

# THE INTUITIVE PSYCHOLOGIST AND HIS SHORTCOMINGS: DISTORTIONS IN THE ATTRIBUTION PROCESS<sup>1</sup>

Lee Ross

STANFORD UNIVERSITY  
STANFORD, CALIFORNIA

I. Introduction to Attribution Theory and Attribution Error . . . . .	174
A. Attribution Theory and Intuitive Psychology . . . . .	174
B. Logical Schemata and Nonlogical Biases . . . . .	179
II. Attributional Biases: Instances, Causes, and Consequences . . . . .	184
A. The Fundamental Attribution Error . . . . .	184
B. The "False Consensus" or "Egocentric Attribution" Bias . . . . .	187
C. Inadequate Allowances for the Role-Biased Nature of Social Data . . . . .	193
D. Overlooking the Informational Value of Nonoccurrences . . . . .	196
III. Attributional Biases in the Psychology of Prediction . . . . .	197
A. Three Heuristics Governing Intuitive Prediction and Judgment . . . . .	198
B. Use of Concrete Instances vs. Abstract Baselines . . . . .	200
C. Sources of Nonconservative (Nonregressive) Prediction . . . . .	201
IV. Perseverance of Social Inferences and Social Theories . . . . .	204
A. Perseverance in Self-Perceptions and Social Perceptions . . . . .	204
B. Perseverance Mechanisms . . . . .	206
C. Perseverance of "Theories" . . . . .	208
D. When Judgments and Theories Change: Overcoming Perseverance Mechanisms . . . . .	209
V. Attributional Biases and Social Science Research . . . . .	212
A. The Researcher's Pursuit of Nonobvious Predictions and Results . . . . .	212
B. The Intuitive Psychologist's Illusions and Insights . . . . .	214
References . . . . .	214

<sup>1</sup>The author gratefully acknowledges the assistance of Teresa Amabile, Daryl Bem, Phoebe Ellsworth, Baruch Fischhoff, David Greene, Larry Gross, Mark Lepper, Richard Nisbett, Julia Steinmetz, Amos Tversky, and Philip Zimbardo, all of whom have contributed useful comments and suggestions concerning both the manuscript and the research and conceptual analysis it reports. The preparation of this chapter and reported research were supported by National Institute of Mental Health Research Grant MH 24134.

## I. Introduction to Attribution Theory and Attribution Error

### A. ATTRIBUTION THEORY AND INTUITIVE PSYCHOLOGY

Attribution theory, in its broadest sense, is concerned with the attempts of ordinary people to understand the causes and implications of the events they witness. It deals with the "naive psychology" of the "man in the street" as he interprets his own behaviors and the actions of others. The current ascendancy of attribution theory in social psychology culminates a long struggle to upgrade that discipline's conception of man. No longer the stimulus-response (S-R) automaton of radical behaviorism, promoted beyond the rank of information processor and cognitive consistency seeker, psychological man has at last been awarded a status equal to that of the scientist who investigates him. For man, in the perspective of attribution theory, is an intuitive psychologist who seeks to explain behavior and to draw inferences about actors and their environments.

To better understand the perceptions and actions of this intuitive scientist we must explore his methods. First, like the academic psychologist, he is guided by a number of implicit assumptions about human nature and human behavior, for example, that the pursuit of pleasure and the avoidance of pain are ubiquitous and powerful human motives, or that conformity to the wishes and expectations of one's peers is less exceptional and less demanding of further interpretation than is nonconformity. The amateur psychologist, like the professional one, also relies heavily upon data. Sometimes these data result from first-hand experience; more often, they are the product of informal social communication, mass media, or other indirect sources. Moreover, the representativeness or randomness of the available data is rarely guaranteed by formal sampling procedures. The intuitive psychologist must further adopt or develop techniques for coding, storing, and retrieving such data. Finally, he must resort to methods for summarizing, analyzing, and interpreting his data, that is, rules, formulas, or schemata that permit him to extract meaning and form inferences. The intuitive scientist's ability to master his social environment depends in large measure upon the accuracy and adequacy of his hypotheses, evidence, and methods of analysis and inference. Conversely, sources of oversight, error, or bias in his assumptions and procedures may have serious consequences, both for the lay psychologist himself and for the society that he builds and perpetuates. These shortcomings, explored from the vantage point of contemporary attribution theory, provide the focus of the present chapter.

While the label "attribution theory" and some of the jargon of its proponents may be relatively new and unfamiliar, its broad concerns—naive epistemology and the social inference process—have a long and honorable history in social psychology. The Gestalt tradition, defying the forces of radical behaviorism, has consistently emphasized the *subject's* assignment of meaning to the events that

unfold in the psychological laboratory and in everyday experience (cf. Asch, 1952). Icheiser (1949) explicitly discussed some fundamental social perception biases and their origins almost 30 years ago. Long before attribution theory's current vogue, Kelly (1955, 1958) brought an attributional perspective to the study of psychopathology and, in fact, explicitly suggested the analogy between the tasks of the intuitive observer and those of the behavioral scientist. Schachter and Singer (1962) and Bem (1965, 1967, 1972) further anticipated current attributional approaches in their respective analyses of emotional labeling and self-perception phenomena.

The broad outlines of contemporary attribution theory, however, were first sketched by Heider (1944, 1958) and developed in greater detail by Jones and Davis (1965), Kelley (1967), and their associates (e.g., Jones, Kanouse, Kelley, Nisbett, Valins, & Weiner, 1972; Weiner, 1974). These theorists emphasized two closely related tasks confronting the social observer. The first task is causal judgment: the observer seeks to identify the cause, or set of causes, to which some particular effect (i.e., some action or outcome) may most reasonably be *attributed*. The second task is social inference: the observer of an episode forms inferences about the *attributes* of relevant entities, that is, either the dispositions of actors or the properties of situations to which those actors have responded.

Causal judgment and social inference tasks have both been the subject of intensive theoretical and empirical inquiry and, until recently, had constituted virtually the entire domain of attribution theory. Lately, however, a third task of the intuitive psychologist has begun to receive some attention; that task is the *prediction* of outcomes and behavior. Episodes characteristically lead the intuitive psychologist not only to seek explanations and to make social inferences but also to form expectations and make predictions about the future actions and outcomes. Thus, when a presidential candidate promises to "ease the burden of the average taxpayer," we do attempt to judge whether the promise might have resulted from and reflected the demands of political expediency rather than the candidate's true convictions. However, we are likely also to speculate about and try to anticipate this candidate's and other candidates' future political actions. The psychology of intuitive prediction, is thus a natural extension of attribution theory's domain.

The three attribution tasks are, of course, by no means independent. Explanations for and inferences from an event are obviously and intimately related, and together they form an important basis for speculation about unknown and future events. Each task, moreover, can reveal much about the assumptions, strategies, and failings of the intuitive psychologist. Each, however, provides some unique problems of interpretation and methodology that we should explore before proceeding.

In describing causal judgments, researchers from the time of Heider's early contributions to the present have relied heavily upon a simple internal-external or disposition-situation dichotomy. That is, they have tried to identify those

configurations of possible causes and observed effects that lead the observer to attribute an event to "internal" dispositions of the actor (e.g., abilities, traits, or motives) or to aspects of the "external" situation (e.g., task difficulties, incentives, or peer pressures).<sup>2</sup> While this seemingly simple dichotomy has undeniable intuitive appeal, it creates a host of conceptual problems and methodological pitfalls (see also Kruglanski, 1975). For instance, attribution researchers (e.g., Nisbett, Caputo, Legant, & Maracek, 1973) frequently require subjects to explain why a particular actor has chosen a particular course of behavior. These attributions are then coded as "situational" or "dispositional" on the basis of the *form* of the subject's response. Thus the statement "Jack bought the house because it was so secluded" is coded as an external or situational attribution, whereas "Jill bought the house because she wanted privacy" is coded as an internal or dispositional attribution. The rationale for such coding seems straightforward: The former statement cites something about the object or situation to which the actor responded while the latter statement cites something about the actor. However, when one attends not to the *form* of the attributer's statement but to its *content*, the legitimacy of many such situation-disposition distinctions becomes more dubious: First, it is apparent that causal statements which explicitly cite situational causes implicitly convey something about the actor's dispositions; conversely, statements which cite dispositional causes invariably imply the existence and controlling influence of situational factors. For instance, in accounting for Jack's purchase of a house the "situational" explanation (i.e., "because *it* was so secluded") implies a disposition on the part of this particular actor to favor seclusion. Indeed, the explanation provided is no explanation at all unless one *does* assume that such a disposition controlled Jack's response. Conversely, the dispositional explanation for Jill's purchase (i.e., because *she* likes privacy) clearly implies something about the house (i.e., its capacity to provide such privacy) that, in turn, governed Jill's behavior. Thus the content of both sentences, notwithstanding their differences in form, communicates the information that a particular feature of the house exists and that the purchaser was disposed to respond positively to that feature. In fact, the form of the sentences could have been reversed without altering their content to read "Jack bought the house because he wanted seclusion" and "Jill bought the house because it provided privacy."

Is there a more meaningful basis for a distinction between situational and dispositional causes? One possibility merits consideration. One could ignore the form of subjects' causal statements and, by attending to content, distinguish

<sup>2</sup>Most contemporary researchers have been concerned with attributional rules or principles that apply commonly to all social perceivers. However, a few investigators [most notably Rotter (1966)] have used a similar dichotomy in discussing individual differences in such strategies (see also Collins, 1974; Collins, Martin, Ashmore, & Ross, 1973; Crandall, Katkovsky, & Crandall, 1965; Lefcourt, 1972).

between (1) explanations that do not state or imply any dispositions on the part of the actor beyond those typical of actors in general, and (2) explanations that do state or imply unique relatively atypical or distinguishing personal dispositions. Thus the causal statements "I was initially attracted to Sally because she is so beautiful" and "I was initially attracted to Sally because her astrological sign is Libra" should be coded differently in terms of the proposed distinction despite their similar form. Specifically, while the former explanation conveys that I, *like* most men, am particularly attracted to beautiful women, the latter implies that I, *unlike* most men, am particularly attracted to women of one specific astrological sign. In a sense, the former statement constitutes a situational explanation because it invokes a widely accepted and generally applicable S-R law; the latter explanation, by contrast, is dispositional because it resorts to an individual difference or distinguishing personality variable.<sup>3</sup>

The interpretation of causal statements in the manner just described is obviously a difficult undertaking and many investigators may favor the second attribution task, i.e., the formation of social inferences. This task, at first glance, seems to offer a far less forbidding but no less rewarding research target. For instance, the subject who learns that Joan has donated money to a particular charity may infer that the relevant act reflected [or, in Jones and Davis' (1965) terms, "corresponded" to] some personal disposition of Joan. Alternatively, the subject may infer that Joan's actions reflected not her personal characteristics but the influence of social pressures, incentives, or other environmental factors. The attribution researcher, accordingly, can measure the subject's willingness to assert something about Joan's traits, motives, abilities, beliefs, or other personal dispositions on the basis of the behavioral evidence provided. Specifically, the subject could be required to characterize Joan by checking a Likert-type scale anchored at "very generous" and "not at all generous" with a midpoint of "average in generosity." An alternative version of the scale might deal with the degree of confidence the rater is willing to express in his social inferences.

Such measures of social inference are, indeed, simple to contrive and simple to score. Nevertheless, nontrivial problems of interpretation do arise. Most obvious is the fact that the meaning of a given point on these scales differs for different subjects. More importantly, that meaning may depend upon subtle features of research context and instruction, features often beyond the experimenter's knowledge or control.

<sup>3</sup>The reader should recognize, however, that the form or structure of a causal statement may have a significance that cannot be predicted from a logical analysis of its contents and meaning. Thus, Mary's statement that she loves John because of his qualities rather than her own needs may be an important reflection of her feelings and an important determinant of their subsequent relationship, notwithstanding the dubious logical status of the implied distinction.

Even subtler problems of interpretation may arise. One common format, for instance, asks subjects to indicate whether the specified person is "generous" or "ungenerous," or that they "can't say, depends upon circumstances." Superficially, the first two options indicate willingness to infer the existence or influence of a personal disposition, whereas the third option suggests unwillingness to do so. But a more careful examination of the rater's perceptions may reveal that the third option reflects the rejection only of a *broad* or *general* dispositional label. Thus, further interrogation might reveal that the rater judged the relevant actor to be unexceptional with respect to the behavioral domain in question, that is, like most actors behaving generously or not as situational pressures and constraints dictate. In such a case it seems that no disposition has been inferred (and that the rater has made a situational rather than a dispositional attribution of relevant behavior). On the other hand, the rater's reluctance to choose either trait label may convey his judgment that the actor is relatively more generous than his fellows in some specific circumstances but less generous in others, i.e., that his generosity is inconsistent or idiosyncratic (cf. Bem & Allen, 1974). In the latter case a disposition *has* been inferred, albeit a relatively specific one, for example, a tendency to be unusually generous to one's employees but not to one's family, or vice versa. In fact, several important papers in the attribution area (e.g., Jones & Nisbett, 1971; Nisbett *et al.*, 1973), have failed to distinguish adequately between the absence of trait inferences and the rejection of broad trait labels in favor of narrow or situation-specified ones. Inevitably, confusion and unwarranted conclusions have been the product of this failure.

The third type of attribution task, prediction of behavior (e.g., Nisbett & Borgida, 1975), permits simple unambiguous questions and produces responses that can be scored objectively. Thus the witness to an ostensibly generous act by Joan might be required to predict Joan's behavior in a series of other episodes that seemingly test an actor's generosity or lack of it. Alternatively, the question put to the social observer might be: "What percentage of students (or of people, or of women, or of Joan's socioeconomic peers, etc.) would have behaved as generously as Joan did?" The logical relationship of the prediction task to the tasks of causal judgment and social inference is worth reemphasizing [although the relevant empirical correlations between attribution measures may prove surprisingly weak; cf. Bierbrauer (1973)]. To the extent that a given action or outcome is attributed to the actor rather than his situation and that some stable disposition is inferred, the attributer should prove willing to make confident and distinguishing predictions about the actor's subsequent behaviors or outcomes. Conversely, to the extent that an act is attributed to situational pressures that would dispose all actors to behave similarly, and to the extent that no inferences are made about the actor's dispositions, the observer should eschew such "distinguishing" predictions; instead, he should invoke the "null hypothesis"

and rely upon his baseline information or estimates about how "people in general" respond in the specified situation.

Prediction measures of attribution processes have a crucial advantage (beyond their simplicity and seeming objectivity). Unlike causal judgments or social inferences, predictions can often be evaluated with respect to their *accuracy*. That is, whenever authentic information is available about the behavior of various actors in more than one situation, the success of the intuitive psychologist's attribution strategy can be measured and the direction of biases can be determined. To illustrate this advantage, research on "nonconservative" prediction biases will be discussed later in this chapter (Section III,C; cf. also Amabile, 1975; Ross, Amabile, Jennings, & Steinmetz, 1976a).

## B. LOGICAL SCHEMATA AND NONLOGICAL BIASES

Contemporary attribution theory has pursued two distinct but complementary goals. One goal has been the demonstration that, by and large, social perceivers follow the dictates of logical or rational models in assessing causes, making inferences about actors and situations, and forming expectations and predictions. The other goal has been the illustration and explication of the sources of imperfection, bias, or error that distort these judgments. We shall consider briefly the so-called logical or rational schemata employed by the intuitive psychologist and then devote the remainder of the chapter to the sources of bias in his attempts at understanding, predicting, and controlling the events that unfold around him.

### 1. Two Logical Schemata

Individuals must, for the most part, share a common understanding of the social actions and outcomes that affect them, for without such consensus, social interaction would be chaotic, unpredictable, and beyond the control of the participants. Introspection on the part of attribution theorists, buttressed by some laboratory evidence, has led to the postulation of a set of "rules" that may generally be employed in the interpretation of behaviors and outcomes. These "commonsense" rules or schemata are analogous, in some respects, to the more formal rules and procedures that social scientists and statisticians follow in their analysis and interpretation of data.

H. Kelley, E. E. Jones, and their associates have distinguished two cases in which logical rules or schemata may be applied: In the *multiple* observation case the attributer has access to behavioral data which might be represented as rows or columns of an Actor  $\times$  Object  $\times$  Situation (or Instance) response matrix. Typically, summary statements are provided rather than actual responses. Thus the potential attributer learns that "Most theatergoers like the new Pinter play," or "Mary can't resist stray animals," or "The only television program that Ann

watches is Masterpiece Theater." In the *single* observation case the attributer must deal with the behavior of a single actor on a single occasion. For instance, he may see Sam comply with an experimenter's request to deliver a painful shock to a peer, or he may learn that "Louie bet all his money on a long shot at Pimlico."

The logical rules or principles governing attributions in these two cases are rather different (Kelley, 1967, 1971, 1973). In the multiple observation case the attributer applies the Covariance Principle; that is, he assesses the degree to which observed behaviors or outcomes occur in the presence, but fail to occur in the absence, of each causal candidate under consideration. Accordingly, the attributer concludes that the new Pinter play is a good one (and attributes praise to the play rather than the playgoer) to the extent that it is liked by a wide variety of playgoers, that it is liked by individuals who praise few plays (e.g., "critics"), and that it is applauded as vigorously on the ninetieth day of its run as on the ninth.

In the single observation case the attributer's assessment strategy involves the application of the Discounting Principle, by which the social observer "discounts" the role of any causal candidate in explaining an event to the extent that other plausible causes or determinants can be identified. This attributional principle can be restated in terms of social inferences rather than causal attributions: To the extent that situational or external factors constitute a "sufficient" explanation for an event, that event is attributed to the situation and no inference logically can be made (and, presumably, no inference empirically *is* made) about the dispositions of the actor. Conversely, to the extent that an act or outcome seems to occur *in spite of* and *not because of* attendant situational forces, the relevant event is attributed to the actor and a "correspondent inference" (Jones & Davis, 1965) is made, i.e., the attributer infers the existence and influence of some trait, ability, intention, feeling, or other disposition that could account for the actor's action or outcome. Thus, we resist the conclusion that Louie's longshot plunge at Pimlico was reflective of his stable personal attributes to the extent that such factors as a "hot tip," a desperate financial crisis, or seven prewager martinis could be cited. On the other hand, we judge Louie to be an inveterate longshot player if we learn that his wager occurred in the face of his wife's threat to leave him if he ever loses his paycheck at the track again, his knowledge that he won't be able to pay the rent if he loses, and a track expert's overheard remark that the favorite in the race is "even better than the track odds suggest."

It is worth noting that the application of these two different principles places rather different demands upon the intuitive scientist. The Covariance Principle requires the attributer to apply rules that are essentially logical or statistical in nature and demands no further insight about the characteristics of the entities in question. Application of the Discounting Principle, by contrast,

demands considerable insight about the nature of man and the impact of such situational forces as financial need, alcohol consumption, and a spouse's threat of abandonment. In a sense, the Covariance Principle can be applied by a mere "statistician," whereas the Discounting Principle requires a "psychologist" able to assess the role of various social pressures and situational forces and even to distinguish intended acts and outcomes from unintended ones (cf. Jones & Davis, 1965).

Evidence concerning the systematic use of commonsense attributional principles comes primarily from questionnaire studies in which subjects read and interpret brief anecdotes about the responses of one or more actors to specified objects or "entities" under specified circumstances (e.g., McArthur, 1972, 1976). Occasional studies of narrower scope have also exposed the attributer to seemingly authentic responses, encounters, and outcomes (e.g., Jones, Davis, & Gergen, 1961; Jones & DeCharms, 1957; Jones & Harris, 1967; Strickland, 1958; Thibaut & Riecken, 1955). Such research has demonstrated that attributers can, and generally do, make some use of the hypothesized principles or rules of thumb. That is, manipulations involving information about either the covariance of causes and effects or the number of potential causes for a given effect have produced statistically significant effects upon subjects' judgments. Some studies have even provided evidence about the *relative* impact of various competing attributional principles or criteria (cf. McArthur, 1972, 1976).

What the methodologies employed to date have not assessed (and, logically, could never assess) is the accuracy of the attributer's judgments and the sufficiency of his judgmental strategies. As we have noted earlier, such determinations become possible only when attributers are presented with authentic information and are required to make predictions or other judgments that can be verified.

## 2. *Motivational and Nonmotivational Sources of Bias*

The central concern of the present chapter, and an increasingly important goal of contemporary research and theory, is not the logical schemata which promote understanding, consensus, and effective social control; instead, it is the sources of systematic bias or distortion in judgment that lead the intuitive psychologist to misinterpret events and hence to behave in ways that are personally maladaptive, socially pernicious, and often puzzling to the social scientist who seeks to understand such behavior.

In speculating about possible distortions in an otherwise logical attribution system, theorists were quick to postulate "ego-defensive" biases through which attributers maintained or enhanced their general self-esteem or positive opinion of their specific dispositions and abilities (Heider, 1958; Jones & Davis, 1965; Kelley, 1967). Attempts to prove the existence of such a motivational bias have generally involved demonstrations of asymmetry in the attribution of positive

and negative outcomes—specifically, a tendency for actors to attribute “successes” to their own efforts, abilities, or dispositions while attributing “failure” to luck, task difficulty, or other external factors. Achievement tasks (e.g., Davis & Davis, 1972; Feather, 1969; Fitch, 1970; Wolosin, Sherman, & Till, 1973) and teaching performances (e.g., Beckman, 1970; Freize & Weiner, 1971; Johnson, Feigenbaum, & Weiby, 1964) have provided most of the evidence for this asymmetry. It has also been shown that actors may give themselves more credit for success and less blame for failure than do observers evaluating the same outcomes (Beckman, 1970; Gross, 1966; Polefka, 1965).

Critics, skeptical of broad motivational biases, however, have experienced little difficulty in challenging such research. [See D. T. Miller and Ross (1975) for a detailed discussion.] First, it is obvious that subjects’ private perceptions and interpretations may not correspond to (and may be either less or more “defensive” than) their overt judgments. Second, asymmetries in the attributions of success and failure or differences in the judgments of actors and observers need not reflect motivational influences. As several researchers have noted, success, at least in test situations, is likely to be anticipated and congruent with the actor’s past experience, whereas failure may be unanticipated and unusual. Similarly, successful outcomes are intended and are the object of plans and actions by the actor, whereas failures are unintended events which occur in spite of the actor’s plans and efforts. Observers, furthermore, rarely are fully aware of the past experiences or present expectations and intentions of the actors whose outcomes they witness.

Challenges to the existence of pervasive ego-defensive biases have been empirical as well as conceptual. Thus, in some studies subjects seem to show “counterdefensive” or esteem-attenuating biases. For example, Ross, Bierbrauer, and Polly (1974), using an unusually authentic instructor-learner paradigm, found that instructors rated their own performances and abilities as more important determinants of failure than of success. Conversely, the instructors rated their learner’s efforts and abilities as less critical determinants of failure than success. In the same study these seemingly counterdefensive attributional tendencies proved to be even more pronounced among professional teachers than among inexperienced undergraduates, a result which contradicted the obvious derivation from ego-defensiveness theory that those most directly threatened by the failure experience would be most defensive.

Researchers who insist that self-serving motivational biases exist can, of course, provide alternative interpretations of studies that seem to show no motivational biases or counterdefensive biases. Indeed, in many respects the debate between proponents and skeptics has become reminiscent of earlier and broader debates in learning theory and basic perception in which the fruitlessness of the search for a “decisive” experiment on the issue of motivational influences (i.e., one that could not be interpreted by the “other side”) became

ever more apparent as data multiplied and conceptual analysis sharpened. One approach favored by many researchers has been an attempt to specify relevant moderator variables that might determine when ego defensiveness will distort the attribution process and when it will not do so. An alternate and perhaps more fruitful strategy, however, may be to temporarily abandon motivational constructs and to concentrate upon those informational, perceptual, and cognitive factors that mediate and potentially distort attributional judgments “in general.” A fuller understanding of such factors, in turn, might well allow us, ultimately, to understand and anticipate the particular circumstances in which attributions of responsibility will unduly enhance or attenuate an attributer’s self-esteem (cf. D. T. Miller & Ross, 1975).

Unfortunately the existing attribution literature provides relatively little conceptual analysis or evidence pertaining to nonmotivational biases. The first identified (Heider, 1958) and most frequently cited bias or error, one which we shall term the *fundamental* attribution error, is the tendency for attributers to underestimate the impact of situational factors and to overestimate the role of dispositional factors in controlling behavior. The evidence for this error and its broader implications for our understanding of social psychological phenomena receive detailed consideration in Section II, A.

Our consideration of other previously cited nonmotivational biases shall be brief. Perhaps the most provocative contribution concerning nonmotivational biases has been Jones and Nisbett’s (1971) generalization regarding the “divergent” perceptions of actors and observers (cf. also Jones, 1976). Essentially, it was proposed that actors and observers differ in their susceptibility to the fundamental attribution error; that is, in situations where actors attribute their own behavioral choices to situational forces and constraints, observers are likely to attribute the same choices to the actors’ stable abilities, attitudes, and personality traits. An interesting and unusual feature of the Jones and Nisbett paper is its careful consideration of underlying processes—informational, cognitive, and perceptual in nature—which might *account for* these divergent perceptions of actors and observers (cf. also Jones, 1976). Another interesting line of investigation (one, incidentally, which promises to subsume Jones and Nisbett’s actor-observer generalization) involves “perceptual focusing” (Duncker, 1938; Wallach, 1959). It appears that whatever or whomever we “focus our attention on” becomes more apt to be cited as a causal agent (Arkin & Duval, 1975; Duval & Wicklund, 1972; Regan & Totten, 1975; Storms, 1973; Taylor & Fiske, 1975).

Other attributional biases that have been proposed in the literature have been less systematically investigated. Our list, although incomplete, is perhaps representative. Jones and Davis (1965), for instance, proposed that actions directed towards the attributer, or having consequences for him, are more likely to be attributed to dispositions of the actor than are acts which do not personally involve or affect the attributer. Walster (1966) reported a question-

naire study suggesting that actors are held more responsible (and "chance" or "luck" less responsible) for acts that have serious consequences than for acts with trivial consequences. Finally Kelley (1971), summarizing the results of several prior questionnaire studies, observed that the actor is also held more responsible for acts which lead to reward than for acts which prevent loss or punishment.

## II. Attributional Biases: Instances, Causes, and Consequences

### A. THE FUNDAMENTAL ATTRIBUTION ERROR

Our exploration of the intuitive psychologist's shortcomings must start with his general tendency to overestimate the importance of personal or dispositional factors relative to environmental influences. As a psychologist he seems too often to be a nativist, or proponent of individual differences, and too seldom an S-R behaviorist. He too readily infers broad personal dispositions and expects consistency in behavior or outcomes across widely disparate situations and contexts. He jumps to hasty conclusions upon witnessing the behavior of his peers, overlooking the impact of relevant environmental forces and constraints. Beyond anecdotes and appeals to experience, the evidence most frequently cited for this general bias (e.g., Jones & Nisbett, 1971; Kelley, 1971) involves the attributer's apparent willingness to draw "correspondent" personal inferences about actors who have responded to very obvious situational pressures. For instance, Jones and Harris (1967) found that listeners assumed some correspondence between communicators' pro-Castro remarks and their private opinions even when these listeners *knew* that the communicators were obeying the experimenter's explicit request under "no choice" conditions. A more direct type of evidence that observers may ignore or underestimate situational forces has been provided by Bierbrauer (1973), who studied subjects' impressions of the forces operating in the classic Milgram (1963) situation. In Bierbrauer's study, participants witnessed a faithful verbatim reenactment of one subject's "obedience" to the point of delivering the maximum shock to the supposed victim. Regardless of the type and amount of delay before judging (see Fig. 1), regardless of whether they actually played the role of a subject in the reenactment or merely observed, and regardless of their perceptual or cognitive "set," Bierbrauer's participants showed the fundamental attribution error; that is, they consistently and dramatically underestimated the degree to which subjects in general would yield to those situational forces which compelled obedience in Milgram's situation. In other words, they assumed that the particular subject's obedience reflected his distinguishing personal dispositions rather than the potency of situational pressures and constraints acting upon all subjects. The susceptibility of observers to the fundamental attribution error has been noted

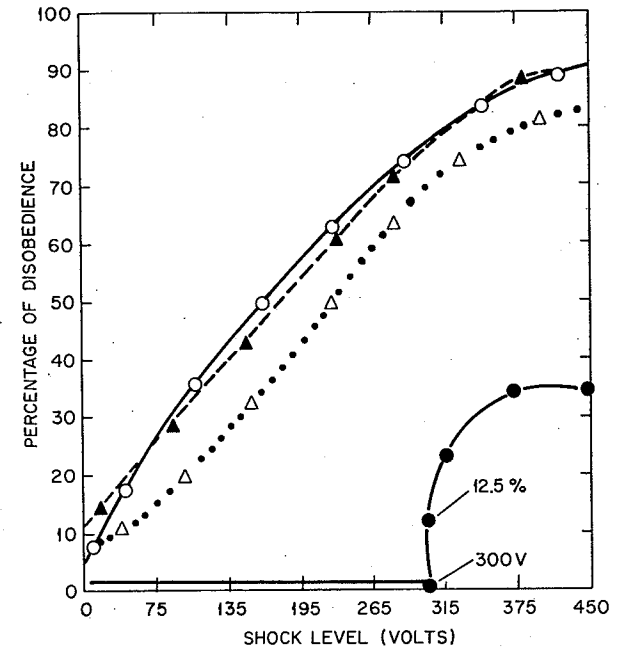


Fig. 1. Comparison of predicted and actual disobedience rates: *open circles*, no delay prediction; *black triangles*, distracted delay prediction; *open triangles*, undistracted delay prediction; *black circles*, rates obtained by Milgram (1963). From Bierbrauer (1973).

by many theorists (e.g., Heider, 1944, 1958; Icheiser, 1949) and disputed by few.<sup>4</sup> The relevance of this error to the phenomena and research strategies of contemporary social psychology, however, has been less widely recognized. To illustrate, we shall first discuss its critical role in mediating the effects of "forced compliance" or "role playing" upon attitude change; then we shall digress slightly to consider the *professional* psychologist's apparent susceptibility to this error.

#### 1. Attribution Error and Forced Compliance Effects

Proponents of cognitive consistency and self-perception theories have regarded demonstrations of attitude change following forced compliance as impor-

<sup>4</sup>Insufficient attention, perhaps, has been given to the possibility that while many situational forces are typically underestimated, there may be others which generally are correctly estimated or even overestimated. Notably, an intriguing series of studies (Deci, 1971; Greene & Lepper, 1974; Lepper & Greene, 1975; Lepper, Greene, & Nisbett, 1973) suggest that actors, in certain circumstances, may inappropriately attribute intrinsically motivated behavior to the influence of salient extrinsic rewards and constraints.

tant evidence for their theoretical viewpoint. Upon closer examination, however, it becomes apparent that these theories "explain" the relevant phenomenon only to the extent that one additionally assumes the operation of the fundamental attribution error. Consider the classic Festinger and Carlsmith (1959) experiment in this regard. Why *does* the compliant actor in the "one dollar" condition experience dissonance? Or, in self-perception terms, why does he incorrectly infer that his compliant behavior reflects liking for the experimental task? Why don't the actors, or the observers in Bem's (1967) interpersonal "simulations," correctly identify the actual "external" causes of the actor's compliance and hence draw no inferences about the actor's attitudes from his counterattitudinal behavior?

The answer is clear: Actors and observers alike must systematically *underestimate* the sufficiency of the particular complex of situational factors in the Festinger and Carlsmith study to produce compliance and must *overestimate* the role played by personal dispositions in producing such behavior. "Correct" attributions, presumably, would produce little dissonance and certainly no erroneous "correspondent inferences" concerning the compliant actor's attitudes. (That is why the "twenty dollar" condition, which presumably facilitates "correct" attributions, produces no attitude change among actors and no tendency for observers to assume congruence between actors private attitudes and their public utterances.) In short, self-perception theory, attribution theory, and dissonance theory "explain" the Festinger and Carlsmith results only if one additionally recognizes the role of the fundamental attribution error.

### 2. Fundamental Attribution Error by Psychologists

The professional psychologist, like the intuitive psychologist, is susceptible to the fundamental attribution error. This susceptibility, in fact, is crucial to the strategy of designing so-called nonobvious research. Many of the best known and most provocative studies in our field depend, for their impact, upon the reader's erroneous expectation that individual differences and personal dispositions will overcome relatively mundane situational variables or "channel factors." Thus Darley and Batson's (1973) study of Good Samaritanism is noteworthy because it seems to contradict our intuition that an individual's ethical principles are more important determinants of bystander intervention than an experimental manipulation of the individual's earliness or lateness for an appointment. Similarly, Bavelas, Hastorf, Gross, and Kite (1965) earn our professional applause and recognition because they show that one can "overcome" those personal qualities which presumably propel the leader into his role through a banal manipulation of the amount of talking done by group members. Most notable of all, perhaps, are the now classic studies showing the vulnerability of actors to situational forces strongly challenging their judgments, preferences, or beliefs (cf. Asch, 1955; Milgram 1963). These studies were professional successes not

because they showed that the relevant target behavior or perceptions could be influenced by a situational manipulation, but because they demonstrated control by a situational factor that the reader had previously assumed to be too weak to exert such control.

In this context it is instructive to consider the heated response of many professionals to Mischel's (1968) summary of evidence indicating that, for most behavior domains of interest to social psychologists, the degree of cross-situational consistency is very modest, making personality scales poor predictors of behavior. Adding insult to injury of the "trait theorists," Mischel and associates (Mischel, 1974; Mischel & Ebbesen, 1970; Mischel, Ebbesen, & Zeiss, 1972) have further demonstrated that in at least one paradigm of general interest—the "delay of gratification paradigm"—relatively subtle situational factors (i.e., the experimenter's suggestion concerning cognitive strategies) overwhelm any individual differences that might be anticipated. Why have Mischel's assertions and demonstrations proven so controversial and prompted such energetic rebuttal research? One reason may be that Mischel's work contradicted not only the formal theories of his critics but also the working assumptions that guide their everyday personal encounters.

The deep conviction that personal dispositions control and are reflected in everyday social behavior will no doubt continue to inspire research in personality theory and personality assessment. Similarly, successful social psychologists will continue to exploit the undue faith of their readers (and their subjects) in the impact of personal beliefs or standards relative to that of situational manipulations. In subsequent sections of this chapter we shall attempt to understand how erroneous trait inferences and trait theories survive or "persevere" in the face of logical challenges and empirical disconfirmations. We shall also explore further the relevance of attributional biases to the tactics and strategy of experimental social psychology.

### B. THE "FALSE CONSENSUS" OR "EGOCENTRIC ATTRIBUTION" BIAS

The professional psychologist relies upon well-defined sampling techniques and statistical procedures for estimating the commonness of particular responses. Where such estimates are relevant to subsequent interpretations and inferences, he can proceed with confidence in his data. Intuitive psychologists, by contrast, are rarely blessed either with adequate "baseline" data or with the means of acquiring such data. To the extent that their systems for interpreting social responses depend upon estimates of commonness or oddity they must, accordingly, rely largely upon subjective impressions and intuitions.

The source of attributional bias that we shall consider next relates directly to the subjective probability estimates of intuitive psychologists. Specifically, we



shall report research demonstrating that laymen tend to perceive a "false consensus," that is, to see their own behavioral choices and judgments as relatively common and appropriate to existing circumstances while viewing alternative responses as uncommon, deviant, and inappropriate. Evidence shall also be reported for an obvious corollary to the false consensus proposition: The intuitive psychologist judges those responses that differ from his own to be more revealing of the actor's stable dispositions than those responses which are similar to his own. Thus, we contend that the person who feeds squirrels, votes Republican, or drinks Drambuie for breakfast will see such behaviors or choices by an actor as relatively common and relatively devoid of information about his personal characteristics. By contrast, another person who ignores hungry squirrels, votes for Democrats, or abstains at breakfast will see the former actor's responses as relatively odd and rich in implications about the actor's personality.

The term *relative* is critical in this formulation of the false consensus bias and requires some clarification. Obviously, the man who would walk a tightrope between two skyscrapers, launch a revolution, or choose a life of clerical celibacy recognizes that his choices would be shared by few of his peers. It is contended, however, that he would see his personal choices as less deviant than would those of us who would *not* walk tightropes, launch revolutions, or become celibate clerics. Similarly, the present thesis concedes that for some response categories virtually all raters' estimates may be biased in the same direction. The incidence of infant abuse, for instance, might be underestimated by abusing and nonabusing parents alike. The relative terms of the false consensus hypothesis leads only to the prediction that abusing parents will estimate child abuse to be more common and less revealing of personal dispositions than will nonabusing parents.

References to "egocentric attribution" (Heider, 1958; Jones & Nisbett, 1971), to "attributive projection" (Holmes, 1968), and to specific findings and phenomena related to false consensus biases have appeared sporadically in the social perception and attribution literatures (cf. Katz & Allport, 1931; Kelley & Stahelski, 1970). Perhaps the most compelling evidence, however, is provided in a series of studies by Ross, Greene, and House (1977b) which we shall review in some detail.

### 1. Empirical Evidence and Implications

In the first study reported, subjects read descriptions of hypothetical conflict situations of the sort they might personally face. One of the four stories, for example, describes the following dilemma:

As you are leaving your neighborhood supermarket a man in a business suit asks whether you like shopping in that store. You reply quite honestly that you do like shopping there and indicate that in addition to being close to your home the

supermarket seems to have very good meats and produce at reasonably low prices. The man then reveals that a videotape crew has filmed your comments and asks you to sign a release allowing them to use the unedited film for a TV commercial the supermarket chain is preparing.

The subjects then were asked to (a) estimate the commonness of the two response alternatives (e.g., signing or refusing to sign the commercial release in the supermarket story); (b) indicate the alternative they, personally, would follow; and (c) assess the traits of the "typical" individual who would follow each of the two specified alternatives.

The relevant estimates and ratings made by subjects strongly supported both the false consensus hypothesis and its corollary. For example, subjects reading the foregoing episode who claimed that they personally would sign the commercial release guessed that more than 75% of their peers would show the same response in the same circumstances; by contrast, subjects who reported that they personally would refuse to sign the release estimated that only 57% would sign. Furthermore, signers made more confident and extreme judgments about the distinguishing traits of the typical nonsigner, while nonsigners reported more confident and extreme impressions about the distinguishing dispositions of the signers.

A second questionnaire study by Ross *et al.* (1977b) dealing with a broad range of habits, preferences, fears, daily activities, expectations, and other personal characteristics greatly extended the apparent domain of the false consensus effect. That is, subjects' estimates of the commonness of the various responses and personal characteristics cited in the questionnaire were consistently biased in accord with their own responses and characteristics.

In a final demonstration by Ross, Greene, and House the hypothetical questionnaire methodology was abandoned and subjects were confronted with a real and consequential conflict situation: Subjects (in the context of a purported experiment on communication techniques) were asked to walk around campus for 30 minutes wearing a large sandwich-board sign bearing the message "EAT AT JOE'S." The experimenter made it clear to subjects that they could easily refuse to participate in the sandwich-board study but that he would prefer that they did participate and thereby "learn something interesting while helping the research project." Subjects were subsequently asked to make their own decision about taking part in the study, to estimate the probable decisions of others, and to make trait inferences about particular peers who agreed or refused to participate.

The results using this "real" conflict situation (Table I) confirmed the findings of earlier questionnaire studies dealing with hypothetical responses. Overall, subjects who agreed to wear the sandwich-board sign estimated that 62% of their peers would make the same choice. Subjects who refused to wear the

TABLE I  
THE FALSE CONSENSUS EFFECT: RATERS' ESTIMATES OF COMMONNESS AND  
TRAIT INFERENCES REGARDING TWO BEHAVIORAL ALTERNATIVES<sup>a</sup>

Raters	Estimated commonness of:		Strength of trait inferences <sup>b</sup> about subject who:	
	Agreement (%)	Refusal (%)	Agrees to wear sign	Refuses to wear sign
Subjects who agree to wear sign ( <i>n</i> = 48)	62	38	120.1	125.3
Subjects who refuse to wear sign ( <i>n</i> = 32)	33	67	139.7	106.8

<sup>a</sup>Summarized from Ross *et al.* (1977b).

<sup>b</sup>Sum of ratings for four traits: higher number indicates more confident and more extreme inferences by rater.

sign estimated that only 33% of their peers would comply with the experimenter's request. Furthermore, as predicted, "compliant" and "noncompliant" subjects disagreed sharply in the relative strength of inferences that they were willing to make about one peer who agreed and one who refused to wear the sandwich board. Compliant subjects made more confident and more extreme inferences about the personal characteristics of the noncompliant peer; non-compliant subjects made stronger inferences about the compliant peer.

Some broad implications of the Ross, Greene, and House demonstrations for our conception of the intuitive psychologist should be clear. His estimates of deviance and normalcy, and the host of social inferences and interpersonal responses that accompany such estimates, are systematically and egocentrically biased in accord with his own behavioral choices. More generally, it is apparent that attributional analyses may be distorted not only by errors in the intuitive psychologist's eventual analysis of social data, but also by earlier biases in sampling or estimating such data.

The present findings are interesting to consider in the light of Jones and Nisbett's (1971) contentions that (1) we see our peers' behavior as the product of broad consistent personal dispositions while attributing our own responses to situational forces and constraints, and (2) we are reluctant to agree that we ourselves possess the type of stable personality traits that we readily apply in characterizing our peers. To explain their results, Jones and Nisbett suggested important differences in the perceptual and informational "perspectives" enjoyed by actors and observers. The Ross *et al.* (1977b) results, however, lead one to speculate that attributional differences of the sort described by Jones and

Nisbett may arise, at least in some measure, simply from attributers' misconceptions about the degree of consensus enjoyed by their own responses and by the alternative responses of their peers.

The derivation is a simple one: To the extent that particular responses by one's peers differ from one's own responses in a given situation, such responses are likely to be seen as relatively odd or deviant—the product, therefore, not of situational forces (which, presumably, guide one's own *contrary* responses) but of distinguishing personality dispositions or traits. Moreover, since any peer responds differently from oneself in at least some situations, it is inevitable that one's peers be seen as the possessors of more numerous and more extreme distinguishing personal characteristics than oneself. The false consensus effect thus allows us to account for many of the phenomena and experimental results that have been mustered in support of Jones and Nisbett's thesis (cf. Jones & Nisbett, 1971; Nisbett *et al.*, 1973) without resorting to the "differing perspective" mechanisms they suggested.

## 2. Sources of the False Consensus Bias

Investigators who have discussed false consensus phenomena or egocentric attributional biases have typically emphasized their motivational status or function for the individual. Such biases, it is contended, both foster and justify the actor's feelings that his own behavioral choices are appropriate and rational responses to the demands of the environment, rather than reflections of his distinguishing personal dispositions. More dynamic interpretations (e.g., Branel, 1962, 1963; Edlow & Kiesler, 1966; Lemann & Solomon, 1952; Smith, 1960) have stressed the ego-defensive or dissonance-reducing function of attributive projection, particularly as a response to failure or negative information about one's personal characteristics.

Several nonmotivational factors, more directly relevant to our present concern with the methods of the intuitive psychologist, may play some role in producing false consensus phenomena. Principal among these are (1) "selective exposure" and "availability" factors, and (2) factors pertaining to the resolution of situational ambiguity.

Selective exposure factors underlying false consensus are fairly straightforward. Obviously, we know and associate with people who share our background, experiences, interests, values, and outlook. Such people *do*, in disproportionate numbers, respond as we would in a wide variety of circumstances. Indeed, our close association is determined, in part, by feelings of general consensus, and we may be inclined to avoid those whom we believe unlikely to share our judgments and responses. This exposure to a biased sample of people and behavior does not demand that we err in our estimates concerning the relevant populations, but it does make such errors likely. More subtle, and more cognitive in character, are the factors which increase our ability to recall,

visualize, or imagine paradigmatic instances of behavior. In a given situation the specific behaviors that we have chosen, or would choose, are likely to be more readily retrievable from memory and more easily imagined than opposite behaviors. In Kahneman and Tversky's (1973) terms, the behavioral choices we favor may be more cognitively "available," and we are apt to be misled by this ease or difficulty of access in estimating the likelihood of relevant behavioral options.

A second nonmotivational source of the false consensus effect arises from the intuitive psychologist's response to ambiguity—both about the nature and magnitude of situational forces and about the meaning and implications of various response alternatives. Attempts to resolve such ambiguity involve interpretation, estimation, and guesswork, all of which can exert a parallel effect on the attributer's own behavior choices and upon his predictions and inferences about the choices of others.

The biasing effect of ambiguity resolution perhaps is most obvious when the attributer's knowledge of a response or situation is secondhand and lacking in important specific details. Consider, for example, the subject who must decide on the precise meaning of such modifiers as *often* or *typically* or of any other potentially ambiguous descriptors encountered in the context of questionnaire items. It is obvious that both the response category to which that subject assigns himself and his categorizations of his peers will be similarly influenced by these decisions about the precise meaning of terms.

Similarly, the subject who reads about a dilemma regarding the signing of a release form for an impromptu television commercial [in Study 1 of Ross *et al.* (1977b)] is forced to imagine the interviewer, the physical setting, and a host of other situational details which might encourage or inhibit the relevant behavioral options. If these imagined details seemingly would encourage one to sign the release, then the subject is more likely to assume that he personally would sign, that a similar decision would be a common response among his peers, and that signing the release would reflect little about the distinguishing dispositions of any particular actor. By contrast, if the details imagined by the subject would inhibit signing of the release, the subject is more likely to assume that he personally would refuse, that his peers typically would do likewise, and that signing of the release would reveal much about the personal dispositions of the relevant actor.

In questionnaire studies this resolution of ambiguities in descriptions of situations and behaviors may seem a troublesome artifact. However, the same factor becomes an important source of bias in everyday social judgments and inferences where attributers may often respond to accounts of situations or actions that are vague and frequently secondhand. The intuitive psychologist constantly is confronted with statements like "Sally hardly ever dates short men" or "John refused to pay the painter's bill when he saw the paint job." In such circumstances he is forced to resolve ambiguities or uncertainties in the

statement. Such resolutions in turn will exert parallel effects upon his assumptions about his own behavior, his impressions about consensus, and his inferences about the dispositions of those whose behavior has been loosely categorized or described.

The false consensus effect demonstrated in Ross *et al.* (1977b), it should be reemphasized, was not restricted to circumstances where raters relied upon ambiguous secondhand descriptions. However, even when attributers actually experience or have fully adequate descriptions of a choice situation, ambiguities remain which inevitably will be resolved differently by different subjects. Thus, subjects who anticipated and feared the ridicule of peers for wearing the "EAT AT JOE'S" sign and regarded the experimenter's wishes and expectations as trivial were likely to refuse to wear the sign, to assume similar refusals by their peers, and to draw strong inferences about the traits of any subject who chose to wear the sign. Opposite priorities, of course, would have produced opposite personal choices and opposite social estimates and inferences.

The false consensus bias, in summary, both reflects and creates distortions in the attribution process. It results from nonrandom sampling and retrieval of evidence and from idiosyncratic interpretation of situational factors and forces. In turn, it biases judgments about deviance and deviates, helps lead actors and observers to divergent perceptions of behavior, and, more generally, promotes variance and error in the interpretation of social phenomena.

#### C. INADEQUATE ALLOWANCES FOR THE ROLE-BIASED NATURE OF SOCIAL DATA

Interpersonal encounters provide an important informational basis for the intuitive psychologist's self-evaluations and social judgments. Often, however, the course of such encounters is shaped and constrained by the formal and informal roles that the various actors must play. More specifically, social roles typically confer unequal control over the style, content, and duration of an encounter; such control, in turn, facilitates displays of knowledge, skill, wit, or sensitivity, while permitting the concealment of deficiencies. Accurate social judgment, accordingly, depends upon the intuitive psychologist's ability to make adequate allowances and adjustments for such role-conferred advantages and disadvantages in self-perception.

In a recent paper, however, Ross, Amabile, and Steinmetz (1977a) have proposed that social perceivers may typically *fail* to make these "necessary" adjustments and, consequently, may draw inaccurate social inferences about role-advantaged and role-disadvantaged actors. In one sense the proposal of Ross *et al.* is simply a special case of the fundamental attribution error described in Section II,A: The fundamental error is a tendency to underestimate the impact of situational determinants and overestimate the degree to which actions and

outcomes reflect the actor's dispositions. The special case proposed by Ross *et al.* deals with the intuitive scientist's underestimation of the effects of roles upon success in self-presentation. In another sense, the proposal of Ross *et al.* contends that the intuitive psychologist is insufficiently sensitive to the biased nature of the data provided by role-constrained encounters and, perhaps, insufficiently sensitive to the problem of sampling bias in general.

The particular roles dealt with in the empirical demonstration reported by Ross *et al.* (1977a) were those of "questioner" and "contestant" in a general knowledge quiz game. The questioner's role obliged the subject to compose a set of challenging general knowledge questions from his or her own store of information, to pose the questions to a contestant, and to provide accurate feedback after each response. Both of these participants (and, in subsequent reenactments, observers as well) were then required to rate the questioner's and the contestant's general knowledge.

The arbitrary assignment and fulfillment of these roles, it should be apparent, forced participants and observers alike to deal with nonrepresentative and highly biased displays of the questioners' knowledge relative to that of contestants. Indeed, the nature of the role-conferred advantages and disadvantages in self-presentation were neither subtle nor disguised. Questioners were allowed and encouraged to display their own wealth of general knowledge by posing difficult and esoteric questions to which their role guaranteed that they would know the answers. The contestant's role, by contrast, prevented any such selective self-serving displays and more displays of ignorance virtually inevitable.

The quiz game contrived by Ross *et al.*, in a sense, provided a particularly stringent test of the intuitive psychologist's proposed insensitivity to role-conferred self-presentation advantages and to corresponding biases in the data samples upon which social judgments frequently are based. For instance, the random nature of the assignment to advantageous and disadvantageous roles was salient and uncontestable. Furthermore, subjects were fully aware of the specific obligations, prerogatives, and limitations associated with each role. In short, subjects seemingly enjoyed an ideal perspective to overcome the proposed source of bias. Nevertheless, the unequal "contest" between questioners and contestants led to consistently biased and erroneous impressions (Table II). As predicted, questioners rated their own general knowledge higher than that of the contestants; moreover, this false impression was shared by the contestants themselves and by uninvolved observers of the encounter.

The narrower as well as the broader implications of the demonstrations of Ross *et al.* should not be overlooked. Indeed, the encounter between advantaged questioners and disadvantaged contestants has obvious parallels within academic settings. Teachers consistently enjoy the prerogative of questioners, and students typically suffer the handicaps of answerers (although some students leap at opportunities to reverse these roles). Consider, as a particularly dramatic in-

TABLE II  
EFFECTS OF QUESTIONER-CONTESTANT ROLE UPON  
SUBJECTS' PERCEPTIONS OF THEIR OWN AND THEIR  
PARTNER'S KNOWLEDGE<sup>a</sup>

Subjects' role	Rating of self <sup>b</sup>	Rating of partner <sup>b</sup>	Self-partner difference
Questioner	54.84	51.66	+3.18
Contestant	40.24	65.16	-24.91

<sup>a</sup>Summarized from Ross *et al.* (1977a).

<sup>b</sup>Higher number indicates belief that the person rated is relatively high in general knowledge (1 = minimum; 100 = maximum).

stance, the role-constrained encounters that characterize the typical dissertation "orals." The candidate is required to field questions from the idiosyncratic and occasionally esoteric areas of each examiner's interest and expertise. In contrast to the examiners, the candidate has relatively little time for reflection and relatively little power to define or limit the domains of inquiry. In light of the present demonstrations, it might be anticipated (correctly so, in this author's experience) that the typical candidate leaves the ordeal feeling more relief than pride, whereas his or her examiners depart with increased respect for each others' insight and scholarship. Such evaluations, of course, may often be warranted. It is worth entertaining the possibility, however, that an alternative procedure for the oral examination, one in which the candidate first posed questions for his examiners and then corrected *their* errors and omissions, would yield more elated candidates and less smug examiners.

There are, of course, countless other contexts in which formal or informal social roles may constrain interpersonal encounters and, in so doing, bias both the data available to the intuitive psychologist and the interpersonal judgments that follow from such data. Thus the employer may dwell upon his personal triumphs, avocations, and areas of knowledge and may avoid mention of his failures, whereas his employee enjoys no such freedom. The physician, likewise, is relatively free to assume with his patient whichever role—stern parent, sympathetic friend, or detached scientist—he wishes. Similarly, the more dominant partner in a personal relationship can disproportionately dictate the rules and arenas for self-presentation and that partner's choice is likely to be self-serving.

If subsequent research confirms the generality of the present thesis, the implications may be all too clear for our understanding of social structures and of the forces that impede social change. Individuals who enjoy positions of power by accident of birth, favorable political treatment, or even their own efforts also tend to enjoy advantages in self-presentation. Such individuals, and

especially their disadvantaged underlings, may greatly underestimate the extent to which the seemingly positive attributes of the powerful simply reflect the advantages of social control. Indeed, this distortion in social judgment threatens to provide a particularly insidious brake upon social mobility whereby the disadvantaged and powerless overestimate the capabilities of the powerful, who in turn inappropriately deem their own caste well suited to the task of leadership.

#### D. OVERLOOKING THE INFORMATIONAL VALUE OF NONOCCURRENCES

The astute Sherlock Holmes directs our attention to a rather subtle but potentially interesting and important shortcoming of the intuitive scientist. In the relevant episode (described in "The Silver Blaze" in *The Memoirs of Sherlock Holmes* by Arthur Conan Doyle) the great detective invites the faithful Dr. Watson to consider "the curious incident of the dog in the night-time." Watson, the conventional behaviorist, remarks correctly that "The dog did nothing in the night-time." Holmes, the inspired behaviorist, triumphantly observes, "That was the curious incident." For the premier practitioner of the science of deduction this *nonincident* or *nonoccurrence* furnishes the key to subsequent interpretations and inferences. (Specifically, Holmes recognizes that a barking dog would have provided no evidence but a silent one proved the intruder in question to be someone well known to the dog.)

The intuitive psychologist, it can more generally be postulated, is like Dr. Watson, a rather conventional behaviorist. He attends to actions or occurrences in forming inferences but neglects to consider the information conveyed when particular responses or events do *not* occur. The author can cite no research directly relevant to the postulated "behaviorist" bias. One source of indirect evidence, however, is provided by findings concerning subjects' use of the observations in a fourfold, presence-absence table. Specifically, Smedslund (1963), Ward and Jenkins (1965), Wason and Johnson-Laird (1972), and others report that only the "present-present" cell strongly influences subjects' inferences regarding covariation (for example, the covariation of diseases and symptoms). Logically, of course, frequency in this cell is no more relevant to the assessment of covariation than are frequencies in any of the other three cells (including the absent-absent cell, which, if our general contention is apt, should prove particularly difficult for subjects to use appropriately in forming impressions of covariation and causal inferences).

Although more directly pertinent research data may be lacking, there are a number of common social experiences which become more explicable in the light of our present speculations. Consider for example the following, rather common episode: Jack meets a new acquaintance, Jill, and after some personal interaction with her, he forms the vague impression that she does not like him.

Such impressions are rarely unfounded. Nevertheless, if Jack searches his memory for specific actions or responses by Jill that reveal her dislike, he will likely be frustrated in his search since, under normal circumstances, acquaintances do not express dislike in overt words or deeds. Thus, if Jack relies upon the sample of evidence he retrieves from memory, he may well conclude that his impression is incorrect and unjustified by the evidence; alternatively, he may cling to his impression but resort to "intuition" or "sixth sense" in order to justify it. If the peerless Mr. Holmes were available for consultation, he doubtless could end Jack's attributional dilemma by focusing his attention on what Jill did *not* do. Jack might well note that Jill did *not* deliberately prolong encounters, did *not* furnish positive nonverbal feedback, and, in general, did *not* show any of the responses that normally signal liking or interest.

In the encounter just described Jack has not been totally oblivious to the information conveyed by Jill's nonresponses—he has correctly discerned her sentiments; he has merely failed to process the information in a manner which facilitates accurate causal inferences or overt verbal expression. In other instances the attributer might even fail to detect or to store the relevant information and might entirely misjudge the sentiments of his acquaintance. The general contention is simply that nonoccurrences are rarely as salient or as cognitively "available" to the potential attributer as are occurrences. As a consequence, recognition, storage, retrieval, and interpretation all become less likely.

The difference between an occurrence and a nonoccurrence can, of course, sometimes be one of semantics. The absence of eye contact can be coded by the potential attributer as the presence of gaze avoidance (e.g., Ellsworth & Ross, 1975). The absence of sexual responsiveness similarly can be coded and interpreted as the presence of frigidity. These seemingly moot semantic distinctions, however, can have nontrivial consequences for the intuitive psychologist. Indeed, if present speculations are warranted, it should be possible to demonstrate that particular absences of response are more noted, more remembered, and more likely deemed as relevant by the attributer when he is provided with positive or active category labels to apply to such absences.

### III. Attributional Biases in the Psychology of Prediction

Implicit expectations and explicit predictions are important products of the intuitive psychologist's collection, coding, storage, retrieval, and interpretation of social data. Often, such expectations and predictions are also crucial mediators of social responses. The attribution theorist, as we have noted earlier, has reason to be concerned with intuitive prediction not only because of its obvious connection to widely studied attributional processes, but also because of its

unique potential for revealing the degree of attributional accuracy and the direction of particular biases. Nevertheless, it has not been attribution researchers but rather two cognitive psychologists, Daniel Kahneman and Amos Tversky, who recently have stirred social psychology's interest in prediction.<sup>5</sup>

#### A. THREE HEURISTICS GOVERNING INTUITIVE PREDICTION AND JUDGMENT

In a very impressive series of papers (Kahneman & Tversky, 1972, 1973; Tversky & Kahneman, 1971, 1973, 1974) these investigators have demonstrated that intuitive predictions and judgments made by typical social observers (and often those made by trained social scientists as well) deviate markedly from the dictates of conventional statistical models. Instead, such predictions seem to reflect the operation of a limited number of "heuristics," or informal decision-making criteria. Among these heuristics are "availability" (Kahneman & Tversky, 1973), "adjustment" (Tversky & Kahneman, 1974), and "representativeness" (Kahneman & Tversky, 1972, 1973). Each heuristic leads the intuitive psychologist to particular errors or biases in subjective estimates and predictions, and each is relevant to the concerns of this chapter.

Use of the *availability* heuristic leads the intuitive psychologist's estimates of the frequency or probability of events to reflect the ease of imagining or remembering those events. Since availability is often poorly correlated with frequency or probability, systematic errors and biases in judgment inevitably result. Thus subjects who heard lists of well-known personalities of both sexes subsequently overestimated the representation of that sex whose members were more famous. The false consensus bias, described earlier, and Chapman and Chapman's (1967, 1969) classic demonstrations of "illusory correlation" in clinical judgment, also seem to reflect the operation of the availability heuristic. Finally, we shall cite the role of availability in our subsequent discussions of impression perseverance and the effects of explanation upon expectation.

Use of the *adjustment* heuristic leads one to make estimates and predictions by "adjusting" either some salient initial value or the result of some partial computation procedure. Such adjustments, however, are rarely sufficient, and the result is typically an "anchoring effect." In one study (Tversky & Kahneman, 1974), for example, subjects were asked to adjust an arbitrary initial estimate of the percentage of African countries in the United Nations. Those starting with anchors of 10% or 65% produced adjusted estimates of 25% and

<sup>5</sup> Unfortunately, the scope of the present chapter forces us to neglect earlier contributions dealing with subjective prediction, expectation, and discrepancies between logical and "psychological" judgment (e.g., Alberoni, 1962; Edwards, 1968; Peterson & Beach, 1967; Slovic & Lichtenstein, 1971; Wheeler & Beach, 1968). Much of this research is acknowledged and described in Kahneman and Tversky's papers and in a recent review by Fischhoff (1976).

45%, respectively. Tversky and Kahneman argue convincingly that overestimation of likelihood for conjunctive events (i.e., the likelihood of *A and B and C* all occurring) and underestimation for disjunctive events (i.e., the likelihood of at least *one* of *A or B or C* occurring) are further results of the intuitive statistician's failure to adequately adjust preliminary or partially computed estimates.

Use of the *representativeness* heuristic is easier to illustrate than to define. It is reflected in the intuitive statistician's tendency to predict that outcome which appears most representative of salient features of the evidence while ignoring conventional statistical criteria such as the reliability, validity, and amount of available evidence, or the prior baseline probabilities associated with the relevant outcomes.

A use of the representativeness heuristic that is particularly striking and pertinent to present concerns is reflected in the intuitive scientist's tendency to give too much weight to *predictor* variables and too little weight to central tendencies in the population distribution of the variable to be *predicted*. Indeed, a sample prediction problem may help the reader to recognize his or her own susceptibility to this bias: *I (the present author) have a friend who is a professor. He likes to write poetry, is rather shy, and is slight of stature. Which of the following is his field: (a) Chinese Studies, or (b) psychology?* The reader who has guessed Chinese Studies, or even seriously entertained the possibility, has fallen victim to the bias described so compellingly by Kahneman and Tversky. Let the unconvinced reader first consider his prediction in light of the number of psychology professors relative to the number of Chinese Studies professors in the overall population. Then let him further consider the more restricted population of the present author's likely friends. Surely *no* psychologist's implicit personality theory about the relationship among avocation, shyness, stature, and academic discipline is sufficiently strong to warrant overlooking such "baseline" considerations.

Errors in parametric prediction problems similarly reflect the use of the representativeness heuristic. Most obvious, perhaps, is the layman's shortcomings in dealing with problems of regression. People expect and predict behaviors and outcomes on variable *Y* to be as "distinctive" or deviant from the norm as the predictor variable *X*, and they are surprised and often disturbed by the phenomenon of "regression to the mean."<sup>6</sup> In fact, they are prone to invent spurious explanations for events that, in reality, are simple regression phenomena. Kahneman and Tversky (1973) describe a relevant anecdote: Israeli flight instructors, urged to make use of positive reinforcement, expressed skepticism. In their experience, they argued, praise of exceptionally good performance typically "led

<sup>6</sup> Nonregressive prediction is, in a sense, a special case of the intuitive statistician's inattentiveness to "baselines" or population distributions for the variable to be predicted. A regressive prediction minimizes error relative to a nonregressive one simply because the former is closer to more observations in the population than is the latter.

to" *diminished* performance on the next trial, while criticism of exceptionally poor performance typically "produced" an immediate *improvement* in performance. On the basis of such firsthand experience, in fact, the instructors concluded that, contrary to accepted psychological doctrine, punishment is more effective than reward.

#### B. USE OF CONCRETE INSTANCES VS. ABSTRACT BASELINES

The relevance of Kahneman and Tversky's work to the general concerns of attribution theory has recently begun to be appreciated (cf. Fischhoff, 1976). Nisbett and Borgida (1975), for example, were quick to note that the weak effects of base rate information on category prediction are analogous to the weak effects of consensus information on attributional judgments (e.g., Cooper, Jones, & Tuller, 1972; McArthur, 1972, 1976; A. G. Miller, Gillen, Schenker, & Radlove, 1973). Pursuing the implications of this observation, Nisbett and Borgida demonstrated that intuitive behavioral predictions, like category predictions, may be relatively impervious to consensus or baseline information. Specifically, subjects were given accurate baseline information about the behavior of previous participants in experiments involving such responses as altruistic intervention and willingness to receive electric shock. As the investigators anticipated, this authentic baseline information did not influence subjects' guesses about the behavior of particular participants in the original experiment. Similarly, this information did not influence subjects' attributions about the causes of such behavior, or their predictions about what their own behavior might be.

Nisbett and Borgida's research design also considered the opposite prediction task, that of estimating overall base rates for behavior on the basis of knowledge provided about the responses of particular individuals. The results were dramatic and consistent with yet another bias described by Tversky and Kahneman (1971). Nisbett and Borgida's subjects' previously demonstrated "unwillingness to deduce the particular from the general was matched only by their *willingness* to infer the general from the particular" (p. 939). Thus, given information that two subjects had behaved in an extreme and counterintuitive fashion (e.g., by taking the maximum possible shock level in a pain threshold experiment), raters predicted that such extreme behavior was *modal* for subjects as a whole.

In attempting to account for these seemingly contradictory but equally nonrational prediction biases, Nisbett, Borgida, Crandall, and Reed (1976) have contrasted the *concreteness* and vividness of specific cases with the pallid *abstract* character of statistical baselines. To illustrate, they invited their readers to participate in the following "thought experiment":

Let us suppose that you wish to buy a new car and have decided that on grounds of economy and longevity you want to purchase one of those solid, stalwart, middleclass Swedish cars—either a Volvo or a Saab. As a prudent and sensible

buyer, you go to *Consumer Reports*, which informs you that the consensus of their experts is that the Volvo is mechanically superior, and the consensus of the readership is that the Volvo has the better repair record. Armed with this information, you decide to go and strike a bargain with the Volvo dealer before the week is out. In the interim, however, you go to a cocktail party where you announce this intention to an acquaintance. He reacts with disbelief and alarm: "A Volvo! You've got to be kidding. My brother-in-law had a Volvo. First, the fancy fuel injection computer thing went out. 250 bucks. Next he started having trouble with the rearend. Had to replace it. Then the transmission and the clutch. Finally sold it in 3 years for junk." [p. 129]

The logical status of this information, Nisbett *et al.* remind the reader, is that the frequency-of-repair record should be shifted by an iota or two on a few dimensions. As they contend, however, the reader's thought experiment is likely to suggest a more dramatic result. The implications of this thought experiment are also borne out in more formal empirical demonstrations. In a series of experiments Borgida and Nisbett (1977) gave undergraduate subjects course-evaluation information and invited them to state their own choices for future enrollment. Some students received summaries of the evaluations of previous course enrollees; others received the information through face-to-face contact with a small number of individuals. As anticipated by the investigators, abstract data-summary information had little impact on course choices, whereas concrete information had a substantial impact.

#### C. SOURCES OF NONCONSERVATIVE (NONREGRESSIVE) PREDICTION

Demonstrations of man's apparent failings as an intuitive statistician promise to capture the attention of attribution theorists. In attempting to clarify the implications of seminal research in this area by Kahneman and Tversky and by others (cf. Fischhoff, 1976), it is important to distinguish between two different sources of bias or error in judgment. Ross *et al.* (1976a) have termed these, respectively, shortcomings in intuitive *psychological theory* and shortcomings in informal *statistical methodology*. The former are misconceptions about the nature of objects and events in the domain of psychological inquiry; the latter are faulty applications of knowledge or information about that domain in making estimates, inferences, and predictions.

Ross, Amabile, and their associates have emphasized this distinction in a series of studies dealing with the intuitive psychologist-statistician's tendency to be "nonconservative" (i.e., nonregressive) in bivariate prediction tasks. In these studies (Amabile, 1975; Ross *et al.*, 1976a) the investigators made use of "authentic" data distributions derived from preliminary studies, student records, and self-report questionnaires. The use of such authentic data, of course, permitted direct assessment of the degree of accuracy and the direction of error in the subjects' predictions and estimates.

Amabile (1975) and Ross *et al.* (1976a) had little difficulty in replicating the basic phenomenon described by Kahneman and Tversky (1972, 1973). Subjects' predictions about the behavior, characteristics, and outcomes of their peers were clearly nonregressive and, by conventional statistical criteria, insufficiently "conservative." For bivariate distributions in which population or large sample correlations were in the range of  $r = 0$  to  $r = +.30$ , subjects made predictions of one variable based on knowledge of the other variable that would have been justified only by correlations in the range of  $r = +.60$  to  $r = +1.00$ .

As anticipated, two distinct sources of error were shown to underlie such nonregressive prediction tendencies. First, it was clear that the intuitive scientists in these studies typically held incorrect assumptions about the strength of the relationship among the observable characteristics and behaviors under consideration. Subjects were required to specify the degree of relationship they believed to exist between particular variables through a variety of "matching tasks" using scatterplots, bivariate charts of numbers, and figures portraying different degrees of covariation between simple physical properties. Using such procedures, subjects consistently overestimated the relevant correlation coefficients and, in so doing, consistently overestimated the degree of cross-situational consistency existing in the relevant behavioral measures and outcomes. This source of bias, accordingly, can be termed *correlation error*. In a sense, it reflects the intuitive psychologist's unwarranted adherence to simple broad "trait theories" of the sort that receive so little support in the systematic investigations of personality theorists (cf. Mischel, 1968, 1969, 1973).

A second source of bias reported by the investigators reflects the intuitive scientist's failure not as a psychologist but as a statistician. This bias, which we may term *regression error*, was reflected in predictions of variable  $Y$  from knowledge of variable  $X$  that were even less regressive than could be justified by the subjects' already-inflated estimates of population correlation. An example from the results reported by Amabile (1975) will help to illustrate the two different sources of nonconservative prediction. In one problem subjects dealt with the relationship between verbal Scholastic Aptitude Test (SAT) scores and subsequent GPA measures for Stanford freshmen. Through examination of a large sample of academic records the investigators estimated the relevant population correlation coefficient (Pearson  $r$ ) to be  $+0.20$ . The subjects' estimate for this relationship, indicated by their choice of appropriately labeled scatterplots, was a correlation coefficient of  $+0.60$  (i.e., correlation error). The subjects' predictions of GPAs from knowledge of SAT scores and vice versa, however, would have been justified by a population correlation not of  $+0.60$  but of  $+0.94$  (i.e., regression error). This pattern of results was replicated using a wide variety of authentic data matrices involving cross-situational consistency in outcomes, personal characteristics, and behaviors.

The data for individual subjects' predictions in the Ross and Amabile studies provide an interesting and more precise view of nonregressive prediction ten-

dencies. Very few individuals rigorously and systematically applied a simple linear "prediction equation." Furthermore, it appears that subjects' departures from consistent linear prediction were governed by the strength of the relationship they believed to exist between  $X$  and  $Y$ . When the relationship was estimated to be strong, their individual predictions were well fit by a simple linear function. When the relationship was believed to be weak, subject's predictions varied widely about a best-fitting regression line. More specifically, the subjects who believed the relationship between  $X$  and  $Y$  to be relatively weak did *not* respond to extreme values of  $X$  with predictions of  $Y$  relatively close to the mean. Instead, they responded by *varying* their predictions (for example, by predicting one more extreme value of  $Y$  and one less extreme value of  $Y$ , given two identically extreme values of the predictor variable  $X$ ).

The data on accuracy were also revealing. Group estimates of  $Y$ , "enlightened" by specific knowledge of  $X$ , consistently yielded a greater mean error (in terms of both absolute and square discrepancies) than would have resulted if the group had averaged across all of those estimates to produce one mean to be offered for all predictions. Similarly the vast majority of individual subjects in the various studies would have decreased the magnitude of their errors by simply repeating their average prediction for  $Y$ , never varying it on the basis of their knowledge of  $X$ . It is difficult to resist the blunt summary that, when it comes to predictions, a little knowledge (i.e., knowledge of a weakly related predictor variable) is a dangerous thing.

Before concluding our discussion of the intuitive psychologist's penchant for nonconservative prediction, a few qualifications, comments, and suggestions concerning future research directions may be in order. First, it is important to recognize that while the term *nonconservative* may be descriptive of the intuitive scientist's judgments in the contexts we have described,<sup>7</sup> it may not accurately describe his intent or his new view of his behavior. Indeed, he may be led to nonregressive prediction through a chain of inferences that seem impeccably conservative.

Consider, for example, a request to predict John's percentile score on a mathematics text in light of information that John scored at the ninetieth percentile in a reading test. The intuitive psychologist-statistician may begin by recognizing that academic abilities tend to be positively correlated and by reasoning that John's math score is likely to be better than average. He may then assume that, having no information about John's math score, he has no basis for predicting whether it will be higher or lower than John's reading score. From the

<sup>7</sup>There is ample evidence, in fact, that in certain judgmental contexts (those involving Bayesian probability) the intuitive psychologist is *overly* "conservative"; that is, he fails to extract sufficient information about population parameters from the data examples available to him (see Peterson & Beach, 1967; Slovic & Lichtenstein, 1971). Such conservatism, perhaps, reflects the operation both of the simple "adjustment" heuristic and of the more general "perseverance mechanisms" to be described later in this chapter.



intuitive scientist's viewpoint it may thus seem conservative to guess that the relevant scores will be equal (a judgment that would be justified, in conventional statistical terms, only by a *perfect* correlation) because such a guess seemingly represents the "middle course" between guessing either that *X* is greater than *Y* or that *Y* is greater than *X*.

A second conceptual issue involves the criterion for optimal prediction. It is entirely possible that subjects may be guided by considerations other than accuracy, or by criteria for accuracy very different from those adopted by the conventional statistician. We certainly have no a priori reason to assume that social observers are particularly concerned with minimizing average *squared* discrepancies, or even with minimizing average absolute discrepancies. Perhaps subjects are concerned with maximizing the number of "exact hits"; perhaps they are willing to increase their "average error" if it will help them to predict the few really deviant or extreme scores in the sample. Some subjects may even be concerned with criteria that are irrelevant to accuracy, for example, making their distribution of predictions reflect the range or variability of the sample scores to be predicted. Any of these nonconventional statistical goals or desiderata may characterize particular subjects or particular data domains (cf. Abelson, 1974). Indeed, one could readily suggest prediction contexts in everyday experience for which each objective would be highly appropriate.

Subsequent research could clarify these issues considerably by pursuing the following research questions: (1) What are subjects' own objectives and what are their criteria for good prediction—in the standard laboratory tasks that have been employed and in a wide range of judgment tasks outside the laboratory? (2) Do subjects recognize the costs of nonregressive prediction strategies, and do they labor under illusions about the possible benefits of such strategies? (3) What kind of feedback in a prediction paradigm, if any, could lead subjects to adopt and to generalize the use of more regressive strategies? Indeed, in what data domains, if any, may subjects already make use of such strategies, and why do they do so?

#### IV. Perseverance of Social Inferences and Social Theories

##### A. PERSEVERANCE IN SELF-PERCEPTIONS AND SOCIAL PERCEPTIONS

In the course of this chapter various biases in the sampling, processing, and interpretation of social data have been described. These biased strategies and procedures produce initial impressions about oneself or other people that typically are premature and often are erroneous. As long as they remain private and

free of behavioral commitment, such first impressions may seem inconsequential, tentative in nature, and free to adjust to new input. A gradually increasing body of theory and research, however, can now be marshaled to suggest the contrary. We shall deal in detail with a pair of "debriefing" experiments reported by Ross, Lepper, and Hubbard (1975). These were designed to provide a simple and dramatic demonstration that errors in initial self-perceptions and social judgments are difficult to reverse and may survive even the complete negation of their original evidential basis (cf. also Walster, Berscheid, Abrahams, & Aronson, 1967; Valins, 1974).

The procedure in the experiments of Ross *et al.* was quite straightforward. Subjects first received continuous false feedback as they performed a novel discrimination task (i.e., distinguishing authentic suicide notes from fictitious ones). In the first experiment reported, this procedure was used to manipulate the subjects' perceptions of their own performance and ability. A second experiment further introduced observers who formed social impressions as they witnessed the false-feedback manipulation. In both experiments, after this manipulation of first impressions had been completed, the experimenter totally discredited the "evidence" upon which the actors' and/or observers' impressions had been based. Specifically, the actor (overheard in Experiment 2 by the observer) received a standard "debriefing" session in which he learned that his putative outcome had been predetermined and that his feedback had been totally unrelated to actual performance. Before dependent variable measures were introduced, in fact, every subject was led to acknowledge explicitly his understanding of the nature and purpose of the experimental deception.

Following this total discrediting of the original source of misinformation, a dependent variable questionnaire was completed dealing with the actors' performances and abilities. The evidence for postdebriefing impression perseverance was unmistakable for actors and observers alike. On virtually every measure (i.e., objective estimates of the actor's just-completed performance, estimates for performance on a future set of discrimination problems, and subjective estimates of the actor's abilities) the totally discredited initial outcome manipulation produced significant "residual" effects upon actors' and observers' assessments (see Table III).

In subsequent related experiments Ross, Lepper, and their colleagues have pursued the perseverance phenomenon using a variety of experimental settings and personal abilities. Although much of this research is still in progress, it is already apparent that the phenomenon is not restricted to the debriefing paradigm or to the suicide note task. For instance, students' erroneous impressions of their "logical problem-solving abilities" (and their academic choices in a follow-up measure 1 month later) persevered even after students learned that good or poor teaching procedures provided a totally sufficient explanation for their success or failure (Lau, Lepper, & Ross, 1976).

TABLE III  
POSTDEBRIEFING PERCEPTIONS OF THE ACTOR'S PERFORMANCE AND ABILITY<sup>a</sup>

Measure	Actor's own perceptions			Observer's perceptions of actors		
	Success	Failure	<i>t</i>	Success	Failure	<i>t</i>
Estimated initial number correct	18.33	12.83	5.91 <sup>e</sup>	19.00	12.42	4.43 <sup>e</sup>
Predicted future number correct	18.33	14.25	4.23 <sup>e</sup>	19.08	14.50	2.68 <sup>c</sup>
Rated ability at task	5.00	3.83	2.65 <sup>c</sup>	5.33	4.00	3.36 <sup>d</sup>
Related abilities at related tasks	4.69	4.53	<1.00	4.69	4.11	1.76 <sup>b</sup>

<sup>a</sup>Summarized from experiment 2 of Ross *et al.* (1975).

<sup>b</sup>*p* < .10.

<sup>c</sup>*p* < .05.

<sup>d</sup>*p* < .01.

<sup>e</sup>*p* < .001.

## B. PERSEVERANCE MECHANISMS

### 1. Distortion and Autonomy

Two related mechanisms have been proposed by Ross and Lepper to account for perseverance phenomena. The first involves *distortion* in the process by which the intuitive psychologist assesses the relevance, reliability, and validity of potentially pertinent data. That is, the weight he assigns to evidence is determined, in large measure, by its consistency with his initial impressions. More specifically, he neglects the possibility that evidence seemingly consistent with his existing impressions may nevertheless be irrelevant or tainted; similarly, he too readily conceives and accepts challenges to contradictory evidence. As a result, data considered subsequent to the formation of a clear impression typically will seem to offer a large measure of support for that impression. Indeed, even a random sample of potentially relevant data "processed" in this manner may serve to strengthen rather than challenge an erroneous impression. The capacity of existing impressions and expectations to bias interpretations of social data is, of course, a well-replicated phenomenon in social psychology (e.g., Asch, 1946; Haire & Grunes, 1950; Hastorf & Cantril, 1954; Jones & Goethals, 1971; Zadny & Gerard, 1974).

The second proposed mechanism involves the *autonomy* achieved by distorted evidence. Once formed, an initial impression may not only be enhanced by the distortion of evidence, it may ultimately be *sustained* by such distortion. The social perceiver, it is contended, rarely reinterprets or reattributes impres-

sion-relevant data when the basis for his original bias in processing that data is discredited. Once coded, the evidence becomes autonomous from the coding scheme, and its impact ceases to depend upon the validity of that scheme. Thus an erroneous impression may survive the discrediting of its original evidential basis because the impression has come to enjoy the support of additional evidence that is seemingly *independent* of that now-discredited basis.

### 2. The Role of Explanation in Impression Perseverance

In accounting for the attributer's reluctance to abandon initial impressions, Ross *et al.* (1975) have emphasized the role of the intuitive psychologist's search for causal explanation. Individuals, they suggest do more than merely aggregate information consistent with their self-perceptions and social perceptions. They also search for antecedents that cause and account for events. These "causal schemata" play a particularly important role in impression perseverance. Once an action, outcome, or personal disposition is viewed as the consequence of known or even postulated antecedents, those antecedents will continue to *imply* the relevant consequence even when all other evidence is removed.

Consider, for example, a subject in the Ross *et al.* (1975) study who has attributed her success in discriminating suicide notes to the insights she gained from the writings of a novelist who committed suicide. Consider, similarly, an observer in that study who has attributed an actor's failure to that actor's manifestly cheerful disposition. Even after debriefing, these attributers retain a plausible basis for inferring the relevant outcome of the discrimination task. Neither participant, of course, has initially considered, or reconsidered after briefing, the many possible antecedents that might have caused and accounted for task outcomes opposite to that contrived by the experimenters.

A series of recent experiments reported by Ross, Lepper, Strack, and Steinmetz (1976c) have provided more direct evidence of the role that causal explanation can play in sustaining discredited impressions and expectations. In these experiments, subjects were presented with authentic clinical case histories. In various experimental conditions they were asked to use this case-study information to explain a significant event in the patient's later life (e.g., suicide, a hit-and-run accident, an attempt to gain elective office, or an altruistic act). In some conditions, subjects wrote their explanations believing that the event had actually occurred, only to learn afterward that the event was hypothetical and that absolutely no authentic information existed concerning the patient's later life. In other conditions, the event to be explained was presented as "merely hypothetical" from the outset. In both experimental conditions, subjects were ultimately asked to estimate the likelihood of the previously explained events and a number of other events as well. (In appropriate control conditions, subjects were given only this final prediction task.)

The results were unambiguous and compelling (Table IV). As hypothesized, the task of identifying case-history antecedents to explain an event increased the

TABLE IV  
RESIDUAL EFFECTS OF "EXPLAINING" AN EVENT ON JUDGED LIKELIHOOD OF THAT EVENT<sup>a</sup>

Event previously explained by subject <sup>b</sup>	Estimated likelihood <sup>c</sup> that patient will:		
	[A] Become involved in hit-and-run accident	[B] Seek election to City Council	Difference [A] - [B]
[A] Patient becomes involved in hit-and-run accident	+1.45	-2.47	+3.92
[B] Patient seeks election to City Council	-0.25	+0.66	-0.91
[C] None	-0.08	-1.55	+1.47

<sup>a</sup>Summarized from Ross *et al.* (1976c).

<sup>b</sup>Data are combined for subjects who explained event initially believing it to be real and for subjects who explained event knowing it to be hypothetical.

<sup>c</sup>More positive number indicates greater belief that the specified event is likely to have actually occurred in patient's life.

subjects' estimates of that event's likelihood. The relevant phenomenon was replicated across a variety of cases and predicted events and was demonstrated under both the "hypothetical" and "nonhypothetical" explanation conditions [see also Fischhoff (1975, 1976) and Fischhoff and Beyth (1975) for a discussion of the "certainty of hindsight knowledge," a phenomenon that may be closely related to the present demonstrations and may depend upon similar mechanisms].

### C. PERSEVERANCE OF "THEORIES"

It should be apparent that the same biased attributional processes which sustain discredited individual inferences may also sustain the discredited attributional strategies that give rise to such inferences. Consider, for instance, the nonregressive or nonconservative prediction strategies discussed earlier. Why does the intuitive scientist continue to believe that correlations reflecting cross-situational consistency are strong when the evidence of his everyday experience will suggest that such correlations are weak? Why does he continue to make nonregressive predictions in a world that, presumably, better rewards more conservative strategies? The answers should be apparent from our foregoing discussion of impression perseverance. First, the intuitive observer selectively codes those data potentially relevant to the relationship between *X* and *Y*. Data points that fit his hypotheses and predictions are accepted as reliable, valid, representative, and free of error or "third-variable influences." Such data points are seen as reflective of the "real" or "paradigmatic" relationship between *X* and

*Y*. By contrast, data points that deviate markedly from the intuitive psychologist's expectations or theory are unlikely to be given great weight and tend to be dismissed as unreliable, erroneous, unrepresentative, or the product of contaminating third-variable influences.

Thus the intuitive scientist who believes that fat men are jolly, or more specifically that fatness causes jolliness, will see particular fat and jolly men as strong evidence for this theory; he will not entertain the hypothesis that an individual's jolliness is mere pretense or the product of a particularly happy home life rather than obesity. By contrast, fat and morose individuals will be examined very carefully before gaining admission to that scientist's store of relevant data. He might, for instance, seek to determine whether the individual's moroseness on the day in question is atypical, or the result of a nagging cold or a disappointing day, rather than the reflection of some stable attribute. It need hardly be emphasized that even a *randomly* generated scatterplot or contingency table can yield a relatively high correlation if coded in the manner just outlined (cf. Chapman & Chapman, 1967, 1969). Indeed, the professional psychologist, like the intuitive one, can readily derive unwarranted support for almost any hypothesis if permitted to delete, post hoc, the data points that offend his thesis. Perseverant beliefs in extrasensory perception in the face of disconfirming experimental evidence may reflect such selective processing of data (see Gardner, 1975).

The autonomy enjoyed by distorted inferences may further contribute to the perseverance of nonoptimal theories and attributional strategies. The intuitive scientist detects more support for his general theory than is warranted and, having thus "coded" or summarized his findings, he is then disposed to maintain his theory in the face of subsequent logical or empirical attacks by "citing" the wealth of seemingly independent empirical support that it enjoys. It is through such means, perhaps, that the intuitive psychologist remains committed to concepts of broad, stable, heuristically valuable, personality traits and perseveres in the use of nonoptimal prediction strategies. Superstitious learning phenomena and the "partial reinforcement effect" similarly may reflect the subject's capacity to selectively attribute instances of reinforcement and nonreinforcement.

### D. WHEN JUDGMENTS AND THEORIES CHANGE: OVERCOMING PERSEVERANCE MECHANISMS

An obvious question begins to emerge from our demonstrations of impression perseverance: Under what circumstances *do* erroneous or unwarranted personal judgments change? Clearly, none of us has exactly the same view of ourselves or of our fellows as he once did; personal experiences do have an impact upon such views. Therapy, education, persuasive arguments, and mass media campaigns also can alter our self-perceptions and social attitudes. Indeed, as we pointed out early in this chapter, psychology's broad view of man has

changed and evolved in response to arguments and evidence presented by the field's vanguard. While a detailed examination of the requisites for such change is beyond the scope of the present discussion, a few observations may be appropriate.

First, it seems clear that neither challenges to specific bits of evidence confirming a belief (or theory) nor the addition of small amounts of contradictory evidence are likely to prove effective in producing overall change. Challenges and additions to data tend themselves to be "selectively" coded in accord with one's biased prior impressions. Such selective coding, in fact, generally is quite rational and reasonable: most of our personal impressions and beliefs are well founded, and confirming evidence usually is more valid, representative, and relevant than disconfirming evidence. But this rational selectivity in interpreting individual "bits" of evidence leads to an irrational result when a whole "batch" of evidence is considered; specifically, virtually any random sample of newly considered evidence processed in this manner will seemingly support the existing belief or theory. In the face of subsequent logical or informational challenges, furthermore, the random sample of newly "processed" data may even help to sustain the incorrect theory which dictated the processing bias. Consider, for instance, an apparently close friendship. Individual acts by either friend will be taken at face value by the other if they seem to reflect sensitivity, concern, affection, or interest. Conversely, particular acts that might seem insensitive, cold, or hostile in the eyes of some disinterested observer will *not* be taken at face value by the other—at least not without the presence of a good deal of corroborating evidence and the absence of potential alternative interpretations. Such biased attributional coding, it should be reemphasized, is not irrational. Our past experiences, and our global views of relationships, generally *do* promote more accurate attributions of specific acts and outcomes. By "rationally" giving our friends the "benefit of the doubt" in our inferences about individual acts, however, we risk irrational interpretations of larger samples of evidence. Specifically, we may be prone to overlook systematic evidence of indifference or resentment until it is overwhelming or until our peers explicitly interpret their behavior for us.

Erroneous impressions, theories, and data-processing strategies, therefore, may not be changed through mere exposure to samples of new evidence. It is not contended, of course, that new evidence can never produce change—only that new evidence will produce *less* change than would be demanded by any logical or rational information-processing model. Thus, new evidence that is strongly and consistently contrary to one's impressions or theories can, and frequently does, produce change, albeit at a slower rate than would result from an unbiased or dispassionate view of the evidence.

It seems clear that the effects of attributional distortion and autonomy can also be overcome without the brute force of consistently disconfirming data.

Dramatic religious and political conversions, for example, presumably are accomplished by other means. Specifically, these conversions seem to be the product not of new data nor attacks on old beliefs but, rather, involve assaults on whole belief systems. Typically, the target of the conversion attempt is not induced to reevaluate the evidence "objectively" or dispassionately. Instead, he is taught to make use of a new and encompassing attributional bias. Often he is also urged to reject all past beliefs and insights as the product of pernicious social, philosophical, or political forces. The attempt to induce a tabula rasa state in the individual and to provide a selective interpretation schema for both the consideration of new evidence and the reconsideration of old evidence is characteristic of strategies for ideological conversion. Insight therapies similarly attempt to overcome impression perseverance through global assaults on belief systems and through the introduction of new explanatory or inferential schemata (although the "working through" of isolated incidents and experiences responsible for perseverant feelings and perceptions is also an important aspect of many therapy regimens).

In a more limited vein, it is worth briefly considering one additional result reported by Ross *et al.* (1975). In one relevant experiment, two different types of debriefing conditions were employed. In the standard "outcome debriefing" condition, subjects were made aware that the prior success-failure manipulation

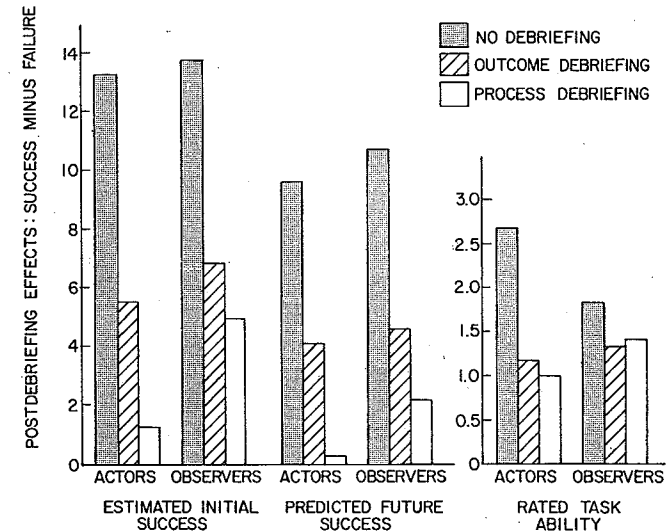


Fig. 2. Postdebriefing differences between success and failure conditions for actors and observers. Summarized from Experiment 2 of Ross *et al.* (1975). Copyright 1975 by The American Psychological Association. Reprinted by permission.

had been a total deception and that their "outcome" had been assigned without regard to their real performance and ability. In special "process debriefing" conditions, subjects also received an explicit discussion of the perseverance phenomenon and of the distortion and autonomy mechanisms which might lead them, personally, to retain inappropriate or inaccurate self-perceptions. Figure 2 presents the relevant data. While the regular debriefing procedures could not eliminate postdebriefing effects of the original outcome manipulation, the special process debriefing conditions were almost totally effective for actors (although less effective for observers). This demonstration, at the very least, suggests process debriefing is an important requirement for the ethical conduct of deception research. It also hints that personal insight concerning one's attributional biases may hasten the achievement of therapeutic goals.

## V. Attributional Biases and Social Science Research

### A. THE RESEARCHER'S PURSUIT OF NONOBVIOUS PREDICTIONS AND RESULTS

Among social psychologists today there is an epidemic of critical self-evaluation and debate about the current status of the field's theory, research strategy, practical contributions, and prospects for the future. In such soul-searching appraisals, the issue of "nonobviousness" has figured prominently (cf. McGuire, 1973). Researchers feel compelled to defend the nonobviousness or counter-intuitiveness of their findings lest they be ridiculed as practitioners of "bubba psychology." Furthermore, results deemed subtle and surprising by the investigator may too often seem obvious to one reader (cf. Fischhoff & Slovic, 1976), gratuitous to a second reader, and simply wrong or misconstrued to a third. The issue of nonobviousness, and its relationship to research strategy, can be reconsidered in the light of our present concern with the shortcomings of the intuitive and the professional psychologist.

It is important, first, to distinguish nonobvious empirical results from non-obvious functional relationships. Rarely does the investigator, in his pursuit of nonobviousness, postulate a direct relationship between two variables that were previously considered to be unrelated or inversely related. The nonobvious nature of most predictions in social psychology, instead, relates to the specific context in which the relationship between variables is tested. By carefully contriving the balance of forces operating in a particular setting, the investigator attempts to prove a deficiency both in the intuitive commonsense psychology of the layman and in the existing conceptual analyses of the investigator's professional peers. The demonstration will be both successful and nonobvious to the extent that those factors identified or manipulated by the investigator are more

potent determinants of behavior than are other factors which would justify the "intuitive," but wrong, hypothesis. In other words, the researcher proceeds from the assumption that his professional audience is prone to some attributional bias that leads it either to ignore the relevance or to underestimate the relative magnitude of a particular set of behavioral determinants. Typically, the investigator also introduces and makes salient in his description of the research setting other potential determinants which he believes will have less impact than his professional peers expect. If the attribution biases illustrated are genuine, and the relevant determinants and experimental context are of broad significance, the research is likely to become highly visible and controversial. If the investigator has incorrectly characterized the relative weightings assigned to behavioral determinants by his peers, or has dealt with very restricted failings in such weightings, then the demonstration will be dismissed as "obvious" or "too limited in its applicability and interest." Needless to say, the research strategy of capitalizing on nonobvious predictions may depend no less on the investigator's abilities as a stage manager and mystery writer than on his ability to recognize inadequacies either in contemporary psychological theory or in the informal attribution theories of his colleagues.

Earlier in this chapter we discussed the special relevance of the fundamental attribution error that leads the intuitive psychologist to underestimate the magnitude of situational factors relative to dispositional factors or individual differences. Now the argument can be generalized to suggest that the most "classic" experiments in social psychology are nonobvious in terms of the *relative magnitude*, rather than the existence, of the situational forces they manipulate or identify. Consider Milgram's (1963) provocative demonstrations of "obedience" or Asch's (1955) studies of conformity. The importance of such research clearly does not lie in the demonstration that subjects obey authority figures or that conformity pressures excited by peers have an impact. That *would* be dull, obvious, and scientifically unproductive. Rather, they demonstrate that these variables are important *relative* to other personal and situational influences that most of us had previously thought to be far more important determinants of behavior (see Ross, Bierbrauer, & Hoffman, 1976b).

In this connection it is worth noting that these celebrated demonstrations share a very unusual characteristic—they employ no control group. Typically, of course, a control group is necessary to establish some kind of baseline from which the experimental group deviates significantly. Asch and Milgram, however, were able to let our intuitions or expectations serve as the baseline condition from which deviations could be noted. Clearly, in their classic studies, the degree of deviation from expectations was sufficiently compelling *without* statistically contrasting the experimental conditions with some control condition. It is not surprising, furthermore, that each of the classic demonstrations cited has led critics to contend that the surprising effects demonstrated were the products of factors other than those proposed by the original investigators. Indeed, such

truly nonobvious demonstrations become a continuing source of inspiration and a challenge for successive generations of scientists.

#### B. THE INTUITIVE PSYCHOLOGIST'S ILLUSIONS AND INSIGHTS

The same attributional biases that provide the basis for nonobvious research demonstrations by misleading the professional scientist may also mislead intuitive psychologists who serve as their subjects. Nisbett and Wilson (1977) recently have described several dramatic instances of the experimental subject's inability to identify accurately the nature and magnitude of the situational features or manipulations which influence his behavior. It is noteworthy that the subject in a nonobvious demonstration or experiment is typically surprised and dismayed to learn that his behavior is so susceptible to the set of situational factors under the experimenter's control. Conforming subjects in the Asch paradigm and obedient subjects in the Milgram paradigm reportedly were shocked, embarrassed, and prone to make inappropriate inferences about themselves, both during the demonstration and afterwards when the experimenter revealed his intent. It is not surprising, moreover, that they were not easily consoled by debriefing procedures, for the experimenter could not restore what the subjects had lost—a satisfying, albeit inaccurate, implicit theory about the relative impact of specific personal and situational determinants of their own behavior.

The intuitive psychologist, shielded by perseverance mechanisms, is no less likely outside the psychological laboratory to remain ignorant of the distortions and inadequacies both in his primary assumptions and in his methods of sampling, coding, and analyzing the data of everyday experience. Sometimes the results of such ignorance are benign or even benevolent, e.g., the social observer attributes his friendships to the particular personal qualities of his friends and overlooks the role of social ecology (Festinger, Schachter, & Back, 1950). At other times, the results may be harmful to the individual or the society, as unjust and maladaptive methods of resource allocation and social control are justified and perpetuated.

The intuitive psychologist, however, cannot be totally insulated from clashes between expectations and observations, between intuitions and evidence. From such clashes he may be led to cynicism, self-doubt, or disappointment. Alternatively, he may be led to new psychological insights and a willingness to reshape his own life and the institutions of his society.

#### REFERENCES

Abelson, R. P. Social psychology's rational man. In G. W. Mortimore & S. I. Benn (Eds.), *The concept of rationality in the social sciences*. London: Routledge & Kegan Paul, 1974.

- Alberoni, F. Contributions to the study of subjective probability. Part I. *Journal of General Psychology*, 1962, 66, 241-264.
- Amabile, T. M. Investigations in the psychology of prediction. Unpublished manuscript, Stanford University, 1975.
- Arkin, R. M., & Duval, S. Focus of attention and causal attributions of actors and observers. *Journal of Experimental Social Psychology*, 1975, 11, 427-438.
- Asch, S. Forming impressions of personality. *Journal of Abnormal and Social Psychology*, 1946, 41, 258-290.
- Asch, S. *Social Psychology*. Englewood Cliffs, N.J.: Prentice-Hall, 1952.
- Asch, S. Opinions and social pressures. *Scientific American*, 1955, 193, 31-35.
- Bavelas, A., Hastorf, A. H., Gross, A. E., & Kite, W. R. Experiments in the alteration of group structure. *Journal of Experimental Social Psychology*, 1965, 1, 55-70.
- Beckman, L. Effects of students' performance on teachers' and observers' attributions of causality. *Journal of Educational Psychology*, 1970, 61, 75-82.
- Bem, D. J. An experimental analysis of self-persuasion. *Journal of Experimental Social Psychology*, 1965, 1, 199-218.
- Bem, D. J. Self-perception: An alternative interpretation of cognitive dissonance phenomena. *Psychological Review*, 1967, 74, 183-200.
- Bem, D. J. Self-perception theory. In L. Berkowitz (Ed.), *Advances in experimental social psychology*. Vol. 6. New York: Academic Press, 1972.
- Bem, D. J., & Allen, A. On predicting some of the people some of the time: The search for cross-situational consistencies in behavior. *Psychological Review*, 1974, 81, 506-520.
- Bierbrauer, G. Effect of set, perspective, and temporal factors in attribution. Unpublished doctoral dissertation, Stanford University, 1973.
- Borgida, E., & Nisbett, R. E. The differential impact of abstract vs. concrete information on decisions. *Journal of Applied Social Psychology*, 1977, in press.
- Bramel, D. A dissonance theory approach to defensive projection. *Journal of Abnormal and Social Psychology*, 1962, 64, 121-129.
- Bramel, D. Selection of a target for defensive projection. *Journal of Abnormal and Social Psychology*, 1963, 66, 318-324.
- Chapman, L., & Chapman, J. The genesis of popular but erroneous psychodiagnostic observations. *Journal of Abnormal Psychology*, 1967, 72, 193-204.
- Chapman, L., & Chapman, J. Illusory correlations as an obstacle to the use of valid psychodiagnostic signs. *Journal of Abnormal Psychology*, 1969, 74, 271-280.
- Collins, B. E. Four components of the Rotter internal-external scale: Belief in a difficult world, a just world, a predictable world, and a politically responsive world. *Journal of Personality and Social Psychology*, 1974, 29, 381-391.
- Collins, B. E., Martin, J. C., Ashmore, R. D., & Ross, L. Some dimensions of the external-internal metaphor in theories of personality. *Journal of Personality*, 1973, 41, 471-492.
- Cooper, J., Jones, E. E., & Tuller, S. M. Attribution, dissonance and the illusion of uniqueness. *Journal of Experimental Social Psychology*, 1972, 8, 45-47.
- Crandall, V. C., Katkovsky, W., & Crandall, V. G. Children's beliefs in their own control of reinforcements in intellectual-academic achievement situations. *Child Development*, 1965, 36, 91-109.
- Darley, J., & Batson, C. D. "From Jerusalem to Jericho": A study of situational and dispositional variables in helping behavior. *Journal of Personality and Social Psychology*, 1973, 27, 100-119.
- Davis, W. L., & Davis, D. E. Internal-external control and attribution of responsibility for success and failure. *Journal of Personality*, 1972, 40, 123-136.

- Deci, E. Effects of externally mediated rewards on intrinsic motivation. *Journal of Personality and Social Psychology*, 1971, 18, 105-115.
- Duncker, K. Induced motion. In W. Ellis (Ed.), *A sourcebook of Gestalt psychology*. New York: Harcourt, 1938. Pp. 161-172.
- Duval, S., & Wicklund, R. A. *A theory of objective self-awareness*. New York: Academic Press, 1972.
- Edlow, O., & Kiesler, C. Ease of denial and defensive projection. *Journal of Experimental Social Psychology*, 1966, 2, 183-191.
- Edwards, W. Conservatism in human information processing. In B. Kleinmuntz (Ed.), *Formal representation of human judgment*. New York: Wiley, 1968.
- Ellsworth, P., & Ross, L. Intimacy in response to direct gaze. *Journal of Experimental Social Psychology*, 1975, 11, 592-613.
- Feather, N. T. Attribution of responsibility and valence of success and failure in relation to initial confidence and task performance. *Journal of Personality and Social Psychology*, 1969, 13, 129-144.
- Festinger, L., & Carlsmith, J. M. Cognitive consequences of forced compliance. *Journal of Abnormal and Social Psychology*, 1959, 58, 203-210.
- Festinger, L., Schachter, S., & Back, K. *Social pressures in informal groups: A study of human factors in housing*. New York: Harper, 1950.
- Fischhoff, B. Hindsight ≠ foresight: The effect of outcome knowledge on judgment under uncertainty. *Journal of Experimental Psychology: Human Perception and Performance*, 1975, 1, 288-299.
- Fischhoff, B. Attribution theory and judgment under uncertainty. In J. Harvey, W. Ickes, & R. Kidd (Eds.), *New directions in attribution research*. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1976. Pp. 419-450.
- Fischhoff, B., & Beyth, R. "I knew it would happen"—remembered probabilities of once-future things. *Organizational Behavior and Human Performance*, 1975, 13, 1-16.
- Fischhoff, B., & Slovic, P. On the psychology of experimental surprises: Outcome knowledge and the journal review process. *Oregon Research Institute Research Bulletin*, 1976, 16, No. 2.
- ✓ Fitch, G. Effects of self-esteem, perceived performance, and chance on causal attributions. *Journal of Personality and Social Psychology*, 1970, 16, 311-315.
- ✓ Freize, I., & Weiner, B. Cue utilization and attributional judgments for success and failure. *Journal of Personality*, 1971, 39, 591-606.
- Gardner, M. Concerning an effort to demonstrate extrasensory perception by machine. *Scientific American*, 1975, 233, 114-118.
- Greene, D., & Lepper, M. Intrinsic motivation: How to turn play into work. *Psychology Today*, 1974, September, 49-54.
- Gross, A. Evaluation of the target person in a social influence situation. Unpublished doctoral dissertation, Stanford University, 1966.
- Haire, M., & Grunes, W. F. Perceptual defenses: Processes protecting an organized perception of another personality. *Human Relations*, 1950, 3, 403-412.
- Hastorf, A. H., & Cantril, H. They saw a game: A case study. *Journal of Abnormal and Social Psychology*, 1954, 49, 129-134.
- Heider, F. Social perception and phenomenal causality. *Psychological Review*, 1944, 51, 358-373.
- Heider, F. *The psychology of interpersonal relations*. New York: Wiley, 1958.
- Holmes, D. S. Dimensions of projection. *Psychological Bulletin*, 1968, 69, 248-268.
- Icheiser, G. Misunderstandings in human relations: A study in false social perception. *American Journal of Sociology*, 1949, 55, Part 2, 1-70.

- Johnson, T. J., Feigenbaum R., & Weiby, M. Some determinants and consequences of the teacher's perception of causation. *Journal of Experimental Psychology*, 1964, 55, 237-246.
- Jones, E. E. How do people perceive the causes of behavior? *American Scientist*, 1976, 64, 300-305.
- Jones, E. E., & Davis, K. E. From acts to dispositions: The attribution process in person perceptions. In L. Berkowitz (Ed.), *Advances in experimental social psychology*. Vol. 2. New York: Academic Press, 1965.
- Jones, E. E., Davis, K. E., & Gergen, K. J. Role playing variations and their informational value for person perception. *Journal of Abnormal and Social Psychology*, 1961, 63, 302-310.
- Jones, E. E., & DeCharms, R. Changes in social perception as a function of the personal relevance of behavior. *Sociometry*, 1957, 20, 75-85.
- Jones, E. E., & Goethals, G. R. Order effects in impression formation: Attribution context and the nature of the entity. In E. E. Jones et al. (Eds.), *Attribution: Perceiving the causes of behavior*. Morristown, N.J.: General Learning Press, 1971.
- Jones, E. E., & Harris, V. A. The attribution of attitudes. *Journal of Experimental Social Psychology*, 1967, 3, 1-24.
- Jones, E. E., Kanouse, D. E., Kelley, H. H., Nisbett, R. E., Valins, S., & Weiner, B. *Attribution: Perceiving the causes of behavior*. Morristown, N.J.: General Learning Press, 1971.
- Jones, E. E., & Nisbett, R. E. The actor and the observer: Divergent perceptions of the causes of behavior. In E. E. Jones et al. (Eds.), *Attribution: Perceiving the causes of behavior*. Morristown, N.J.: General Learning Press, 1971.
- Kahneman, D., & Tversky, A. Subjective probability: A judgment of representativeness. *Cognitive Psychology*, 1972, 3, 430-454.
- Kahneman, D., & Tversky, A. On the psychology of prediction. *Psychological Review*, 1973, 80, 237-251.
- Katz, D., & Allport, F. *Student's attitudes*. Syracuse: Craftsman Press, 1931.
- Kelley, H. H. Attribution theory in social psychology. In D. Levine (Ed.), *Nebraska symposium on motivation*. Vol. 15. Lincoln: University of Nebraska Press, 1967.
- Kelley, H. H. Attribution in social interaction. In E. E. Jones et al. (Eds.), *Attribution: Perceiving the causes of behavior*. Morristown, N.J.: General Learning Press, 1971.
- Kelley, H. H. The process of causal attribution. *American Psychologist*, 1973, 28, 107-128.
- Kelley, H. H., & Stahelski, A. The social interaction basis of cooperators' and competitors' beliefs about others. *Journal of Personality and Social Psychology*, 1970, 16, 66-91.
- Kelly, G. *The psychology of personal constructs*. New York: Norton, 1955. 2 vols.
- Kelly, G. Man's construction of his alternatives. In G. Lindzey (Ed.), *Assessment of human motives*. New York: Holt, 1958.
- Kruglanski, A. The endogenous-exogenous partition in attribution theory. *Psychological Review*, 1975, 82, 387-406.
- Lau, R., Lepper, M. R., & Ross, L. Persistence of inaccurate and discredited personal impressions: A field demonstration of attributional perseverance. Unpublished manuscript, Stanford University, 1976.
- Lefcourt, H. M. Internal vs. external control of reinforcement revisited: Recent developments. In B. A. Maher (Ed.), *Progress in experimental personality research*. Vol. 6. New York: Academic Press, 1972.
- Lemann, T. B., & Solomon, R. L. Group characteristics as revealed in sociometric patterns and personality ratings. *Sociometry*, 1952, 15, 7-90.
- Lepper, M. R., & Greene, D. Turning play into work: Effects of adult surveillance and extrinsic rewards on children's intrinsic motivation. *Journal of Personality and Social Psychology*, 1975, 31, 479-486.

- Lepper, M. R., Greene, D., & Nisbett, R. E. Undermining children's intrinsic interest with extrinsic reward: A test of the "overjustification" hypothesis. *Journal of Personality and Social Psychology*, 1973, 28, 129-137.
- McArthur, L. A. The how and what of why: Some determinants and consequences of causal attribution. *Journal of Personality and Social Psychology*, 1972, 22, 171-193.
- McArthur, L. A. The lesser influence of consensus than distinctiveness information on causal attributions: A test of the person-thing hypothesis. *Journal of Personality and Social Psychology*, 1976, 33, 733-742.
- McGuire, W. J. The yin and yang of progress in social psychology: Seven koan. *Journal of Personality and Social Psychology*, 1973, 26, 446-456.
- Milgram, S. Behavioral study of obedience. *Journal of Abnormal and Social Psychology*, 1963, 67, 371-378.
- Miller, A. G., Gillen, B., Schenker, C., & Radlove, S. Perception of obedience to authority. *Proceeding, 81st Annual Convention, American Psychological Association*, 1973, 8, 127-128.
- Miller, D. T., & Ross, M. Self-serving biases in the attribution of causality: Fact or fiction? *Psychological Bulletin*, 1975, 82, 213-225.
- Mischel, W. *Personality and assessment*. New York: Wiley, 1968.
- Mischel, W. Continuity and change in personality. *American Psychologist*, 1969, 24, 1012-1018.
- Mischel, W. Towards a cognitive social learning reconceptualization of personality. *Psychological Review*, 1973, 80, 252-283.
- Mischel, W. Processes in delay of gratification. In L. Berkowitz (Ed.), *Advances in experimental social psychology*. Vol. 7. New York: Academic Press, 1974.
- Mischel, W., & Ebbesen, E. B. Attention in delay of gratification. *Journal of Personality and Social Psychology*, 1970, 16, 329-337.
- Mischel, W., Ebbesen, E. B., & Zeiss, A. R. Cognitive and attentional mechanisms in delay of gratification. *Journal of Personality and Social Psychology*, 1972, 21, 204-218.
- Nisbett, R. E., & Borgida, E. Attribution and the psychology of prediction. *Journal of Personality and Social Psychology*, 1975, 32, 932-943.
- Nisbett, R. E., Borgida, E., Crandall, R., & Reed, H. Popular induction: Information is not always informative. In J. Carroll & J. Payne (Eds.), *Cognitive and social behavior*. Potomac, Md.: Lawrence Erlbaum Associates, 1976.
- Nisbett, R. E., Caputo, C. G., Legant, P., & Maracek, J. Behavior as seen by the actor and as seen by the observer. *Journal of Personality and Social Psychology*, 1973, 27, 154-164.
- Nisbett, R. E., & Wilson, T. D. Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, 1977, in press.
- Peterson, C. R., & Beach, L. R. Man as an intuitive statistician. *Psychological Bulletin*, 1967, 68, 29-46.
- Polefka, J. The perception and evaluation of responses to social influences. Unpublished doctoral dissertation, Stanford University, 1965.
- Regan, D. T., & Totten, J. Empathy and attribution: Turning observers into actors. *Journal of Personality and Social Psychology*, 1975, 32, 850-856.
- Ross, L., Amabile, T. M., Jennings, D. L., & Steinmetz, J. L. Non-conservative and non-optimal prediction strategies: Experiments on the psychology of intuitive prediction. Unpublished manuscript, Stanford University, 1976. (a)
- Ross, L., Amabile, T. M., & Steinmetz, J. L. Social roles, social control, and biases in social perception processes. *Journal of Personality and Social Psychology*, 1977, in press. (a)
- Ross, L., Bierbrauer, G., & Hoffman, S. The role of attribution processes in conformity and dissent: Revisiting the Asch situation. *American Psychologist*, 1976, 31, 148-157. (b)

- Ross, L., Bierbrauer, G., & Polly, S. Attribution of educational outcomes by professional and non-professional instructors. *Journal of Personality and Social Psychology*, 1974, 29, 609-618.
- Ross, L., Grèene, D., & House, P. The false consensus phenomenon: An attributional bias in self perception and social perception processes. *Journal of Experimental Social Psychology*, 1977, in press. (b)
- Ross, L., Lepper, M., & Hubbard, M. Perseverance in self perception and social perception: Biased attributional processes in the debriefing paradigm. *Journal of Personality and Social Psychology*, 1975, 32, 880-892.
- Ross, L., Lepper, M. R., Strack, F., & Steinmetz, J. L. The effects of real and hypothetical explanation upon future expectations. Unpublished manuscript, Stanford University, 1976. (c)
- Rotter, J. B. Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs*, 1966, 80, No. 609.
- Schachter, S., & Singer, J. E. Cognitive, social and physiological determinants of emotional state. *Psychological Review*, 1962, 69, 379-399.
- Slovic, P., & Lichtenstein, S. Comparison of Bayesian and regression approaches to the study of information processing in judgment. *Organizational Behavior and Human Performance*, 1971, 6, 649-744.
- Smedslund, J. The concept of correlation in adults. *Scandinavian Journal of Psychology*, 1963, 4, 165-173.
- Smith, A. The attribution of similarity: The influence of success and failure. *Journal of Abnormal and Social Psychology*, 1960, 61, 419-423.
- Storms, M. Videotape and the attribution process: Reversing actors' and observers' points of view. *Journal of Personality and Social Psychology*, 1973, 27, 165-175.
- Strickland, L. H. Surveillance and trust. *Journal of Personality*, 1958, 26, 200-215.
- Taylor, S. E., & Fiske, S. T. Point of view and perceptions of causality. *Journal of Personality and Social Psychology*, 1975, 32, 439-445.
- Thibaut, J. W., & Riecken, H. W. Some determinants and consequences of the perception of social causality. *Journal of Personality*, 1955, 24, 113-133.
- Tversky, A., & Kahneman, D. Belief in the law of small numbers. *Psychological Bulletin*, 1971, 76, 105-110.
- Tversky, A., & Kahneman, D. Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, 1973, 5, 207-232.
- Tversky, A., & Kahneman, D. Judgment under uncertainty: Heuristics and biases. *Science*, 1974, 185, 1124-1131.
- Valins, S. Persistent effects of information about internal reactions: Ineffectiveness of debriefing. In H. London & R. E. Nisbett (Eds.), *Thought and feeling: Cognitive modification of feeling states*. Chicago: Aldine, 1974.
- Wallach, H. The perception of motion. *Scientific American*, 1959, 201, 56-60.
- Walster, E. Assignment of responsibility for an accident. *Journal of Personality and Social Psychology*, 1966, 3, 73-79.
- Walster, E., Berscheid, E., Abrahams, D., & Aronson, V. Effectiveness of debriefing following deception experiments. *Journal of Personality and Social Psychology*, 1967, 6, 371-380.
- Ward, W. D., & Jenkins, H. M. The display of information and the judgment of contingency. *Canadian Journal of Psychology*, 1965, 19, 231-241.
- Wason, P. C., & Johnson-Laird, P. N. *Psychology of reasoning: Structure and content*. London: Batsford, 1972.



- Weiner, B. *Achievement motivation and attribution theory*. Morristown, N.J.: General Learning Press, 1974.
- Wheeler, G., & Beach, L. R. Subjective sampling distributions and conservatism. *Organizational Behavior and Human Performance*, 1968, 3, 36-46.
- Wolosin, R. J., Sherman, S. J., & Till, A. Effects of cooperation and competition on responsibility attribution after success and failure. *Journal of Experimental Social Psychology*, 1973, 9, 220-235.
- Zadny, J., & Gerard, H. B. Attributed intentions and informational selectivity. *Journal of Personality and Social Psychology*, 1974, 10, 34-52.