or feelings on relevant occasions. The same trait name can sometimes serve in both functions. The statement that "Jane is assertive" can be understood as saying either that she is more assertive than most people or that her behavior is assertive on most occasions. In the latter interpretation, the word "assertive" is a category label, which evokes exemplars of assertive behaviors.

Trait labels that have both an absolute and a relative sense are potentially ambiguous. This ambiguity appears to underlie the tendency of people to accept general descriptions as uniquely relevant to them, also known as the Barnum effect (Snyder, Shenkel, & Lowery, 1977). Barnum statements typically evoke both absolute and relative interpretations. For example, the statement "You are shy in the presence of strangers" can be recognized by most people as a valid description of themselves—if shyness indicates that one is less comfortable with strangers than with familiar others. In its relative sense, of course, the description is applicable only to the minority of people who are

sufficiently extreme to deserve special mention. It is the validity of the Barnum statement in its absolute sense that makes it believable to the individual, but it is the unwarranted extension to the interpersonal comparison that makes it interesting. A similar mixture of meanings was noted by Higgins and Winter (cited in Kraut & Higgins, 1984) in their analysis of trait ambiguity. They asked subjects what percentage of people possessed various personality traits (e.g., friendliness, aggressiveness) and found many traits that were assumed to apply to an absolute majority of people. It appears that the assignment of these traits involves a mixture of comparative and absolute criteria.

Trait descriptions vary in the degree to which they lend themselves to interpersonal and intrapersonal interpretation. For example, the expression "He is not very intelligent" evokes an interpersonal norm, but the statement "I like Coke" has an intrapersonal reference (more than other beverages, not necessarily more than other people). The difference has a predictable effect on the impact that information about others has on self-description. Whether a person is more or less intelligent than her reference group will influence how intelligent she judges herself to be (Davis, 1966), but whether a person consumes more or less of a drink than her reference group will not influence her expressed liking for the drink (Hansen & Donoghue, 1977; Nisbett, Borgida, Crandall, & Reed, 1976).

Norms From Single Elements

The present model of category norms assumes both that a single instance suffices to set up a norm and that a reference to the category label serves to restrict consideration to members of that category. These assumptions entail nonregressive predictions and a tendency to neglect relevant base rates.

To illustrate, imagine that you were shown a single exemplar of an unfamiliar species of insect, which was larger than any other insect you have ever seen. How large would you expect another exemplar of that species to be? Would it occur to you that the single insect you saw is likely to be larger than most of its conspecifics? The canons of inference prescribe that the single insect you saw is likely to be larger than most others in its species. Predictions of the size of the next member of the new species should accordingly regress toward the general insect norm. Intuitive expectations, in contrast, are firmly centered on the single observed value, which constitutes the norm for that category. In the context of social judgment, the ease with which category norms are established leads to radical generalization from a single observation of behavior to an interpersonal norm for the behavior of other people in the same setting, or to a norm for a person's behavior on future occasions (Nisbett & Ross, 1980; S. Read, 1984).

Radical generalization to interpersonal norms was neatly demonstrated by Hamill, Wilson, and Nisbett (1980), who showed that subjects centered their norms for an entire social group (prison guards, in one study) on an observation of a single member, even when he was described as atypical. In the same vein, Nisbett and Borgida (1975) exposed observers to brief and innocuous interviews with two students, allegedly participants in an experiment on helping behavior. From the information that these two individuals had not helped a stranger in distress, the observers generalized that most people would also fail to help in the same situation. This generalization occurred despite the fact that the unhelpful behavior contradicted the observers' prior beliefs and expectancies.

Radical generalization from observed behavior to an intrapersonal norm is manifest in the nonregressiveness of behavioral predictions (Kahneman & Tversky, 1973; Ross & Anderson, 1982). A single observed instance of unusually generous tipping appears sufficient to set up a norm that will be consulted in predictions of a person's future tips. The interpersonal base rate for the behavior of other people is effectively excluded from consideration when thinking about this person's tips, just like one's knowledge of the size of insects in the previous example.

Radical generalizations can be made with high and unwarranted confidence. Experimental results suggest that subjective confidence depends mainly on the consistency of available evidence rather than on its quality or quantity (Einhorn & Hogarth, 1982; Kahneman & Tversky, 1973; Tversky & Kahneman, 1971). Subjective confidence in a prediction is likely to reflect the *breadth,* or variability, of the category norm on which the judgment is based (see Figure 1). Confidence will be high if all elements cluster around a single value of the attribute, independently of the number of these elements. However, it would be incorrect to conclude that judgments are entirely insensitive to the quantity of evidence. The model of Figure 1 implies, in particular, that the susceptibility of a norm to change depends on the absolute *height* of the availability profile, and thus on the number of elements in the norm. A category norm that is based on one or two elements can support confident judgments, but it can also be altered relatively easily under the impact of new evidence.

Many inferences bridge across situations and across attributes. In the absence of better evidence, people readily predict success in graduate school from an IQ test score, research productivity from performance in a colloquium, or the size of a mother's graduation gift to her daughter from the size of a tip that she gave to a waiter. Such generalizations are involved both in deliberate predictions of future behaviors and in spontaneous inferences about traits and about unobserved aspects of situations.

Inferences that bridge attributes are no more regressive than direct generalizations (Nisbett & Ross, 1980; Ross & Anderson, 1982). Nonregressive predictions can be generated in two ways. First, the predicted value may be chosen so that its position matches that of the known attribute in their respective interpersonal norms (Kahneman & Tversky, 1973). Thus, a predicted graduation gift can be chosen that is as extreme in the norm for gifts as the observed tip is in the interpersonal norm for tips. Second, nonregressive predictions can also be generated by matching descriptive labels (e.g., if a tip is remembered as very generous, what graduation gift would be considered equally generous?). The intention to predict a behavior elicits a search for relevant incidents or for pertinent descriptive labels. The search is guided by the similarity of potential elements to the target attribute, and it is probably concentric: The nearest incidents or labels that turn up in the search are used, nonregressively, to generate a prediction. In the absence of better data, people are willing to make extreme predictions from evidence that is both flimsy and remote. The process of concentric search yields radical (and overconfident) predictions from observations of dubious relevance. However, the same process also makes it likely that distant labels or incidents will be ignored when evidence that is closer to the target attribute is available. Thus, generous tipping

habits will not be given much weight in predicting a mother's graduation gift to her daughter if pertinent incidents of their interaction can be retrieved.

Unwanted Elements of Norms

The elicitation of a norm has been described here as a process of parallel activation of multiple representations that are recruited by a stimulus or by a category. We have assumed that the process of recruitment is rapid, automatic, and essentially immune to voluntary control after its initiation. In particular, it is not possible for the individual to sift through activated elements and discard irrelevant or misleading ones. This limitation on the voluntary control of norms helps explain two well-documented phenomena of social judgment: the perserverance of discredited beliefs and the correspondence bias in person perception.

Imagine a discussion of a Canadian athlete, in which someone who is unfamiliar with metric measures reads from a sheet: "Brian weighs 102 kg. That's 280 lb, I think. No, it's actually about 220 lb." Will the speaker's initial error affect listeners' subsequent responses to questions about Brian's size and strength? The literature on perseverance of discredited beliefs (C. A. Anderson, 1983; Fleming & Arrowood, 1979; Ross & Anderson, 1982; Ross & Lepper, 1980; Schul & Bernstein, 1985) suggests that it will. The message of this literature is that traces of an induced belief persist even when its evidential basis has been discredited. The discarded message is not erased from memory, and the norm elicited by a subsequent question about Brian's weight could therefore contain the original message as well as its correction. Thus, a listener might "know" immediately after the message that Brian's true weight is 220 lb, and this value would presumably retain an availability advantage, but the category norm associated with Brian's weight would still be biased toward the erroneous value of 280 lb. Judgments that depend indirectly on the activation of the norm would be biased as well.

Some aspects of the phenomenon known as the *correspondence bias* (Jones, 1979) or the *fundamental attribution error* (Ross, 1977) could be explained in similar terms. Many studies have shown that people make unwarranted inferences concerning personal traits and attitudes from observations of behavior that is in fact entirely constrained by the situation. Subjects in one famous series of experiments in this tradition observed an individual who was explicitly instructed to write an essay or to read aloud a speech advocating an unpopular position (Jones, 1979; Jones & Harris, 1967; Jones & McGillis, 1976). In response to subsequent questions, observers commonly attributed to target persons an attitude consistent with the position that these persons had been constrained to advocate.

In these experiments, as in studies of discredited beliefs, a behavioral observation is accompanied by information that challenges its validity. As in a theatrical performance, the actor in the experiments of Jones and his colleagues engages in a behavior (e.g., advocating the regime of Fidel Castro) that does not have its usual significance because of the special demands of the situation. We propose that both in the theater and in these experiments two traces are laid down: (a) a literal memory of the person expressing pro-Castro sentiments and (b) a memory of the behavior in the context of the situational constraints. Both memories are elements of the set that will be evoked by further observations of the actor's political opinions or by a question

concerning those opinions. As in the preceding example, this norm will be biased even for an observer who "knows" that the actor's behavior is constrained, and therefore uninformative. Quite simply, pro-Castro behaviors are more normal for the actor than for a random stranger. Belief perseverance, generalizations from atypical examples, and failures of discounting all illustrate the same principle: Any observation of behavior—even if it is discounted or discredited—increases the normality of subsequent recurrences of compatible behaviors.

Causal Questions and Answers

In this section we consider the role of norms in causal judgments. This topic was chosen to illustrate the concepts of presupposition and norm coordination that were introduced earlier. We begin by examining a routine conversational exchange, which provides a conceptual model for much attribution research. A questioner, whom we call Quentin (Q), asks a why question and receives an answer from Ann (A). An example might be:

Q: "Why did Joan pass *this* math exam?"
A: "She used the Brown textbook."

We focus on two issues: (a) the inferences that A must make about Q's norms to interpret a why question and (b) the constraints that this interpretation of the question places on appropriate answers.

Norms and Causal Questions

Causal questions about particular events are generally raised only when these events are abnormal. The close connection between causal reasoning and norms is evident in the rules that govern the homely why question as it is used in conversations about particular events (Lehnert, 1978). The why question implies that a norm has been violated. Thus, the question "Why was John angry?" indicates Q's belief that it was normal for him not to be, and the question "Why was John not angry?" indicates the contradictory belief. Even the question "Why is John so normal?" implies that he is normal to an abnormal degree. A why question, then, presupposes that some state X is the case, and also implies an assertion that not-X was normal. The strongest indication of the implicit assertion of a norm is that the why question, unlike most others, can be denied. The denial can be expressed by a question, as in the familiar exchange:

Q: "Why do you so often answer a question with a question?"
A: "Why not?"

The why-not retort denies the assertion that not-X is normal. It is a legitimate answer that, if accepted, leaves nothing to be explained.

We suggest that why questions (at least those of the deniable variety, for which "why not?" is a sensible answer) are not requests for the explanation of the occurrence or nonoccurrence of an *event*. A why question indicates that a particular event is surprising and requests the explanation of an *effect*, defined as a contrast between an observation and a more normal alternative. A successful explanation will eliminate the state of surprise. This is commonly done in one of three ways. First, A may deny the implied assertion that X is abnormal in the light of what Q already knows. This is the why-not answer, which invites Q to

change his opinion that X is abnormal. Second, A may inform Q of some fact of which Q was previously ignorant, which makes the occurrence of X normal. For example, Q may not have known that Joan used the Brown text, which he knows to be excellent. Third, A may indicate that there is a causal link, of which Q was presumably ignorant, between the effect X and some known aspect of the situation. For example, Q may know that Joan had used the Brown textbook, but he may need to be told that it is excellent.

The choice of causal feature is constrained in many ways, which have been extensively discussed by philosophers (for particularly relevant treatments, see Hart & Honoré, 1959; Mackie, 1974) and by psychologists (Einhorn & Hogarth, 1986; Kelley, 1967; Schustack & Sternberg, 1981). We shall not be concerned with the factors that determine impressions of causal efficacy. We focus here on a constraint that relates directly to the notion of norm: A cause must be an event that could easily have been otherwise. In particular, a cause cannot be a default value among the elements that the event X has evoked. The rule that a default value cannot be presented as a cause was noted by Hart and Honoré (1959), who observed that the statement "It was the presence of oxygen that caused the fire" makes sense only if there were reasons to view the presence of oxygen as abnormal. It is important to note, however, that a property need not be statistically unusual to serve as an explanation; it is only precluded from being a default value. Peculiar behaviors of cars observed on the road are frequently "explained" by reference to the drivers being young, elderly, or female, although these are hardly unusual cases. The default value for an automobile driver appears to be middle-aged male, and driving behavior is rarely explained by it.

Ambiguities in Causal Questions

Conversations in general, and answers to questions in particular, are governed by subtle rules that determine what is said and what is presupposed or implicated (Clark, 1979; Grice, 1975). The situation is especially complicated when the conversation is actually a test, as is frequently the case in psychological experiments. The unique feature of a test is that the questioner is not ignorant or puzzled, as questioners usually are. When the question is ambiguous, the respondent faces the bewildering task of choosing a state of ignorance for the questioner.

The why question appears to be especially susceptible to ambiguities. Consider a perennial favorite of attribution research: "Why did Ralph trip on Joan's feet?" (McArthur, 1972). The event in question is clearly specified, but the effect—defined as a contrast between the event and its norm—is not. To answer this question, the respondent must first identify what it is that the experimenter considers surprising. In everyday conversations intonation provides a potent cue to the intended meaning of a question and to the violated presupposition that underlies it. It is instructive to read the question about Ralph and Joan aloud several times, each time stressing a different word. The location of the major stress substantially reduces the number of possible interpretations, although it does not suffice to disambiguate the question completely. For example, the reading "Why did *Ralph* trip over Joan's feet?" suggests either that (a) it is unusual for Joan's partners to trip over her feet or that (b) although Joan's

partners usually trip over her feet, there was special reason to expect Ralph to be more fortunate.

An experimental demonstration of the ambiguity of why questions was described by Miller (1981). Several groups of student and graduate nurses were asked to explain their decision to enter the nursing profession. Different versions of the same basic question were used in the different groups. The basic question was "Why did you go into nursing?" An analysis of the answers to this question indicated that student nurses cited significant aspects of nursing (e.g., "it is a respected profession") more often than did graduate nurses. On the other hand, the graduate nurses were more likely to cite personal qualities (e.g., "I like people"). The critical finding of Miller's study was that the differences between students and graduate nurses vanished when they were asked questions that explicitly specified the relevant norm: "Why did you decide to go into nursing rather than some other profession?" or "Why did you decide to go into nursing when most of your friends did not?" As expected, the former elaboration yielded a majority of answers for both groups that referred to nursing, whereas the answers to the second question referred predominantly to personal dispositions. In view of this result, the contrasting answers of the two groups to the unelaborated why question appear to reflect different interpretations of an ambiguous question rather than different causal beliefs.

Questioners convey cues that broadly specify the content of the causal answers that they wish to receive (Lehnert, 1978). For example, the questions "Why did Carter lose the 1980 election?" and "Why did Reagan win the 1980 election?" refer to the same event but differ in the explanation that they request: some noteworthy fact about Carter in the first question, about Reagan in the second. In the absence of indications to the contrary, the subject of the sentence is supposed to be its focus (Pryor & Kriss, 1977), and the syntactical form of the question suggests an equivalent form for the answer.

Perspective Differences

The coordination of the norms that apply to an effect and to a proposed cause is illustrated in an example discussed by the legal philosophers Hart and Honoré in their classic *Causation in the Law* (1959):

> A woman married to a man who suffers from an ulcerated condition of the stomach might identify eating parsnips as the cause of his indigestion. The doctor might identify the ulcerated condition as the cause and the meal as a mere occasion. (p. 33)

The causes chosen by his wife and by the physician refer to the same event but explain different effects. It is evident from her answer that the wife is concerned with an exception to an intrapersonal norm: "Why does he have indigestion today but not on other days?" The physician, on the other hand, is concerned with an interpersonal norm: "Why does this patient suffer from indigestion when others do not?" The difference could reflect the role of availability in the recruitment of norm elements: The wife is likely to retrieve many memories of recent days on which her husband, although known to have an ulcer, did not suffer indigestion. These memories, which resemble the present occasion in most respects, will define a norm for it. The physician, of course, is unlikely to have had the same amount of exposure

to the patient. According to the rule that coordinates causes to effects (and explanations to questions), the wife chooses as a cause a property that distinguishes this particular day from other days, and the physician selects a feature that distinguishes this patient from other patients.

The situational attribution made by the wife and the dispositional attribution made by the physician in Hart and Honoré's example recall the actor–observer differences described by Jones and Nisbett (1971). Actors often explain their actions and attitudes by reference to eliciting properties of situations, whereas observers of the same actions and attitudes attribute them to the actor's distinctive characteristics. As in Hart and Honoré's example, the situational attribution corresponds to an intrapersonal norm, whereas a dispositional attribution of the same behavior relates it to an interpersonal norm. The intuitions about differential availability that make the indigestion example so compelling apply as well to the case of actors and observers. The question "Why do you like this particular girl?" appears to favor the recruitment of thoughts about the respondent's attitude toward other girls. The question "Why does he like this particular girl?" is more likely to evoke in an observer thoughts of the attitudes of other people toward that girl (Nisbett, Caputo, Legant, & Maracek, 1973). The different elements that are evoked produce quite different questions: "Why do you like this girl more than most other girls?" and "Why does he like this girl more than most others do?" Each question, in turn, constrains appropriate answers to factors that vary among the elements of the evoked set—other girls for the actor, other individuals for the observer.

The intuitions about the wife–physician example cannot be reduced to the accounts commonly offered for actor–observer differences. The contrast could not be explained by difference of knowledge (Jones & Nisbett, 1971) or of perceptual salience (Arkin & Duval, 1975; Storms, 1973). It is not explained by the distinction between a state of self-consciousness and other states of consciousness (Duval & Wicklund, 1972). Nor is it compatible with the hypothesis that the focus of attention is assigned a dominant causal role (Fiske, Kenny, & Taylor, 1982; Ross, 1977; Taylor & Fiske, 1978), inasmuch as the husband surely plays a more focal emotional role for the wife than for the physician.

The hypothesis of the present treatment is that the same event evokes different norms in the wife and the physician of the example, and in actors and observers in other situations. Different descriptions of the same event can appear to provide conflicting answers to the same question, when in fact they are concerned with different questions. This proposal can be subjected to a simple test: Do the observers actually disagree? A negative answer is suggested by several studies. Nisbett et al. (1973) found that subjects easily adopt a typical observer perspective in reporting how their choice of a girlfriend would be described by a close friend. Other data confirm the ability of actors to mimic observers (Miller, Baer, & Schenberg, 1979). On the other hand, observers instructed to empathize with one of the participants in a dialogue tend to adopt an actor perspective in explaining that person's behavior (Regan & Totten, 1975).

In summary, we contend that there are a number of advantages to viewing the process of causal reasoning from the perspective of norms. First, our analysis provides an account of the antecedents of causal reasoning. A search for explanation may occur spontaneously when a significant event violates a norm that it evokes (see also Hastie, 1984; Weiner, 1985). Causal search can also be prompted by a question that presupposes a violated norm (see Lalljee & Abelson, 1983). Second, the present approach draws attention to the difficulty of assessing the accuracy of attributers who differ in their perspectives (Funder, 1982; Monson & Snyder, 1977). It is important to distinguish real disagreements in causal attributions from specious disagreements that arise when people answer different questions. Finally, the present analysis identifies a necessary feature of any factor that is considered a possible cause of a surprising event: A cause cannot be a default value of the norm that the consequence has evoked.

Concluding Remarks

This essay has proposed a theory of norms. The two main functions of norms are the representation of knowledge of categories and the interpretation of experience. We have challenged the conception of norms as precomputed structures and have suggested that norms—and sometimes even their elements—are constructed on the fly in a backward process that is guided by the characteristics of the evoking stimulus and by the momentary context. In this regard our treatment resembles other approaches that emphasize the role of specific episodes and exemplars in the representation of categories (Barsalou, in press; Brooks, 1978; Hintzman, in press; Jacoby & Brooks, 1984; McClelland & Rumelhart, 1985; Medin & Schaffer, 1978; Schank, 1982). A distinctive aspect of the present analysis is the separation of normality and post hoc interpretation, on the one hand, from probability and anticipation, on the other. Another distinctive aspect is our attempt to identify the rules that determine which attributes of experience are immutable and which are allowed to vary in the construction of counterfactual alternatives to reality. Our closest neighbor in this enterprise is Hofstadter (1979, 1985), with his highly evocative treatment of what he calls slippability. Like him, we believe that it is "very hard to make a counterfactual world in which counterfactuals were not a key ingredient of thought" (Hofstadter, 1985, p. 239).

Our current understanding of the rules for the retrieval of norm elements far exceeds our understanding of the rules for the construction of unrealized alternatives. We believe that the roles of presuppositions and mutability in counterfactual thought define a promising area for future research. Norms and cognate concepts have often been applied to the study of comparative judgment and personality description, but we have argued that the concept is also central to numerous other processes, including affective reactions and causal reasoning.

References

Alba, J. W., & Hasher, L. (1983). Is memory schematic? *Psychological Bulletin, 93*, 203–231.

Anderson, C. A. (1983). Abstract and concrete data in the perseverance of social theories: When weak data lead to unshakeable beliefs. *Journal of Experimental Social Psychology, 19*, 93–108.

Anderson, J. R. (1983). A spreading activation theory of memory. *Journal of Verbal Learning and Verbal Behavior, 22*, 261–295.

Anderson, N. H. (1981). *Foundations of information integration theory.* New York: Academic Press.

Appley, M. H. (Ed.). (1971). *Adaptation-level theory: A symposium.* New York: Academic Press.

Arkin, R. A., & Duval, S. (1975). Effects of focus of attention on the causal attributions of actors and observers. *Journal of Experimental Social Psychology, 11,* 427–438.

Avant, L. L., & Helson, H. (1973). Theories of perception. In B. B. Watson (Ed.), *Handbook of general psychology* (pp. 419–448). Englewood Cliffs, NJ: Prentice Hall.

Barsalou, L. W. (1983). Ad hoc categories. *Memory and Cognition, 11,* 211–227.

Barsalou, L. W. (1985). Ideals, central tendency, and frequency of instantiation as determinants of graded structure in categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 11,* 629–654.

Barsalou, L. W. (in press). The instability of graded structure: Implications for the nature of concepts. In U. Neisser (Ed.), *Concepts reconsidered: The ecological and intellectual bases of categories.* New York: Cambridge University Press.

Birnbaum, M. H. (1982). Controversies in psychological measurement. In B. Wegener (Ed.), *Social attitudes and psychophysical measurement* (pp. 401–485). Hillsdale, NJ: Erlbaum.

Brooks, L. R. (1978). Nonanalytic concept formation and memory for instances. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization* (pp. 169–215). Hillsdale, NJ: Erlbaum.

Brooks, L. R. (in press). Decentralized control of categorization: The role of prior processing episodes. In U. Neisser (Ed.), *Concepts reconsidered: The ecological and intellectual bases of categories.* New York: Cambridge University Press.

Brown, D. R., & Reich, C. M. (1971). Individual differences in adaptation-level theory. In M. H. Appley (Ed.), *Adaptation-level theory: A symposium* (pp. 215–231). New York: Academic Press.

Clark, H. H. (1979). Responding to indirect speech acts. *Cognitive Psychology, 11,* 430–474.

Clark, H. H., & Haviland, S. E. (1977). Comprehension and the given-new contract. In R. O. Freedle (Ed.), *Discourse production and comprehension* (pp. 1–40). Norwood, NJ: Ablex.

Coren, S., & Girgus, J. S. (1978). *Seeing is deceiving: The psychology of visual illusions.* Hillsdale, NJ: Erlbaum.

Coren, S., & Miller, J. (1974). Size contrast as a function of figural similarity. *Perception and Psychophysics, 16,* 355–357.

Davis, J. A. (1966). The campus as a frog pond: An application of the theory of relative deprivation to career decisions of college men. *American Journal of Sociology, 72,* 17–31.

Duval, S., & Wicklund, R. A. (1972). *A theory of objective self-awareness.* New York: Academic Press.

Eich, J. M. (1982). A composite holographic associative recall model. *Psychological Review, 89,* 627–661.

Einhorn, H. J., & Hogarth, R. M. (1982). *A theory of diagnostic inference: I. Imagination and the psychophysics of evidence.* Center for Decision Research, Graduate School of Business, University of Chicago.

Einhorn, H. J., & Hogarth, R. M. (1986). Probable cause: A decision making framework. *Psychological Bulletin, 99,* 000–000.

Fischhoff, B. (1975). Hindsight is not equal to foresight: The effects 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: Heuristics and biases in hindsight. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 335–351). New York: Cambridge University Press.

Fiske, S. T., Kenny, D. A., & Taylor, S. E. (1982). Structural models for the mediation of salience effects on attribution. *Journal of Experimental Social Psychology, 18,* 105–127.

Fleming, J., & Arrowood, A. J. (1979). Information processing and the perseverance of discredited self-perceptions. *Personality and Social Psychology Bulletin, 5,* 201–205.

Fried, L. S., & Holyoak, K. J. (1984). Induction of category distributions:

A framework for classification learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10,* 234–257.

Funder, D. C. (1982). On the accuracy of dispositional vs. situational attributions. *Social Cognition, 1,* 205–222.

Garner, W. R. (1962). *Uncertainty and structure as psychological concepts.* New York: Wiley.

Garner, W. R. (1970). Good patterns have few alternatives. *American Scientist, 58,* 34–42.

Goldstein, W. M. (1982). *Inconsistent assessments of preference: Effects of stimulus presentation method on the preference reversal phenomenon.* Unpublished manuscript, University of Michigan.

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics: Speech acts* (pp. 41–58). New York: Academic Press.

Hamill, R., Wilson, T. D., & Nisbett, R. E. (1980). Insensitivity to sample bias: Generalizing from atypical cases. *Journal of Personality and Social Psychology, 39,* 578–589.

Hansen, R. D., & Donoghue, J. M. (1977). The power of consensus: Information derived from one's own and other's behavior. *Journal of Personality and Social Psychology, 35,* 294–302.

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

Hastie, R. (1981). Schematic principles in human memory. In E. T. Higgins, C. P. Herman, & M. P. Zanna (Eds.), *Social cognition: The Ontario symposium* (Vol. 1, pp. 39–88). Hillsdale, NJ: Erlbaum.

Hastie, R. (1984). Causes and effects of causal attribution. *Journal of Personality and Social Psychology, 46,* 44–56.

Hastie, R., & Kumar, P. A. (1979). Person memory: Personality traits as organizing principles in memory for behaviors. *Journal of Personality and Social Psychology, 37,* 25–38.

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

Helson, H. (1964). *Adaptation level theory: An experimental and systematic approach to behavior.* New York: Harper.

Helson, H., & Kozaki, A. (1968). Effects of duration of series and anchor stimuli on judgments of perceived size. *American Journal of Psychology, 81,* 291–302.

Higgins, E. T., & King, G. (1981). Accessibility of social constructs: Information-processing consequences of individual and contextual variability. In N. Cantor & J. F. Kihlstrom (Eds.), *Personality, cognition and social interaction* (pp. 69–121). Hillsdale, NJ: Erlbaum.

Higgins, E. T., & Lurie, L. (1983). Context, categorization, and memory: The "change-of-standard" effect. *Cognitive Psychology, 15,* 525–547.

Hinton, G. E., & Anderson, J. A. (Eds.). (1981). *Parallel models of associative memory.* Hillsdale, NJ: Erlbaum.

Hintzman, D. L. (in press). "Schema abstraction" in a multiple-trace memory model. *Psychological Review.*

Hintzman, D. L., & Ludlam, G. (1980). Differential forgetting of prototypes and old instances: Simulation by an exemplar-based classification model. *Memory and Cognition, 8,* 378–382.

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

Hofstadter, D. R. (1985). *Metamagical themas: Questing for the essence of mind and pattern.* New York: Basic Books.

Jacoby, L. L., & Brooks, L. R. (1984). Nonanalytic cognition: Memory, perception, and concept learning. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 18, pp. 1–47). New York: Academic Press.

Johnson, M. K. (1983). A modular model of memory. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 17, pp. 81–123). New York: Academic Press.

Jones, E. E. (1979). The rocky road from acts to dispositions. *American Psychologist, 34,* 107–117.

Jones, E. E., & Harris, V. A. (1967). The attribution of attitudes. *Journal of Experimental Social Psychology, 3,* 1–24.

Jones, E. E., & McGillis, D. (1976). Correspondent inference and the attribution cube: A comparative appraisal. In J. Harvey, W. Ickes, & R. Kidd (Eds.), *New directions in attribution research* (Vol. 1, pp. 389–420). Hillsdale, NJ: Erlbaum.

Jones, E. E., & Nisbett, R. E. (1971). *The actor and the observer: Divergent perceptions of the causes of behavior.* Morristown, NJ: General Learning Press.

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

Kahneman, D., & Tversky, A. (1982). 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.

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.

Kraut, R. E., & Higgins, E. T. (1984). Communication and social cognition. In R. S. Wyer & T. K. Srull (Eds.), *Handbook of social cognition* (Vol. 3, pp. 81–127). Hillsdale, NJ: Erlbaum.

Lakoff, G. (in press). *Women, fire, and dangerous things.* Chicago: University of Chicago Press.

Lalljee, M., & Abelson, R. P. (1983). The organization of explanations. In M. Hewstone (Ed.), *Attribution theory: Social and functional extensions* (pp. 65–80). Oxford, England: Blackwell.

Landman, J. (1984). *Regret and undoing: Retrospective assessment of hypothetical and real life events.* Unpublished doctoral dissertation, University of Michigan.

Lehnert, W. (1978). *The process of question answering.* Hillsdale, NJ: Erlbaum.

Lerner, M. J. (1980). *The belief in a just world.* New York: Plenum.

Lerner, M. J., & Miller, D. T. (1978). Just world research and the attribution process: Looking back and ahead. *Psychological Bulletin, 85,* 1030–1051.

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

Lockhead, G. R., & King, M. C. (1983). A memory model of sequential effects in scaling. *Journal of Experimental Psychology: Human Perception and Performance, 9,* 461–473.

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

Mandler, G. (1967). Organization and memory. In K. W. Spence & J. A. Spence (Eds.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 1, pp. 328–372). New York: Academic Press.

Mandler, G. (1975). Memory storage and retrieval: Some limits on the reach of attention and consciousness. In P. M. A. Rabbitt & S. Dornic (Eds.), *Attention and performance V* (pp. 499–516). New York: Academic Press.

Mandler, J. M. (1984). *Stories, scripts, and scenes: Aspects of schema theory.* Hillsdale, NJ: Erlbaum.

McArthur, L. A. (1972). The how and what of why: Some determinants and consequences of causal attribution. *Journal of Personality and Social Psychology, 22,* 171–193.

McClelland, J. L. (1985). Putting knowledge in its place: A scheme for programming processing structures on the fly. *Cognitive Science, 9,* 113–146.

McClelland, J. L., & Rumelhart, D. E. (1985). Distributed memory and the representation of general and specific information. *Journal of Experimental Psychology: General, 114,* 159–188.

Medin, D. L. (1983). Structural principles in categorization. In T. J. Tighe & B. E. Shepp (Eds.), *Perception, cognition, and development: Interactional analyses* (pp. 203–230). Hillsdale, NJ: Erlbaum.

Medin, D. L., Altom, M. W., Edelson, S. M., & Freko, D. (1982). Correlated symptoms and simulated medical classification. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 8,* 37–50.

Medin, D. L., & Schaffer, M. M. (1978). Context theory of classification learning. *Psychological Review, 85,* 207–238.

Mellers, B. A. (1982). Equity judgments: A revision of Aristotelian views. *Journals of Experimental Psychology: General, 111,* 242–270.

Mellers, B. A., & Birnbaum, M. H. (1983). Contextual effects in social judgment. *Journal of Experimental Social Psychology, 19,* 157–171.

Metcalfe Eich, J. (1985). Levels of processing, encoding specificity, elaboration, and CHARM. *Psychological Review, 92,* 1–38.

Miller, A. G., Baer, R., & Schenberg, P. (1979). The bias phenomenon in attitude attribution: Actor and observer perspectives. *Journal of Personality and Social Psychology, 37,* 1421–1431.

Miller, D. T. (1981, August). *Changes over time in the selection of causes and the definitions of effects.* Paper presented at the meeting of the American Psychological Association, Los Angeles.

Miller, D. T., & McFarland, C. (1986). *Counterfactual thinking and victim compensation.* Manuscript in preparation.

Monson, T. C., & Snyder, M. (1977). Actors, observers and the attribution process: Toward a reconceptualization. *Journal of Experimental Social Psychology, 13,* 89–111.

Murdock, B. B. (1982). A theory for the storage and retrieval of item and associative information. *Psychological Review, 89,* 609–626.

Murphy, G. L., & Medin, D. L. (1985). The role of theories in conceptual coherence. *Psychological Review, 92,* 289–316.

Nisbett, R. E., & Borgida, E. (1975). Attribution and the psychology of prediction. *Journal of Personality and Social Psychology, 32,* 932–943.

Nisbett, R. E., Borgida, E., Crandall, R., & Reed, H. (1976). Popular induction: Information is not always informative. In J. Carrol & J. Payne (Eds.), *Cognition and social behavior* (pp. 227–236). Hillsdale, NJ: Erlbaum.

Nisbett, R. E., Caputo, C., Legant, P., & Maracek, J. (1973). Behavior as seen by the actor and as seen by the observer. *Journal of Personality and Social Psychology, 27,* 154–164.

Nisbett, R. E., & Ross, L. (1980). *Human inference: Strategies and shortcomings of social judgment.* Englewood Cliffs, NJ: Prentice-Hall.

Osherson, D. N., & Smith, E. E. (1982). Gradedness and conceptual combination. *Cognition, 12,* 299–318.

Parducci, A. (1956). Direction of shift in the judgment of single stimuli. *Journal of Experimental Psychology, 51,* 169–178.

Parducci, A. (1965). Category judgment: A range-frequency model. *Psychological Review, 72,* 407–418.

Parducci, A. (1974). Contextual effects: A range-frequency analysis. In E. C. Carterette & M. P. Friedman (Eds.), *Handbook of perception: Psychophysical judgment and measurement* (Vol. 7, pp. 127–141). New York: Academic Press.

Posner, M. I., & Keele, S. W. (1968). On the genesis of abstract ideas. *Journal of Experimental Psychology, 77,* 353–363.

Pryor, J. B., & Kriss, N. (1977). The cognitive dynamics of salience in the attribution process. *Journal of Personality and Social Psychology, 35,* 49–55.

Ratcliff, R., & McKoon, G. (1981). Does activation really spread? *Psychological Review, 88,* 454–457.

Read, D. (1985). *Determinants of relative mutability.* Unpublished research, University of British Columbia, Vancouver, Canada.

Read, S. J. (1984). Analogical reasoning in social judgment: The importance of causal theories. *Journal of Personality and Social Psychology, 46,* 14–25.

Regan, D. T., & Totten, J. (1975). Empathy and attribution: Turning observers into actors. *Journal of Personality and Social Psychology, 32,* 850–856.

Restle, F. (1978a). Assimilation predicted by adaptation-level theory with variable weights. In N. J. Castellan, Jr., & F. Restle (Eds.), *Cognitive theory* (Vol. 3, pp. 75–91). Hillsdale, NJ: Erlbaum.

Restle, F. (1978b). Relativity and organization in visual size judgments. In E. Leeuwenberg & H. Buffart (Eds.), *Formal theories of visual perception* (pp. 247–263). New York: Wiley.

Rips, L. R., & Turnbull, W. (1980). How big is big? Relative and absolute properties in memory. *Cognition, 8,* 145–174.

Rosch, E. (1973). On the internal structure of perceptual and semantic categories. In T. E. Moore (Ed.), *Cognitive development and the acquisition of language* (pp. 111–144). New York: Academic Press.

Rosch, E. (1978). Principles of categorization. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization* (pp. 27–48). Hillsdale, NJ: Erlbaum.

Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology, 7,* 573–605.

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology, 8,* 382–439.

Ross, L. (1977). The intuitive psychologist and his shortcomings. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 10, pp. 173–220). New York: Academic Press.

Ross, L., & Anderson, C. A. (1982). Shortcomings in the attribution process: On the origins and maintenance of erroneous social assessments. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 129–152). New York: Cambridge University Press.

Ross, L., & Lepper, M. R. (1980). The perseverance of beliefs: Empirical and normative considerations. In R. A. Shweder & D. Fiske (Eds.), *New directions for methodology of social and behavioral science: Fallible judgment in behavioral research* (pp. 17–36). San Francisco: Jossey-Bass.

Roth, E. M., & Shoben, E. J. (1983). The effect of context on the structure of categories. *Cognitive Psychology, 15,* 346–378.

Rumelhart, D. E. (1980). Schemata: The building blocks of cognition. In R. J. Spiro, B. C. Bruce, & W. F. Brewer (Eds.), *Theoretical issues in reading comprehension: Perspectives from cognitive psychology, linguistics, artificial intelligence, and education* (pp. 33–58). Hillsdale, NJ: Erlbaum.

Sarris, V. (1976). Effects of stimulus-range and anchor value on psychophysical judgment. In H. G. Geissler & Y. M. Zabrodin (Eds.), *Advances in psychophysics* (pp. 253–268). Berlin, East Germany: UEB Deutscher Verlag der Wissenschaften.

Schank, R. C. (1982). *Dynamic memory: Learning in computers and people.* New York: Cambridge University Press.

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

Schustack, M. W., & Sternberg, R. J. (1981). Evaluation of evidence in causal inference. *Journal of Experimental Psychology: General, 110,* 101–120.

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

Slovic, P. (1985). Violations of dominance in rated attractiveness of playing bets. *Decision Research Report 85-6.* Eugene, OR: Decision Research.

Smith, E. E., & Medin, D. L. (1981). *Categories and concepts.* Cambridge, MA: Harvard University Press.

Snyder, C. R., Shenkel, R. J., & Lowery, C. R. (1977). Acceptance of personality interpretations: The "Barnum effect" and beyond. *Journal of Consulting and Clinical Psychology, 45,* 104–114.

Storms, M. (1973). Videotape and the attribution process: Reversing actors' and observers' points of view. *Journal of Personality and Social Psychology, 27,* 165–175.

Suls, J. M., & Miller, R. L. (Eds.). (1977). *Social comparison processes: Theoretical and empirical perspectives.* Washington, DC: Halsted-Wiley.

Taylor, S. E., & Fiske, S. T. (1978). Salience, attention and attribution: Top of the head phenomena. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 11, pp. 249–288). New York: Academic Press.

Tversky, A., & Kahneman, D. (1971). Belief in the law of small numbers. *Psychological Bulletin, 76,* 105–110.

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology, 5,* 207–232.

Tversky, A., & Kahneman, D. (1982). Causal schemas in judgments under uncertainty. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 117–128). New York: Cambridge University Press.

Tversky, A., & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review, 90,* 293–315.

Ward, L. M. (1979). Stimulus information and sequential dependencies in magnitude estimation and cross-modality matching. *Journal of Experimental Psychology: Human Perception and Performance, 5,* 444–459.

Weiner, B. (1985). "Spontaneous" causal thinking. *Psychological Bulletin, 97,* 74–84.

Wyer, R. S., & Gordon, S. E. (1984). The cognitive representation of social information. In R. S. Wyer & T. K. Srull (Eds.), *Handbook of social cognition* (Vol. 2, pp. 73–150). Hillsdale, NJ: Erlbaum.

Wyer, R. S., & Srull, T. K. (1981). Category accessibility: Some theoretical and empirical issues concerning the processing of social stimulus information. In E. T. Higgins, C. P. Herman, & M. P. Zanna (Eds.), *Social cognition: The Ontario symposium* (Vol. 1, pp. 161–197). Hillsdale, NJ: Erlbaum.

Zadeh, L. A. (1965). Fuzzy sets. *Information and control, 8,* 338–353.

Zadeh, L. A. (1976). A fuzzy-algorithmic approach to the definition of complex or imprecise concepts. *International Journal of Man-Machine Studies, 8,* 249–291.