

# The Processes of Causal Attribution<sup>1</sup>

HAROLD H. KELLEY *University of California, Los Angeles*<sup>2</sup>

This article summarizes and synthesizes my several papers on attribution theory (Kelley, 1967, 1971, 1972). Lest the reader incorrectly conclude that these writings define attribution theory, note must be made of other statements of the theory. These include, most importantly, Fritz Heider's (1958) book, *The Psychology of Interpersonal Relations*, which has played a central role in the origination and definition of attribution theory and continues to be the major source of ideas; Ned Jones and Keith Davis' (1965) systematic hypotheses about the perception of intention in their essay on "From Acts to Dispositions"; and Daryl Bem's imaginative work on the theory of self-perception (summarized in a recent paper, 1972). Deserving attention as a broad survey of the field is the recent volume edited by Jones, Kanouse, Kelley, Nisbett, Valins, and Weiner (1972), entitled *Attribution: Perceiving the Causes of Behavior*.

Attribution theory is a theory about how people make causal explanations, about how they answer questions beginning with "why?" It deals with the information they use in making causal inferences, and with what they do with this information to answer causal questions.

The theory has developed within social psychology primarily as a means of dealing with questions of *social perception*: If a person is aggressively competitive in his behavior, is he this kind of person, or is he reacting to situational pressures? If a person advocates a certain political position, does this reflect his true opinions, or is it to be explained

in some other way? If a person fails on a test, does he have low ability, or is the test difficult? In all such instances, the *questions* concern the causes of observed behavior and the *answers* of interest are those given by the man in the street. Thus, attribution theory concerns what Heider has called "naive psychology."

The theory also deals with questions of *self-perception*. The interest here stems in the first place from Festinger's (1954) theory of social comparison processes and Schachter's subsequent work on emotions (1959; Schachter & Singer, 1962). A second major impetus here has been provided by Bem's (1965, 1967, 1972) statements of the conditions under which one's attitudes are known to oneself. The common problems in this area concern the person's judgments of his own ability, feelings, attractiveness, etc.

Beyond these matters of social and self-perception, attribution theory is related to a more general field that might be called *psychological epistemology*. This has to do with the processes by which man "knows" his world and, more importantly, *knows that he knows*, that is, has a sense that his beliefs and judgments are veridical. The ascription of an attribute to an entity amounts to a particular causal explanation of effects associated with that entity—reactions or responses to it, judgments and evaluations of it, etc. So all judgments of the type "Property X characterizes Entity Y" are viewed as causal attributions.

These examples will make it clear why attribution theory has developed within social psychology, for therein is found the concern with the interpretation of other persons' behavior. But it will also be clear that attribution theory is relevant to other fields of psychology, particularly those in which self-concepts are regarded as important. And as a general conception of the way people think about and analyze cause-effect data, attribution theory might have emerged from any of the classical fields

---

<sup>1</sup> This article was prepared with the support of Grant GS-1121X from the National Science Foundation. Portions were presented at the 20th annual meeting of the Japan Group Dynamics Association, Kyoto, August 1972, and as a Distinguished Scientific Contribution Award Address at the annual meeting of the American Psychological Association, Honolulu, September 1972.

<sup>2</sup> Requests for reprints should be sent to Harold H. Kelley, Department of Psychology, University of California, Los Angeles, California 90024.

of psychology concerned with perception, judgment, and thinking.

If I repeatedly refer to attribution *theory*, it should not lead the reader to expect too much in the way of a systematized set of assumptions, propositions, and deductions. "Theory" is used here in a broad and, I would insist, entirely appropriate sense, to refer to a more or less plausible set of general principles offered to explain certain observed phenomena. The examples mentioned above define a large class of questions and problems to which attribution theory affords a common approach.

As we consider some of the generalizations of attribution theory, the reader also must be warned that these will strike him as very simple and obvious ideas, very much a part of common sense. If I have been clear so far, this will come as no surprise, for it is precisely common sense with which attribution theory is concerned. This being the case, and inasmuch as the reader possesses common sense, he will find that he already "knows" much of attribution theory. However, it is our experience that the careful explication and systematization of what at first seems obvious eventually carries us into realms of discovery and insight. I believe social psychologists finally are realizing that their proper role is not to confound common sense but rather to analyze, refine, and enlarge on it.

### Two Systematic Statements of Attribution Theory

The ideas in attribution theory deal with two different cases, distinguishable in terms of the amount of information available to the attributor (the person making the attribution):

Turn	CASE I		CASE II	
	Bill	Bill's Partner	Bob	Bob's Partner
1st	Cooperative		Cooperative	
2nd		Cooperative		Competitive
3rd	Competitive		Competitive	
Result:	84% judge Bill to be competitive		62% judge Bob to be competitive	

FIG. 1. Inference of cooperativeness from moves in the Prisoner's Dilemma Game.

Case 1: The attributor has information from multiple observations.

Case 2: The attributor has information from only a single observation.

Two sets of principles have evolved to deal with these different cases. The first case permits the attributor to observe and respond to the *covariation* between an observed effect and its possible causes. (I have written about this case in the 1967 Nebraska Symposium paper, and in the 1971 module on "Attribution in Social Interaction" under the "Covariation Principle.")

The second case requires the attributor to respond to the set of conditions present at a given time. It is necessary for him to take account of the *configuration* of factors that are plausible causes for the observed effect. (This case is considered in the 1971 module under the "Discounting Principle," and more systematically in the 1972 module on "Causal Schemata and the Attribution Process.") These two cases will now be considered in turn.

#### COVARIATION CONCEPTS

These concepts can be illustrated best by the covariation principle which is as follows: *An effect is attributed to the one of its possible causes with which, over time, it covaries.* The principle applies when the attributor has information about the effect at two or more points in time. In some studies, the length of the series of observations is minimal—two. For example, in the Kelley and Stahelski (1970a) study of inferences of cooperativeness from moves in the Prisoner's Dilemma Game, only two actions of the stimulus person are given: his first one and, after he knows the partner's move, his second one. Thus, in Figure 1, we see Bill's cooperative move on his first turn and, later, his second move on the third turn after he knows his partner's first move. Similarly for Bob, we know his cooperative move on the first turn and his competitive move after he knows his partner's move. The results provide a confirmation of the covariation principle. The person who shows a shift from cooperative to competitive behavior that is covariant with the partner's competitive move (Bob) is judged less often to be competitive than is the person who shows the same shift independently of the partner's (cooperative) action (Bill). Apparently Bob's covariant shift is attributed to

the partner's influence or example, and Bill's independent shift, to Bill's own intentions.<sup>3</sup>

In a number of self-perception studies, the attributor has had a longer series of observations over which to notice the covariation between an effect and possible causes. For example, Valins (1966) presented his male subjects with feedback of their heartbeats which were made to appear to change in rate covariantly with the presentation of the photographs of certain nude females. His results suggest that the covariant pattern induced an attribution of special properties to the pictured figures. The women associated with heart rate changes were rated as more attractive, and their photographs were chosen more often by the subjects as rewards for their experimental participation.<sup>4</sup>

Implicit in the covariation principle is the important and little investigated problem of the exact temporal relations assumed to exist between a cause and its effect. The notion of covariation implies temporal contiguity, that is, that there are occasions when both an effect and its cause are present and other occasions when both are absent. Michotte's (1963) remarkable studies of the per-

---

<sup>3</sup> In the Prisoner's Dilemma Game, the partner's competitive action is a plausible cause for one's competitive behavior, so Bob's switch to competition is attributed readily to the partner. On the other hand, the partner's cooperative action is not clearly a plausible cause for one's cooperative behavior. His cooperation provides as much a reason for one's own competition (in order to exploit him) as for one's own cooperation (to establish a working relationship with him). Consequently, as our results show, the person who shifts to cooperation after his partner's cooperative act is judged no less cooperative than is the person who shifts to cooperation independently of the partner's (competitive) move. This ambiguity in interpreting the causal role of the partner's cooperative move was pointed out to us by Wayne Osgood. The results raise the problem, discussed later, of the interplay between the attributor's preconceptions of causal factors (e.g., their plausibility) and the current information he has about covariation between an effect and possible causes.

<sup>4</sup> To be perfectly accurate, Valin's procedure may only have raised the possibility that the particular females had special properties. Barefoot and Straub (1971) subsequently showed that the effect Valins reported occurs only when subjects are allowed ample time to observe the pictures. Apparently, the covariant changes in heart rate create some attributional instability or uncertainty that motivates informational search. When such search reveals good reasons for the effect (as Valins' subjects report to be the case in viewing the nudes), there occurs the appropriate changes in attribution to the object and in the evaluation of it. This problem is discussed in detail by Nisbett and Valins (1971).

ception of causality suggest that not only is a close temporal relation essential to a causal interpretation, but also that the ordinal relation is important. Heider and Simmel (1944) also observed that "The unit which moves first is . . . more likely to be seen as the origin [p. 256]." Apparently, effects ordinarily are assumed to occur *closely after* their causes. A first investigation of these temporal factors as they operate in interpersonal perception has been conducted by Bavelas and his colleagues (Bavelas, Hastorf, Gross, & Kite, 1965).

It is convenient to conceptualize the process under discussion, the inference of cause from the observation of covariation, in terms of the analysis of variance as employed in psychology to interpret experimental results. Following Heider's suggestion that causal analysis is "in a way analogous to experimental methods [p. 297]," I developed the analysis of variance analogue in an earlier paper (Kelley, 1967). The assumption is that the man in the street, the naive psychologist, uses a naive version of the method used in science. Undoubtedly, his naive version is a poor replica of the scientific one—incomplete, subject to bias, ready to proceed on incomplete evidence, and so on. Nevertheless, it has certain general properties in common with the analysis of variance as we behavioral scientists use it.<sup>5</sup>

In this analysis of variance, the salient possible causes constitute the independent variables and the effect constitutes the dependent variable. For a wide range of attribution problems, the classes of possible causes are as shown in Figure 2: persons, entities, and times. An example will illustrate

---

<sup>5</sup> Much earlier, Karl Duncker (1945) had expressed the same view in these terms:

The essential cause of an effect b—or that which "leads to b"—can be grasped by the abstraction of those further factors which all b-situations have in common, and of those which all comparable non-b-situations lack in common. Such an "abstracting induction" takes place everywhere in the practical life of man and animal, only less systematically than in the sciences [p. 64, italics omitted].

Duncker noted too, as I have here, the close relation to the covariation criterion of the factor of contiguity: "We generally perceive as 'cause' of an event, of a singularity, another singularity which coincides spatially and above all temporally with the first [p. 67]." Later in this article, I will describe Duncker's further suggestion that certain cause-effect connections are "intelligible" because there is something about the form or material of the effect that singles out its cause from other conceivable causes.

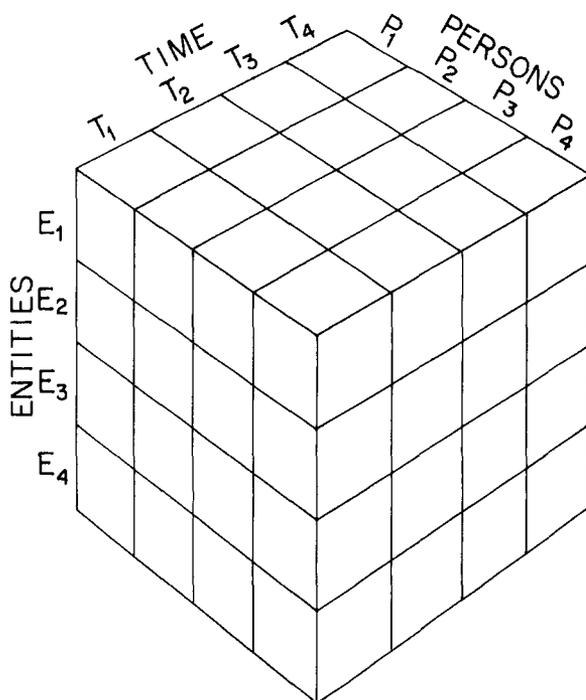


FIG. 2. The analyses of variance framework for making causal inferences.

the model. Consider the following information (adapted from McArthur, 1972): "Paul is enthralled by a painting he sees at the art museum. Hardly anyone who sees the painting is enthralled by it. Paul also is enthralled by almost every other painting. In the past, Paul has almost always been enthralled by the same painting." Given that information, the subject is asked what probably caused this event—Paul's being enthralled by the painting—to occur? Was it something about Paul (person), something about the painting (entity), something about the particular circumstances (time), or some combination of these factors?

The information providing the context for interpreting the effect can be summarized by the data pattern in Figure 3. The effect occurs only for  $P_2$  (Paul), but it occurs for him at various times and for various entities. This pattern suggests that the effect, "enthralment," is dependent on Paul. It is "caused" by Paul—some property, characteristic, or predisposition of him. In McArthur's data, 85% of her college student subjects explain items such as this in terms of something about the person, and almost none of her subjects, in terms of something about the entity.

In contrast, consider the following information

pattern: "Sue laughed at the comedian she heard the other night. Almost everyone who hears the comedian laughs at him. Sue does not laugh at almost any other comedian. In the past, Sue has almost always laughed at the same comedian." The covariation pattern is summarized in Figure 4, which has the effect distributed over the top layer of the analysis of variance cube. The pattern suggests that Sue's laughter is caused by the comedian. He is apparently a funny or, in other ways, a laughter-provoking entity. (In McArthur's data, 61% of the subjects gave an entity explanation for this kind of covariation pattern and only 12% gave a person explanation.)

Not surprisingly, the explanation in terms of "the particular circumstances"—a package of unspecified but transient causal factors—is most frequent (on the order of 70%) when the data pattern is that shown in Figure 5: few other people laugh at the comedian, Sue does not laugh at other comedians, and she has not laughed at this one in the past.

Frieze and Weiner (1971) presented a very similar pattern of results from a study of the explanations given for success and failure. The effect, either success or failure, was presented in different

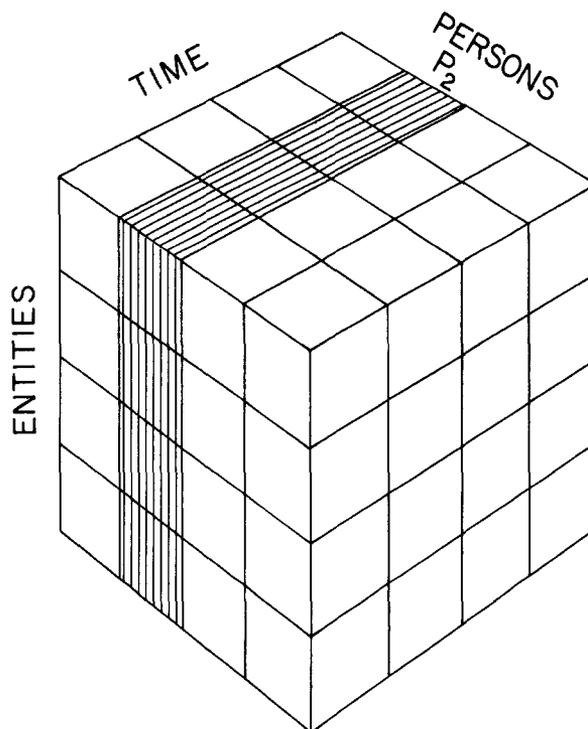


FIG. 3. Data pattern indicating attribution to the person.

informational contexts, and the subjects judged the degree to which various personal and impersonal factors were causes for the effect. One informational context was similar to that in Figure 3. The person had experienced the same outcome, say, success, on this particular task and on similar tasks in the past, but very few other persons had the same outcome. This information resulted in high ratings of causal factors within the person (his ability and his effort) and low ratings of external causal factors (the task and luck). The information pattern corresponding to Figure 4 described the person as having, for example, succeeded on similar tasks in the past but as having, along with most other persons, consistently failed the present task. His latest failure was judged to be caused by the particular task and not by his own properties. Finally, corresponding to Figure 5, the person's present failure on a task was presented in the context of his always having succeeded on it and similar tasks in the past and all other persons' having succeeded on it as well. In this context, the present failure was attributed to causes that, like McArthur's "particular circumstances," are unstable and fleeting, namely, bad luck and lack of effort.

One implication of the analysis of variance con-

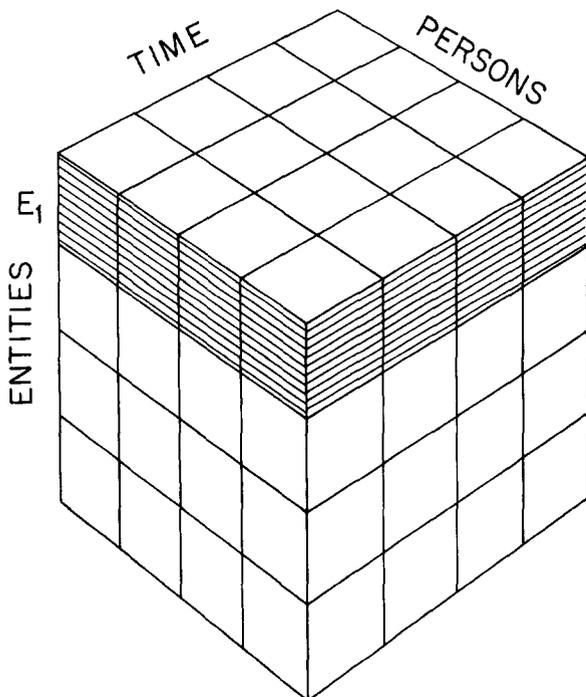


FIG. 4. Data pattern indicating attribution to the entity.

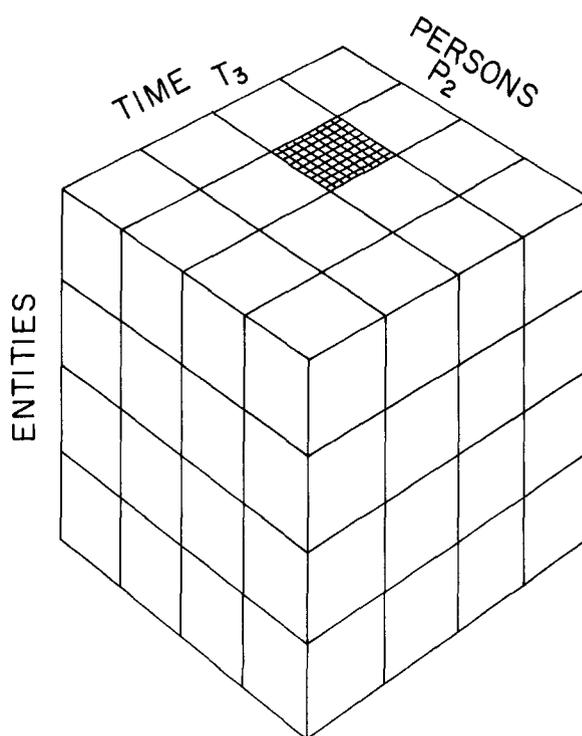


FIG. 5. Data pattern indicating attribution to the "circumstances."

ception is that not all patterns of data will be equally easy for the attributor to interpret. Certain patterns (corresponding to "main effects") are familiar to him as indicating effects due to a particular person or entity. An Entity  $\times$  Person pattern also seems to be rather readily interpretable. Frank consistently enjoys a particular "rock" record and only that one record, while none of his friends like it. This suggests a peculiar affinity between Frank and the specific record. More complex patterns entail greater ambiguity. For example, existing evidence suggests that the attributor has considerable difficulty coming to a single interpretation of information such as this: Steve admires his mathematics teacher on a particular occasion although he has never thought well of him in the past. However, he admires all his other teachers and all the other students admire this particular one. It is not that no explanation comes to mind for this effect. Quite to the contrary, several different interpretations seem to compete for the attributor's attention.

Of the various uses to which the analysis of variance conception of the attribution process can be put, one of the most important has to do with

the *phenomenology of attribution validity*. Here we deal with the particular aspect of self-knowledge that, following Lady Burton, might be described as wisdom: the wise man is "he who knows and knows he knows." How does a person know that his perceptions, judgments, and evaluations of the world are correct or true? This can be answered in terms of when it is he can confidently make an entity attribution for a perception, judgment, or evaluation. More specifically, the answer is defined by the Person  $\times$  Entity  $\times$  Time data framework: I know that my response to a particular stimulus is a *valid* one if (a) my response is associated *distinctively* with the stimulus, (b) my response is similar to those made by other persons to the same stimulus (there is *consensus*), and (c) my response is *consistent* over time—on successive exposures to the stimulus and as I interact with it by means of different sensory and perceptual modalities. The criteria of distinctiveness, consensus, and consistency correspond, of course, to the pattern in Figure 4 which supports an entity attribution.

Most of the terms in this formulation of subjective validity are familiar ones. *Consensus* has been shown in many experiments to afford a basis for confidence in one's judgment. Support from others tends to increase adherence to one's opinions, and disagreement with others tends to reduce certainty and to increase the likelihood of change (Hare, 1962, pp. 37–41). With regard to *temporal consistency*, John Harvey and I<sup>6</sup> have shown recently that it too provides a basis for confidence in the correctness of one's judgments. An earlier study by Gerard (1963) can be interpreted as making the same point. Several studies by Frank Irwin and his colleagues deal with the relation between *distinctiveness* and *consistency*. Irwin, Smith, and Mayfield (1956) required statistically naive subjects to estimate which of two decks of cards had, on the average, the larger numbers on their faces. They found that confidence in this judgment increased with increasing magnitude of difference between the means of the two decks (distinctiveness) and with decreasing magnitude of the variances of the decks (con-

sistency). In subsequent experiments, Irwin and Smith (1956, 1957) showed that the same two variables (distinctiveness and consistency) affect the number of observations subjects made before deciding whether the mean of a deck was larger or smaller than zero. More data were required the closer to zero was the deck's mean and the larger was its variance. Apparently the ratio of distinctiveness to consistency permitted the subjects to "know that they know" which mean is larger or how a mean compares with zero. Irwin's studies explicitly test and support the idea that naive subjects treat information informally in a manner similar to the way statisticians treat it formally.

The three criteria of validity—distinctiveness, consensus, and consistency—suggest a means of indexing an individual's *level of information* regarding any portion of his external world. An attributionally appropriate index consists of the ratio of the between-entity distinctions of which he is capable to the within-entity variance among his own responses and those of other persons—a ratio analogous to the familiar *F* ratio in statistics.<sup>7</sup> Information level, defined in this manner, provides a convenient approach to many of the classic problems of interpersonal influence based on information or expertise. The details are summarized elsewhere (Kelley, 1967; Kelley & Thibaut, 1969). Suffice it here to note that this analysis suggests:

---

<sup>7</sup> A suggestive parallel to this index is found in the concept of *codability* proposed by Brown and Lenneberg (1954). In naming a set of colors, codability is high if the instances are from discrete locations on the color continuum and do not fall near the boundaries between color categories. As Brown and Lenneberg show, high codability is indicated by interpersonal agreement in naming responses and interpersonal consistency over time, as well as by short reaction time and short naming responses, the last fact implying high distinctiveness. The index of information level proposed here would be high for this set of entities, and, accordingly, each person would be sure that he knows what each color "is." In the color-naming task, low distinctiveness appears in the use of long, compound color descriptions. Thus, for a highly similar set of colors, the descriptions given the various colors would have elements in common with each other ("green" versus "green with a tinge of blue" versus "blue-green"). And such descriptions would not be given with consistency or consensus. Consequently, confidence in the correctness of one's judgments of the colors would be low. Brown and Lenneberg show that the scientist can use codability as evidence of interiorized social norms. The present argument is that the various aspects of codability have special private significance to the person in relation to his sense of knowing.

---

<sup>6</sup> J. H. Harvey and H. H. Kelley. Sense of Own Judgmental Competence as a Function of Temporal Pattern of Stability-Instability in Judgment. In preparation.

(a) when a person will be susceptible to and, indeed, looking for informational influence (as suggested by the Irwin and Smith experiments mentioned above); (b) what the various bases of informational influence are; and (c) what factors affect the persistence of such influence. The Person  $\times$  Entity  $\times$  Time framework also affords an approach to the systematic analysis of an audience's reactions to a persuasive communicator, the framework affording in this case the basis for making causal inferences about the communicator's behavior and message.

#### CONFIGURATION CONCEPTS

The attribution process implied by the analysis of variance model is undoubtedly somewhat on the idealized side. It would be foolish to suggest that anything like a large data matrix is filled out with effect observations before a causal inference is made. The framework should be regarded as simply the context within which some limited and small sample of observations is interpreted. Beyond that, it is obvious that the individual is often lacking the time and the motivation necessary to make multiple observations. In these circumstances, he may make a causal inference on the basis of a *single observation* of the effect. In doing so, he is rarely acting in complete ignorance. Ordinarily he has observed similar effects before and has some notions about possibly relevant causes and how they relate to this type of effect. And, of course, his information about the present instance is likely to indicate the presence of certain plausible causes.

A first, simple statement about how the attributor thinks about such cases is provided by the discounting principle: *The role of a given cause in producing a given effect is discounted if other plausible causes are also present.* This principle is illustrated in many experiments, beginning with Thibaut and Riecken's (1955) demonstration that the "behavioral compliance" a subject evokes from a lower status target person is attributed less to him than is the similar compliance evoked in a higher status target person. In this experiment, a subject induces both a lower status person and a higher status person to comply with a request. In situations of this sort, it is probable that the subject assumes there are different sets of plausible causes for the two target persons' compliance. *Our*

EXPERIMENTAL CONDITION	PLAUSIBLE CAUSES	EFFECT
Low Status	Internal, External	---> Compliance
High Status	Internal	---> Compliance

FIG. 6. Plausible causes assumed by subject in the Thibaut and Riecken experiment.

assumptions about the *subject's* assumptions are shown in Figure 6. He assumes that the low-status person's compliance is caused either by his *internal* attitudes and dispositions (e.g., his helpfulness) or by the *external* pressure or force applied to him by the subject's request for help, or by both. On the other hand, the latter factor (the external pressure) is not assumed to be a plausible cause for the high-status person's compliance, he presumably being more powerful than the subject. Accordingly, his compliance is attributed to his internal properties. And, as suggested by the discounting principle, appropriate internal properties of the low-status person are inferred less strongly from *his* compliance inasmuch as there are plausible external reasons for it. Numerous other studies make basically the same point (Baldwin & Baldwin, 1970; Jones, Davis, & Gergen, 1961; Jones & Harris, 1967; Strickland, 1958; to mention a few).

The discounting paradigm is essentially equivalent to Bem's (1967) account of self-perceptions in forced-compliance experiments. Looking back on his own compliance (e.g., his delivery of a counter-attitudinal communication), the subject draws different inferences from it as to what his own attitudes on the issue must be, depending on the circumstances. If there was high external justification for the compliance, it is treated as in the top line of Figure 6, and his own internal causes (his own attitudes) are discounted as possibly causing the compliance. With low external justification, the situation is as in the bottom line of Figure 6, and one's own compliance is taken as a basis for inferring one's own attitudes.

When examined closely, the discounting principle involves many more subtle and complex points than can be explained here, but one example is worth noting. The external cause may be an *inhibitory* cause for the observed effect, that is, a cause that acts to suppress the observed effect. In this case, the presence of the external cause (given that the effect is observed to occur) serves to heighten the impression that an internal (facilita-

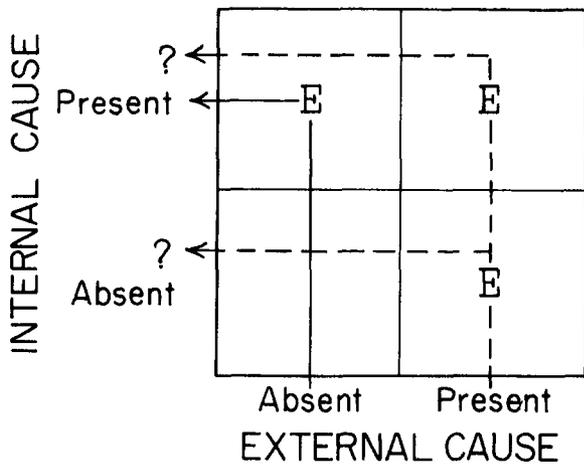


FIG. 7. Causal schema for multiple sufficient causes.

tive) cause is present and a potent force. (I have referred to this as the augmentation principle, Kelley, 1971.) For example, consider Frank who works on a very difficult task and succeeds, and compare him with Tony who works on a task of moderate difficulty and also succeeds. The presence of a cause counterindicative of success (the difficult task) provides a basis for a stronger, more confident attribution of ability to Frank than to Tony. The augmentation principle refers to the familiar idea that when there are known to be constraints, costs, sacrifices, or risks involved in taking an action, the action once taken is attributed more to the actor than it would be otherwise.<sup>8</sup>

One of the interesting consequences of making the distinction between *facilitative* and *inhibitory* external causes is that we then become aware of a class of instances in which there is considerable ambiguity about whether a given cause is assumed by the attributor to be one or the other. For example, is an outright payment to perform an immoral act assumed to be a *facilitative* or an *inhibitory* cause? Depending on which meaning it is given, the inference drawn from the performance of the act will be quite different. If taken as a *facilitative* cause, the payment will reduce the attribution of the immoral act to the actor and, therefore, will presumably reduce the negative

<sup>8</sup> Another version of this inference is the social desirability hypothesis of Jones and Davis (1965). Behavior of low social desirability is attributed more to the person than is behavior of high social desirability. The former usually implies an action contrary to social norms, that is, a behavior enacted despite inhibitory external causes.

evaluations and sanctions directed toward him. In contrast, interpreted as an *inhibitory* cause, the payment will increase the attribution of the act to the actor with a resulting increase in the negative moral evaluation made of him.

These uses of the discounting principle (and of its variant—the augmentation principle) assume that as the attributor witnesses an effect, he also notes the factors present that might plausibly have caused it, and then *takes account of them* in interpreting the effect. The question is, How does he take them into account? The two principles seem to imply somewhat different ways of doing so. It is natural to ask (a) whether there are *other* ways to take plausible causes into account and (b) what might be *all* the possible forms of such “taking into account.” It is with these matters that my recent paper (Kelley, 1972) on “Causal Schemata and the Attribution Process” deals. There, an attempt is made to specify the possible ways a person may take account of causes in relation to effects.

A causal schema refers to the way a person thinks about plausible causes in relation to a given effect. It provides him with a means of making causal attributions given only such limited information as is suggested by the preceding examples. For instance, the discounting principle implies a multiple sufficient cause schema. In the Thibaut and Riecken situation, the subject assumes that *either* the external pressure *or* the internal disposi-

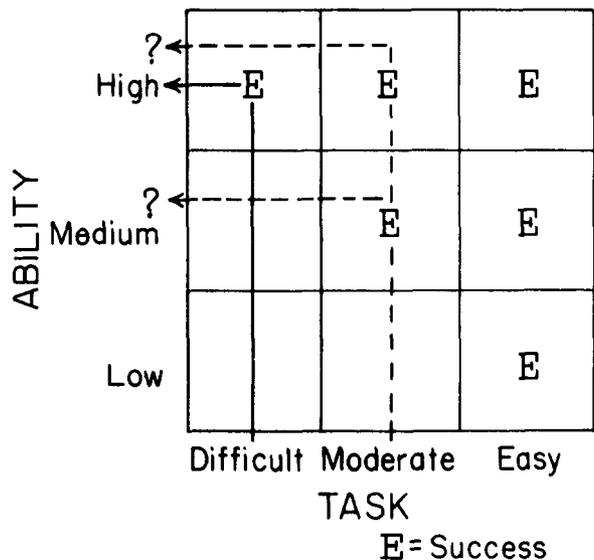


FIG. 8. Causal schema for compensatory causes. (E = success.)

tion is adequate to produce compliance. From the configuration of effects in relation to causes implied by this statement, as shown in Figure 7, the familiar inferences are implied. Given the effect and the external cause absent, an inference of the presence of the internal cause is indicated unequivocally, as shown by the solid arrow. Given the effect and the external cause present, there is uncertainty as to whether or not the internal cause was also present (shown by the dotted arrow).

The augmentation principle implies a slightly different, though related, configuration, one I have called a *compensatory cause schema*. As shown in Figure 8, it describes the trade-off between two quantitatively graded causes. In this case, the effect, success, depends on high ability or low task difficulty. The effect occurs if either cause is maximally favorable or if both are moderately so. Given success and the strong inhibitory cause present (high task difficulty), the internal cause, ability, is inferred to be present and strong. In the absence of the inhibitory cause, as with moderate task difficulty, success is ambiguous as to its implication for ability.

These examples illustrate more concretely what is meant by a *causal schema*:

[It] is a conception of the manner in which two or more causal factors interact in relation to a [given] effect. A schema is derived from experience in observing cause and effect relationships, from experiments in which deliberate control has been exercised over causal factors, and from implicit and explicit teachings about the causal structure of the world. . . . The mature individual . . . has a repertoire of [such] abstract ideas about the operation and interaction of causal factors. These conceptions [enable him to make] economical and fast attributional analysis, by providing a framework within which bits and pieces of relevant information can be fitted in order to draw reasonably good causal inferences [Kelley, 1972, p. 2].

More specifically (and this is the link between the analysis of variance and the schemata ideas), the causal schema can be viewed as "an *assumed pattern of data* in a complete analysis of variance framework." It is within this assumed configuration of data that the single observation is fitted and interpreted. "Given information about a certain effect and two or more possible causes, the individual tends to assimilate it to a specific assumed analysis of variance pattern, and from that to make a causal attribution [Kelley, 1972, p. 2]."

The "Causal Schemata" paper attempts to identify (a) what the possible schematic configurations

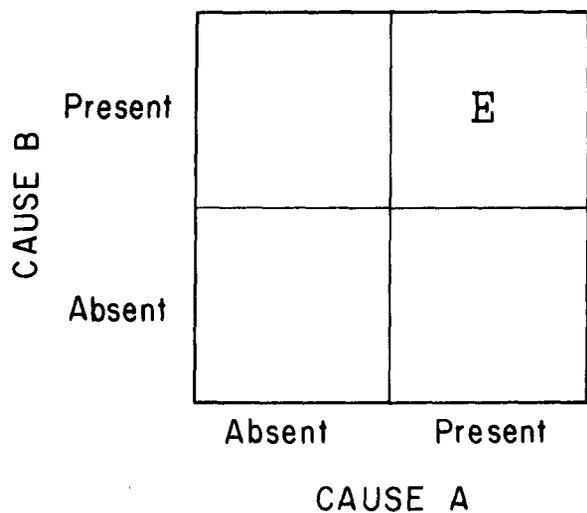


FIG. 9. Causal schema for multiple necessary causes.

are, (b) what inferential implications each one has, and (c) what seem to be the conditions under which each one is elicited. The other types of configurations, beyond the two described already, can only be mentioned here. Figure 9 shows the multiple necessary cause schema. Both causes must be present or favorable if the effect is to occur. The reader will readily see the different inferences possible with this pattern in comparison with those implied by the multiple sufficient cause schema. Most important is that the effect affords a basis for inferring the existence of, say, an internal cause *even though the external one is known also to be present*. This schema has been hypothesized as being evoked for extreme or unusual effects.

Linear patterns, such as are shown in Figure 10, will be recognized as those associated with person attributions or entity attributions. On the left is the pattern the attributor would assume if he took the effect to be caused by Entity X, and on the right, the pattern he would assume if he took the effect to be caused by Person B. The one-to-one pattern (Figure 11) represents an assumption appropriate for exclusivity relationships, such as prevail in ownership, marriage, etc.

Simple symmetrical patterns as in the left portion of Figure 12, symmetrical around the diagonal from upper left to lower right, represent an assumption of pairing or reciprocal relationships, such as are characteristic of sociometric choices. Finally, patterns characterized by both symmetry and transitivity, as in the right portion of Figure

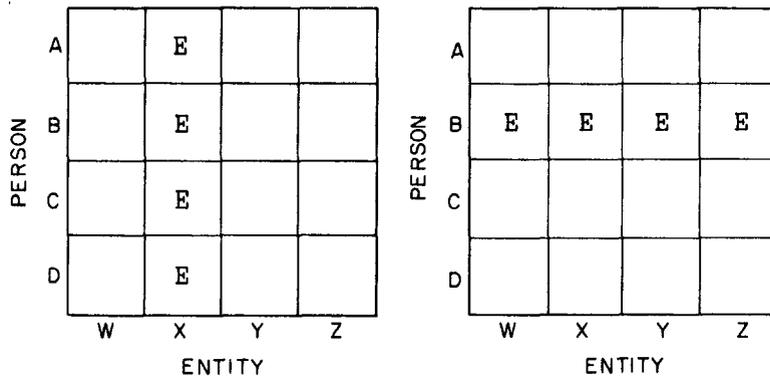


FIG. 10. Person-entity schemata assuming, respectively, entity attribution and person attribution.

12, probably are assumed in interpreting effects related to social groupings: the effect occurs as between members of the same group, but not between members of different groups. Results from DeSoto and Kueth (1959) suggest that the pairing schema usually is assumed for such interpersonal effects as *disliking*, *hating*, and *confiding*. The grouping schema tends to be evoked for such effects as *likes*, *trusts*, and *lies to*. That is, *dislike* or *hate* is an effect assumed to occur between pairs of people; *like* or *trust* is assumed to occur within groups of people. These results are far from conclusive, but they indicate the nature of the empirical generalizations eventually to be made about

the kinds of interpersonal effects eliciting each schema.

The types of inferences afforded by these symmetric configurations are perhaps obvious. Given information that Bill hates Jack, if the pairing schema is assumed by the attributor, he will draw the inference of reciprocation, that Jack also hates Bill. The grouping schema has more powerful implicational possibilities, deriving from the property of transitivity: Knowing that Jim trusts Jack and that Jack trusts George, the grouping schema permits the attributor to infer that Jim trusts George.

The schematic analysis suggests many other interesting problems and hypotheses. For example, a given attributional problem may contain conflicting cues as to the appropriate schema. This is probably true of one condition in Jones and Harris' (1967) experiment. The subject is told about a student who prepares an essay in which he advances an extreme and unusual point of view (suggesting the multiple necessary cause schema to be appropriate) but he does so under direct pressure from his instructor (suggesting the multiple sufficient cause schema to be appropriate). The first schema would lead to an inference from the student's essay that he holds corresponding opinions. The second schema would imply that the essay may have been due to the external pressure, so his opinions are in doubt. The high variability of attributional response to this information, reported by Jones and Harris, suggests that some subjects may base their attributions on the one schema, and other subjects, on the other schema.

A related problem concerns the different generalizations to be drawn as *extremity* or *magnitude* of

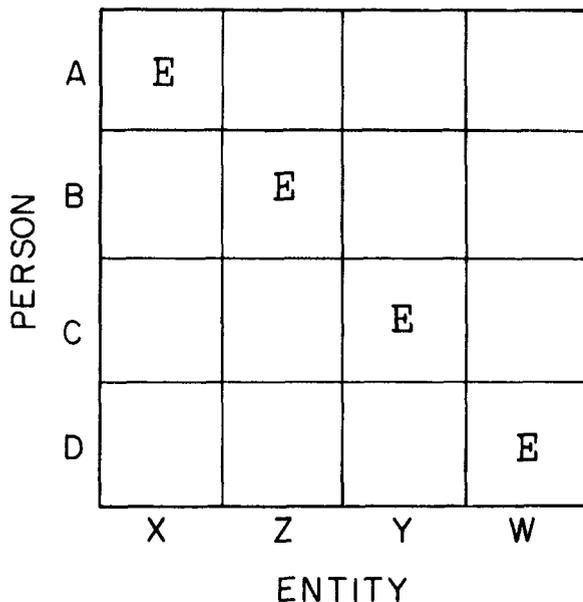


FIG. 11. Person-entity schema assuming one-to-one pairing.

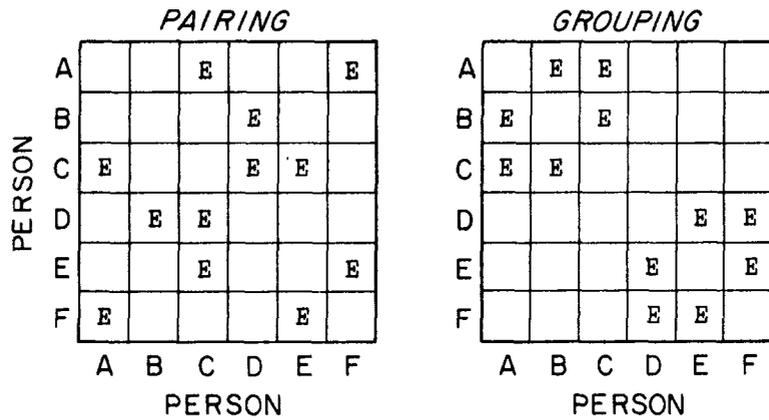


FIG. 12. Person-person schemata assuming, respectively, symmetry, and symmetry and transitivity.

effect increases. John Cunningham and I<sup>9</sup> have tested the hypothesis, mentioned earlier, that whereas effects of moderate magnitude are assumed to reflect the multiple sufficient cause schema, extreme effects are referred to the multiple necessary cause schema for their interpretation. We presented a sample of interpersonal events (e.g., liking, dominance) of varying magnitudes and measured their attribution to the actor and target person. Our results clearly confirm the hypothesis. They also suggest that some extreme interpersonal events are treated as *emergent* effects, unique to a particular actor and target person and not apparent in their other relationships. Other extreme interpersonal events are treated as *resultant* effects, being seen as summations of effects the causes produce separately. Thus, Bob's ecstatic love for Carol is likely to be seen as an emergent effect reflecting the unique conjunction of causal factors characteristic of the multiple necessary cause pattern—the special “chemistry” of their relationship. An observer is not likely to assume Bob will respond similarly to other young ladies, or that other men will respond similarly to Carol. In contrast, Bob's thorough and total victory over Ted in their tennis match, suggestive of the extreme effect in a compensatory cause (and additive effect) schema, as in Figure 13, implies that Bob will produce a similar effect in relation to other

tennis opponents and that Ted will suffer a similar fate at the hands of his other opponents.

This last schema, involving compensatory causes and additivity of effects, is worth special attention. It has wide applicability, for example, to person-task interactions with degree of success as the effect, and also to person-person interactions with dominance as the effect (the above example of Bob's victory over Ted). The pattern is essentially a combination of simpler schemata. In Figure 13, for example, the effect occurs to at least some degree if either cause is strongly facilitative or if both are weakly so. The effect increases in magnitude (the number of Es indicating magnitude of the effect) as these conditions favorable to its

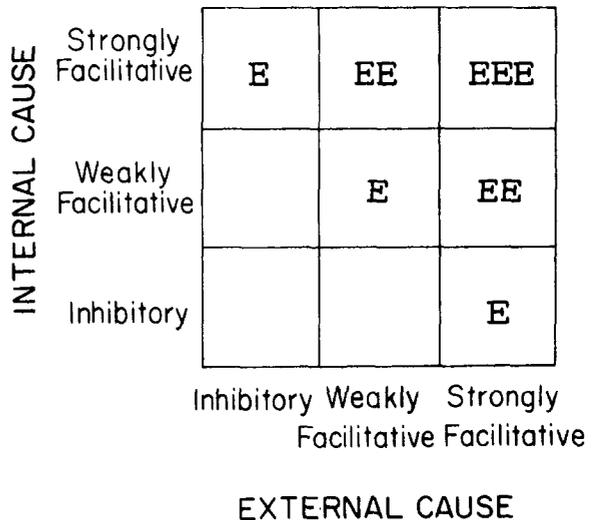


FIG. 13. Schema for compensatory causes and additive effects.

<sup>9</sup>J. D. Cunningham and H. H. Kelley. Causal Attributions for Social Events of Various Magnitudes. In preparation.

minimal occurrence cumulate in number. This configuration permits many interesting inferences. If the effect occurs in the presence of the inhibitory state of the external cause, a strong state of the internal cause is inferred—the augmentation principle described earlier. If the effect occurs to an extreme degree, *both* causes are inferred to be strongly facilitative—the extremity hypothesis we have confirmed recently. Thus, this pattern implies greater attribution to the person with increased magnitude of effect. This implication is consistent with various ideas and evidence that Hastorf, Schneider, and Polefka (1970) summarized in these terms: “It appears that the perceiver has increasing need to attribute responsibility to someone as the outcomes become more severe [p. 79].” Finally, the pattern has the subtle implication that a small effect produced in the presence of a strong external cause will lead to the inference of an *inhibitory* internal state. Evidence of this has been presented by Jones, Worchel, Goethals, and Grumet (1971). A person expressing a weakly favorable attitude when under pressure to express a strongly favorable one was inferred to have an unfavorable attitude.

#### IMPLICATIONS OF THE SCHEMATIC ANALYSIS

The preceding examples illustrate the possible uses of the schematic analysis. If we know which configuration a given causal problem will evoke for the attributor, we can predict, for various given patterns of information, what inferences he will make and with what certainty he will make them. As the examples show, he may be given an effect and one cause and then be asked about the state or magnitude of another cause. Alternatively, he may be given data only about the effect (and, perhaps, its magnitude) and be asked about certain causes. And, of course, he may be given information only about the causes and be asked to infer the effect.

The reader probably will find it plausible to assume that the layman has a repertoire of thought models for dealing with causal problems, a repertoire of causal schemata. The assumption seems reasonable in view of the typical person's experience with a wide variety of causal phenomena. The assumption gains plausibility as one tries to make sense (as I have, Kelley, 1972) of the many

studies in which causal assumptions and inferences have been assessed. However, in the context of the last 15 years of thought within social psychology, the notion of a *repertoire* of thought models is rather radical in its implications. This period has been characterized by proposals that the layman has one model or another and by the attempt to demonstrate the operation of each model. Thus, typical thought is said to be characterized by balance, or by assumed similarity, or by hierarchical organization, or by assumed consistency. Thus, it is said that the layman assumes that interpersonal relationships will be in balance. He expects to find them to be characterized by this property: if the information permits, he will see them and remember them as being so characterized, and if this property is missing or violated, he will show signs of discomfort. Alternatively (but similarly), it is proposed that the individual assumes that other persons hold attitudes similar to his own, he will tend to see them as doing so, and he is uncomfortable if they do not. And so on, for his assumption that another person will behave consistently in a variety of situations, for his assumption that the power relations among a set of persons permit them to be placed in a simple linear ordering, etc.

The present conception is, of course, that each and all of these models are reflected in the person's thinking—each at specific (and specifiable) times, and all, over a variety of occasions. The assumption of triadic balance is a manifestation of configurations characterized by symmetry and transitivity. Assumed similarity in attitude reflects the operation of an entity-attribution schema.<sup>10</sup> Assumed consistency of behavior, whether for one's self or for another, reflects the operation of a person-attribution schema. The assumption of a

---

<sup>10</sup> The causal schemata provide a useful way to interpret one of the most commonly observed departures from the predictions of balance theory. This is that two persons often are assumed to have the same attitude about a third person or entity even though they dislike each other (e.g., Rodrigues, 1968). This observation implies that whereas the disliking between the two is referred to a pairing or grouping schema, their attitudes toward certain other objects are referred to the linear, entity-attribution schema. When the subject thinks of personal disliking, he fits it into one of the schemata in Figure 12, but when he thinks of evaluations of other entities, he assimilates it to the left schema in Figure 10.

simple social hierarchy reflects a compensatory causal schema in which the same set of persons is ordered along the "person" and "entity" axes. Not all of the possible relations and parallels have been worked out, but the point is probably clear.

The research implications are probably also clear. Our initial problem is not one of proving or disproving the operation of one model or another. Rather it is one of identifying the entire set of models that are commonly or importantly used. Then, we must determine the conditions under which each of the set is evoked and the implications and consequences of its evocation. Some of the theoretical problems will be interesting ones, such as those arising when, in the presence of multiple cues, two or more schemata are brought into play. However, I suspect that much of the necessary research entails rather lackluster problems—problems that by their very nature do not have aesthetically appealing solutions. I have in mind such questions as when does a person assume there will be a general consensus on an issue, or when does he assume there will be a different consensus within each of several groups (simple entity-attribution schema versus Group  $\times$  Entity pattern). Closely related is the problem of objective versus subjective truths—entities assumed to elicit the same reaction from everyone *versus* entities assumed to be reacted to differently by different people, involving "matters of taste." These problems tend to be lacking in appeal because their answers depend on the specifics of time and place. Perhaps for this reason, they are some of the most frequently recurrent and most elusive problems within social psychology. A good example is the never satisfactorily answered question of who constitutes a reference group for a given person on a given issue.

If these are unappealing research questions, they are not unimportant ones. Social psychologists must deal with the particular substance and content of thought, and not just with its form, if they are to fulfill their mission in relation to the other behavioral sciences. This mission, as I understand it, is to provide the necessary theoretical tools and data for predicting how specific persons, at specific times and places, will react to their social environments.

## *Some Problems for Attribution Theory*

### THE INTERPLAY BETWEEN PRECONCEPTIONS AND NEW INFORMATION

The distinction drawn here, between two kinds of attribution processes, raises the following questions: What is the interplay between the two kinds of processes, the one based on observation and analysis of present information, and the other based on causal preconceptions and stereotypes (described here as causal schemata)? How do a priori causal beliefs affect the intake and processing of further information bearing on the attribution problem?

These questions are raised from another quarter by a superlative pair of studies conducted by Loren and Jean Chapman (1967, 1969), studies that also, incidentally, show one of the many implications of attributional analysis for the professional and scientific practices of psychologists themselves. The Chapmans begin with the specific question of why clinical psychologists persist in the use of projective test diagnostic signs that have been shown to be without validity. Their results bear on the more general problem of how causal beliefs affect the perception of covariation. They first show that clinically naive undergraduate students, when shown projective protocols in random pairings with statements of patients' problems, rediscover the same relations between test responses and problems that clinicians report observing in their practice. For example, even though not at all covariant in the experimental materials, drawing a large head on the Draw-A-Person Test is seen as being associated with worry about one's intelligence. Similarly, male patients' reports of anal openings and buttocks on the Rorschach are judged to covary with having sexual feelings toward other men.

It seems reasonable to characterize these "illusory" covariances as a priori causal beliefs—beliefs about what personal disposition is to be inferred as the cause for a given test behavior. One remarkable fact is how widely these beliefs are shared. The undergraduates tend to agree among themselves in what covariances they see in the experimental data, and they also agree with the relationships that clinicians report observing in the presumably equally random data they see. Of great importance for attribution theory, the Chapmans show in further experiments that these beliefs are strongly resistant to covariance data that discon-

firm their validity. Moreover, the beliefs interfere with the detection of other covariances existing between clinically valid signs and various personal problems.

There is some evidence in the Chapmans' data as to the origin of the beliefs they observe. The signs and problems seen erroneously to covary were found to have relatively strong "associative connections." Thus, the subjects reported a stronger tendency for homosexuality to "call to mind" sexual organs, or rectum and buttocks, or part man-part woman, than for it to call to mind monsters or part animal-part human. These connections probably reflect to some degree prior experience and information. Men who are worried about their manliness *are* concerned about their bodies, so it is plausible to expect them to draw broad-shouldered, muscular figures on the Draw-A-Person Test (although statistically they are not especially likely to do so). On the other hand, to some degree these beliefs probably reflect tendencies to make attributions on the basis of a simple correspondence or fittingness between cause and effect. It is fitting that homosexual feelings should cause the perception in inkblots of figures that have both female and male characteristics, so this cause-effect relationship is seen even without evidence of covariation. There is no such correspondence between homosexuality as a cause and perceiving animalized human figures as an effect.<sup>11</sup> So the latter covariance, one that has an objective existence, is easily overlooked unless (a) it is strongly *present* in the information sample and (b) the invalid stereotypic signs are *absent*.

The Chapmans' (1969) results may be cast in terms of the interplay between causal schemata and covariance information. Their experimental

procedure encourages the subject to treat the data in terms of person-causation patterns (cf. Figure 3 and right portion of Figure 10). The information is presented as coming from patients with different problems, and the task is to report "the kinds of things seen by patients with each kind of problem [p. 276]." Thus, although the information actually may contain no covariance patterns, the person-attribution schema is strongly implied and, as the results show, the hypothesis of this configuration tends to be confirmed. Of course, there is more to the Chapmans' results than this. The person schema is not confirmed in some arbitrary way or in different ways by different subjects. They all tend to assume the same plausible consistency across situations and times in the behavior of persons with certain characteristics. They do not assume, as they should, that behavior is caused jointly by person and situation and, therefore, that the diagnostic sign may bear little correspondence to the problem behavior. Furthermore, the fact that all persons make the same cause-effect assumptions may be an important factor in the maintenance of the assumptions even when they are incorrect. This is to suggest, as the Chapmans do, that the objective agreement among their subjects is known to the subjects themselves, on the basis of earlier communication about the properties of homosexuals, persons with inadequacy feelings, etc. So the causal beliefs probably have an assumed consensual basis for the various observers. These comments perhaps will indicate why the Chapmans' procedure evokes such strong and clear causal schemata and make it more understandable why these schemata are so resistant to counter-evidence.

The Chapmans' results have an important lesson for the development of attribution theory. Prior beliefs about causation affect the intake of information about covariation between various causes and effects. Our model of the attribution process must somehow take account both of the fact of prior beliefs and of their effect on the processing of new covariance information. This problem appears within psychology in many forms, for example, primacy versus recency, proactive versus retroactive inhibition, and assimilation versus accommodation. Further development of attribution theory requires an account of this conflict between existing cognitive structures and new data, and

---

<sup>11</sup> The phenomenon of attribution on the basis of correspondence between cause and effect will be discussed further in the next section of this article. It is not simply a matter of similarity between cause and effect. Drawing a big head is not similar to worry about intelligence, nor is drawing exaggerated sexual organs similar to impotence. Rather, these assumed effects are fitting or appropriate manifestations of these causes according to some implicit personality theory. Such implicit theories include not only instances in which the behavior is similar to a need (wearing feminine attire to express a desire to assume a female role), but also instances in which the behavior is compensatory for, or fulfills or completes, the personal property (stressing muscular development as a reflection of concern about masculinity).

of the process by which they interact and become reconciled.

#### SIMPLE VERSUS COMPLEX SCHEMATA

As repeatedly mentioned here, one major task for attribution theory is to specify when a given schema is evoked. Existing evidence shows that this depends in part on the type of causes and effects witnessed by the attributor. Beyond this, there may be a tendency for the attributor to prefer simple schemata over complex ones. Simple schemata are best illustrated by the linear or "main effect" patterns reflecting person attributions or entity attributions. They are essentially single-cause patterns, all the variation in effect being associated with variation in one cause or one type of cause. They are also patterns that assume similarity among the effects a given cause produces under various conditions. They do not admit of Person  $\times$  Entity patterns ("interaction effects" in the statistical sense). Thus, they correspond to such attributional stereotypes and beliefs as the following: "Good people cause good effects"; "Poor performance on an IQ test reflects low intelligence"; "A child who disobeys his parents is a bad child"; and "A man who likes to dress in women's clothes is a homosexual."

Simple beliefs such as these have been described by Duncker (1945) in his discussion of phenomenal causation based on correspondence of form or material. He notes that certain cause and effect connections are intelligible to the attributor without evidence of covariation: The track resembles the animal's foot, the wetness of rain becomes the wetness of the street, heavy things make "heavy" noises, and dainty things move daintily. The many such instances of intelligibility encountered by the attributor probably encourage him to adopt it as a widely useful principle that properties of the cause can be inferred on the basis of being correspondent with or similar to properties of the effect.

Heider (1958) referred to these simple beliefs, as they operate in interpersonal perception, in his oft-quoted phrase, that "behavior . . . tends to engulf the total field [p. 54]." Personal properties are inferred directly from behavior without its being interpreted in relation to the situation in which it occurs. Simple person attributions also are incorporated in balance theory, when certain

interpersonal configurations characterized by both sentiment and unit (causal) relations are considered. If P does not like O and does not like a Dessert X he is served at O's home, P can balance the P-O-X triad by thinking that O prepared the desert. Thus, affective similarity provides the basis for linking cause and effect. The good act is caused by the good person, and the bad act, by the bad one. One is reminded of Zillig's (1928) experiment in which two groups of youngsters performed a calisthenics drill before an audience of their classmates. Mistakes made during the performance were attributed by the audience to a disliked group of performers even though, as a result of the experimenter's deliberate intervention, the mistakes had actually been made by the liked group of children.

As to whether simple schemata, of the type these examples illustrate, are preferred to more complex ones, we may first note that simple schemata have a developmental precedence over complex ones. The belief that the individual is responsible in a simple way for what his actions bring to pass is characteristic of the developmental stage Piaget (1948) described as objective responsibility: "actions are evaluated in terms of material result and independently of motive . . . [p. 119]." Consequences are linked directly to the actor rather than being viewed as a joint function of him and the situation. The emphasis here is on the *evaluation* of behavior rather than *attribution*, but the concepts are presumably closely related. In his parallel treatment of levels of responsibility, Heider (1958) characterized this level as one at which "anything that is caused by  $p$  is ascribed to him [p. 113]." Impersonal or physical causality forms the primary basis for attribution. Personal causality, with a consideration of intentions in relation to anticipated consequences, plays little role.

Among the various studies deriving from these descriptions of developmental levels, those of Baldwin and Baldwin (1970; and also Baldwin, Baldwin, Castillo-Vales, & Seegmiller, 1971) show most clearly the increasing complexity of attributional assumptions with increasing age. This research was conceived largely within an attributional framework. Children were asked to make judgments of a boy's kindness on the basis of a story about his actions, the circumstances, and the consequences. Kindness is conceptualized as a motivation, that is, an internal causal factor. Some of the stories

most relevant here concern behavior that benefits another child and is enacted in either the presence or the absence of (a) an external cause (the mother's request that the boy perform the act), (b) an irrelevant internal cause (the beneficial act is reciprocated or is performed for a bribe), or (c) an inhibitory cause (the beneficial act involves self-sacrifice). In all of these cases, almost all eighth graders and college students in the Baldwins' samples make the judgment of kindness that takes account of the circumstances. The helpful act performed without the external cause or the irrelevant internal cause, and in the presence of the inhibitory cause, provides a basis for judging the actor as being kinder. In contrast, their kindergartners respond to the problems on a chance basis without apparently taking account of the circumstances. For these younger subjects, it seems that "behavior is behavior" and the circumstances are not important. They do not distinguish between the presence and absence of other causes and do not uniformly infer greater kindness from the one case than from the other.

Several of the results from these studies suggest that something more may be at work than merely a failure to take account of other possible causes of behavior in their attributions about the person. For example, a substantial number of children describe the obedient child as kinder than the spontaneously helpful one. This is particularly clear in a sample of children from the Yucatán where, the authors note, obedience to parents is a strong, intensely socialized value. There are also many children (a significant majority at the age of five) who interpret the helpful action done to return a favor as kinder than similar help given a person who had earlier refused to do a favor. In both cases, the children explain their judgments in terms of "objective responsibility": the act that conforms to social rules of fairness or obedience is the act that reveals kindness. Thus, the attribution is on the basis of affective similarity: An action that is "good" (by any criterion) indicates a "good" person.

It is proposed here that the preference for simple rather than complex causal explanations not only is characteristic of children but also that it persists into adulthood. The experiments by Chapman and Chapman (1967, 1969) described earlier might be interpreted in these terms. Their subjects tended to see covariances involving cor-

respondence between cause (the patient's problem) and effect (his projective test responses), even though these relations were not present in the data, and they tended to overlook covariances in which the symptom bore no simple correspondence to the problem. Another possible example is provided by the Jones and Harris (1967) study, which showed that the judgment of a student's opinion was influenced very much by an essay he wrote even though it had been written under strong, legitimate external pressure. While there are various interpretations to place on this result, certainly one of them is that some adult attributors take behavior as direct and simple evidence of internal cause—"A person's words express his attitudes."

These results are not presented as proving the existence of a preference for simple causal schemata, but simply as illustrations of the form such proof might take. Adults undoubtedly employ complex schemata more frequently than children do. There is little question that they are more able to do so than are children. However, the two studies mentioned above suggest that the tendency to employ simple schemata also may exist among adult attributors, even under circumstances where the use of such schemata is in conflict with other evidence in the situation. Perhaps there is an indication here of a preference for simple explanations—a preference that manifests itself only when the hypothesis of simple causation is not too strongly contraindicated by the data. It also might be suggested that this preference will appear only when the attribution task is clearly *not* structured as one in which the subject's subtlety and sophistication as an attributor are under evaluation. Probably rather few of the usual attribution experiments fulfill this last condition.

#### ATTRIBUTION OF COVARIATION AMONG CAUSES

The main theme of attribution theory, as described in this article, is that causal inferences generally are made in a manner that takes account of the joint contribution of multiple causes to a given effect. If, as described in the foregoing, our attributor often bases his inference on simpler assumptions, I believe he also *occasionally* bases it on more complex ones. The hypothesis proposed here is that he sometimes treats the various possible causes as themselves being interdependent and exercising influence over each other. He does

not always treat them as independent, as the covariance and schematic paradigms would imply. There is very little evidence on this point, so I will present it on logical and theoretical grounds.

As scientists, we know that causes do not occur independently and in all combinations. We conduct experiments precisely for the purpose of creating such circumstances. That is, we identify the independent variables and manipulate them orthogonally, and keep them separated from the dependent variables. At the same time, we know that such separation and independence are not characteristic of real life. It is for this reason that (as scientists) we are usually quite uneasy about interpreting naturally covarying data in cause and effect terms. It seems reasonable to assume that the lay attributor shares this awareness of the possible interdependence among causal factors and often assumes that certain causes of a given effect are themselves effects of other causes of the same effect.

If and when the layman makes such assumptions, the inferences he will draw from a given pattern of information may shift in a drastic manner. For example, following the implications of the schema for compensatory causes, we would portray our attributor as being uncertain about the difficulty of a task if all he knows is that a capable person has succeeded on it. Persons with high ability tend to succeed on all tasks, so an instance of this sort provides no information about the specific task. However, our attributor may make the assumption that ability and task difficulty are not independent but tend to covary. High ability tends to go with difficult tasks, and low ability, with easy ones. On this basis, and quite without regard to success or failure, information that the person is capable may lead to the inference that the task is difficult.

There are two reasons why a person might assume such covariance to exist between internal and external causes such as ability and difficulty: (a) one factor affects the other; for example, a person elects to work on tasks that are in keeping with his ability; and (b) some third remote causal factor affects both proximal causes; for example, the social system—teachers, parents, employers—pairs off persons and tasks so that ability tends to match difficulty.

Another example will serve further to illustrate the difference it makes for the attributor to assume covariation among causes. According to the mul-

tiply sufficient causal schema, the person who expresses a certain opinion when external causes are present is seen to a lesser degree to hold a corresponding attitude than is the person who expresses the same opinion in the absence of such external causes. This is asserted by Bem (1967) to be the basis of the *inverse* relationship between incentive to express a counterattitudinal opinion and subsequent self-report of the attitude. This inverse relation is the effect predicted by dissonance theory and has been found in a number of forced-compliance experiments. In contrast, if the attributor assumes the internal and external causes to covary, then a *positive* relationship might be anticipated between incentive and attitude report. For example, the attributor may assume that the high incentive induces a more favorable attitude, and, on that basis, infer the attitude directly from the external cause without reference to the stated opinion. Or a different assumption of interdependence among causes might form the basis for the inverse relation. The external pressure represented by the high incentive might be seen as dampening the person's internal interest in the matter, or the high incentive might be seen as selected by the experimenter in order to overcome the strong internal resistance. Working from either of these assumptions, the presence of high incentive would support an inference of strongly negative attitudes.

Here, again, a problem is posed for a complete theory of the attribution processes. Our description must include such complex assumptions as our attributors commonly make and provide generalizations about the conditions under which various assumptions are made. Even more difficult problems are posed for our methodology. An accurate elaboration of the theory along these lines will be possible only when we have clear evidence about what assumptions our subjects do in fact make for various specific attributional problems.

#### ERRORS, ILLUSIONS, AND DISCREPANCIES IN CAUSAL ATTRIBUTIONS

As is true of all perceptual and cognitive processes, the attribution process is subject to error. A review of the pertinent evidence by Peterson and Beach (1967) makes it clear that the man in the street can detect *covariation* between variables and can take account of the degree of *consistency* in a sample of data, but, also, that there are

"systematic discrepancies between normative and intuitive inferences [p. 43]." The major discrepancy seems to be that subjects are too conservative: "subjects apparently fail to extract all the information latent in samples of data [p. 43]." In relation to the present context, this result means that subjects do not gain as much confidence from a series of events characterized by consistency as a probability model suggests they should. They give too much weight to exceptions from general trends and treat successive observations as being interdependent, when, in fact, they are independent.

To complicate matters a bit, what appears to be the opposite tendency—to gain more certainty from a sample of data than it warrants—has been reported by Tversky and Kahneman (1971). They find that there exists among psychologists a tendency to view a sample as a representation of the population, "similar to the population in all essential characteristics." Thus, the results from small samples are expected to replicate to an unreasonable degree.

These two sets of observations appear to tell us something about biases in the processing of information relating to consistency, but I am not able to reconcile them. They are particularly difficult to interpret jointly because the conservatism typically has been shown with samples of statistically naive subjects, whereas Tversky and Kahneman report the beliefs of persons with statistical training. From the perspective of the present discussion, one wonders whether the various deviations from the statistical model may not themselves reflect attributional motives. That is to say, the tendency to assume there to be causes for all observed events may interfere with treating a sample of information in probabilistic terms. This possibility is suggested by Tversky and Kahneman (1971). They are writing of the psychological investigator, but the characterization is surely apropos of the naive scientist: "He rarely attributes a deviation of results from expectations to sampling variability, because he finds a causal explanation for any discrepancy. Thus, he has little opportunity to recognize sampling variation in action [p. 109]."

Certain errors in attribution appear to be analogous to what in other realms of perception are known as *illusions*. These consist of percepts the erroneous nature of which is apparent to the observer on his successive observations of the

stimulus or is easily made apparent to him by providing him with new information. In terms of the analysis of variance conception, some illusions, such as the reversible figures, are apparent as inconsistencies over time in the appearance of the stimulus. Others, such as the Müller-Lyer illusion, are only apparent as inconsistencies between different modalities of interaction with the stimulus (judging it by eye versus measuring it with a ruler). For both types, the illusional percept is produced by an interplay among the various components of information present at a given time. Thus, geometric illusions are errors in the perception of part of a figure as a consequence of how it is related to the whole figure. All persons with the same information tend to be subject to a given illusion, so the analysis of variance pattern potentially provides a consensual validation of the illusional percept.

Temporally unstable attributions, analogous to reversible figures, have not been studied yet. The conditions for attaining flip-flop attributions probably can be created by the presentation of two sets of cues that are both inconsistent with each other (see earlier examples) and difficult to attend to simultaneously.

Illusions of the second sort, where the attribution shifts with a change in informational modality, probably are illustrated by the "illusion of freedom." This has been proposed as a mechanism essential to the success of forced-compliance experiments (Kelley, 1967). In these studies, all subjects are induced to comply, but (at least in the crucial conditions) each subject is left feeling that he was acting freely. Thus, the attribution of external causation that an observer to all of the experimental sessions would make is quite different from the internal attribution for his behavior that each subject is led to make. A review of the forced-compliance procedures suggests that the subject's illusion of freedom is induced by such things as (a) leading him to believe that not all subjects, and perhaps few others, comply;<sup>12</sup> (b) labeling the situation as one involving free choice; (c) implying that the compliance is mildly socially undesirable; and (d) concealing the external pressure by exerting it in subtle, confusing, and nonmemorable

---

<sup>12</sup> A recent study by Cooper, Jones, and Tuller (1972) suggests that this factor may not be a necessary condition for the success of the forced-compliance procedure.

ways. In short, "strong situational demands, entirely sufficient to produce total or near-total compliance, are successfully camouflaged by a network of cues as to self-determination [p. 229]."

Although this analysis may not give us a complete understanding of the forced-compliance procedure, it does illustrate an important point. This is that, through proper manipulation of causes and information about causes, actors can be uniformly led to highly erroneous views about the reasons for their own behavior. The concepts outlined earlier suggest the terms in which these manipulations can be described. Thus, the true pattern of covariance between possible causes and effects may be concealed and/or distorted. Through special characterization of the behavior, the actor may be encouraged to assimilate it to and interpret it within one causal schema rather than another. Similar schematic manipulation may entail making salient the presence of certain plausible causes and interfering with the actor's attention to others.

Thinking in these terms, eventually we will be able to identify other attributional illusions and to describe the conditions necessary for their production. These will include the illusion of responsibility, whereby a person assumes to an unrealistic degree that he is personally responsible for consequences that other persons would attribute largely to circumstances or luck, and the illusion of external constraint, in which the person attributes to the situation or other persons consequences that are generated by his own behavior. The latter is illustrated by the competitive person's tendency to overlook his role in eliciting competitive behavior from others and his related belief that everyone is competitive and that is why he too must be competitive (Kelley & Stahelski, 1970b).

It is possible to speak of errors and illusions when there is some criterion of objective accuracy, that is, when the investigator has attributional procedures in which he is willing to place more confidence than in those of his subjects. Without such a criterion, it is still possible to identify systematic discrepancies between the attributions different persons or different types of persons make. A very important discrepancy of this sort has been suggested by Jones and Nisbett (1971). (Once again, the idea derives its importance in part from its implications for the attributional practices of psychologists themselves.) They propose that there is a general tendency for *actors* to attribute their

own actions to situational requirements but for *observers* of the same actions to explain them by reference to stable personal dispositions. Jones and Nisbett present their hypothesis as an "actuarial proposition" to which there are many exceptions. Nisbett and his colleagues (Nisbett, Caputo, Legant, & Marecek, in press) recently have presented further instances of the bias and, in a study by Storms (in press), have begun to test some of the suggestions advanced by Jones and Nisbett about the conditions under which the bias occurs.

These authors suggest that the bias is due to differences between actor and observer in the information available to them and in the salience of that information. This is interpretable within the analysis of variance conception of attribution in terms of differences between actor and observer in their respective information about possible causes and covariations. The person enacting a bit of behavior in a given situation ordinarily has more information about his own prior behavior in similar and related circumstances than does an observer of the present behavior. The actor is often keenly aware of ways in which the present behavior is inconsistent and distinctive, and out of this awareness he is inclined to make an external attribution for it. Also, because of the necessity that he respond appropriately to the present situation, the distinctive properties of that situation are highly salient to the actor. In contrast, the observer is relatively more informed about the person dimension of the analysis of variance cube—the ways in which the actor's behavior departs from that of other persons (including, perhaps, the observer's own) under similar conditions. This impression of uniqueness of the observed behavior provides the basis for the observer to attribute it to the actor.

This line of reasoning indicates when the proposed attributional differences between actor and observer will be absent or reversed. The reversal (person attribution by actor and situation attribution by observer) should occur for cases in which the actor feels his behavior to be both socially unique and consistent over time, but where the observer finds it to be totally inconsistent with past actions. This might characterize an aberrant sexual act that the actor finds highly consistent with his past and current behavioral fantasies (which he believes few other people share) but which an acquaintance finds to be totally out of keeping with the actor's prior behavior. Another condition un-

der which reversal of the actor-observer difference will occur relates to the differential salience of behavior versus situation for the actor and observer. A recent study by Storms (1972) shows that the actor-observer discrepancy can be reversed (the observer now making more situational attributions than the actor) by using videotape to provide the actor with the observer's view of the interaction, and the observer, with the actor's view.

The importance of the Jones and Nisbett proposition for the practice of psychology is perhaps obvious to the reader. Inasmuch as we psychologists are ordinarily observers of other persons' actions, we may tend to attribute too much causality to persons for the behavior we observe—too much causality to personality, ability, and other personal dispositions. This, of course, is to prejudge the difficult question of who is generally more accurate in his attributions—the actor or the observer. Nevertheless, the Jones and Nisbett hypothesis serves as an important warning to psychologists that the strength of their belief in and the intensity of their search for personal causes of behavior may not be justified entirely.

#### THE CONSEQUENCES OF ATTRIBUTIONS

The reader may be asking himself at this point why we are interested in causal attributions. After all, they are only subjective interpretations and inferences about what causes what. Of what importance are they in relation to what the person feels and does? It is appropriate to conclude with a mention of the many studies that show the role that attribution plays in relation to behavior. Some comment on this issue is especially important in view of where a recent article by Bem (1972) leaves the matter.

It must be emphasized first that attribution theory deals only with the processes by which attributions are derived from informational input. As Bem (1972) writes, "Self-perception theory [read "attribution theory"] can get us from the stimulus manipulation to the attribution. It cannot get us from the attribution to anything beyond that [p. 47]." But then the question arises of whether attributions are related to subsequent behavior and, if so, the nature of the relation. It is on the first of these questions that Bem's otherwise excellent article is seriously misleading, emphasizing as it does a handful of studies in which

there are discrepancies between the experimentally induced changes in attributions and the related behavioral effects. What is overlooked is a wide variety of studies in which the actions taken after two different sets of informational inputs correspond closely to the attributions that follow from those inputs.<sup>13</sup>

A few examples will serve to make the point. A long line of studies shows that the degree of anger and aggression expressed in reaction to another person's frustrating behavior is related to how much the available information about his behavior affords a basis for attributing it to him. Similarly, the reciprocation of benefit has been shown to depend on the attribution to the actor of his helpful behavior. (See Kelley, 1971, pp. 14-15, on these two points.) The effectiveness of a communication in changing opinions has been shown to depend on attributing it to the communicator's own beliefs rather than to external factors (Mills & Jellison, 1967; Steiner & Fields, 1960). Weiner and his colleagues present an impressive array of evidence to the effect that expectancies about future task performance depend on whether recent outcomes are attributed to stable or unstable factors, and that affective reactions to successful or unsuccessful task performance depend on whether it is attributed to internal or external causes (Weiner, Frieze, Kukla, Reed, Rest, & Rosenbaum, 1971). These authors, as well as Lanzetta and Hannah (1969), also show that reward and punishments are delivered to other persons according to the particular causal factor to which their task performance is attributed. Recent studies by Deci (1971) and Kruglanski, Friedman, and Zeevi (1971) suggest the role that self-attribution plays in relation to task persistence, creativity, and enjoyment.

In regard to the second question, concerning the nature of the relation between attributions and subsequent behavior, as the preceding examples suggest, the link seems to be characterized by reasonableness and plausibility. The person ordinarily takes actions appropriate to the meaning his causal interpretation gives to his own or others'

---

<sup>13</sup> Or, at least, the behaviors correspond to the respective attributions we would expect to follow from the two sets of inputs. The attributions are not always measured explicitly. Also, these studies generally illustrate very simple attributional ideas and provide little of interest about the determinants of various attributions.

behavior. The attribution process can be incorporated readily within a decision-making model of behavior, much as Weiner and his colleagues (1971) have for achievement situations. The causal attribution identifies the causes of certain effects and forms the basis for decisions about how to act in order to bring about the continuance or discontinuance of those effects. Beyond this, the theoretical statements are quite vague. As Bem (1972) aptly characterizes them, they are generally of the form "If one has managed to alter an individual's attitude or self-attribution, it is not unreasonable to expect that this will induce consequent changes in other response systems [p. 45]." I suspect Bem is quite correct in his belief that much of both practical and theoretical importance is concealed beneath these "it is not unreasonable to expect" statements.

There is much evidence, then, that attributions do matter. Man's concern with the reasons for events does not leave him "lost in thought" about those reasons. Rather, his causal explanations play an important role in providing his impetus to action and in his decisions among alternative courses of action. When the attributions are appropriate, the person undoubtedly fares better in his decisions and actions than he would in the absence of the causal analysis. It is not only true that (as Cicero tells us) "The causes of events always interest us more than the events themselves." It is undoubtedly also true (as another Roman, Virgil, records) that "Happy is he who has been able to perceive the causes of things."

#### REFERENCES

- BALDWIN, A. L., BALDWIN, C. P., CASTILLO-VALES, V., & SEEGMILLER, B. Cross-cultural similarities in the development of the concept of kindness. In W. W. Lambert & R. Weisbrod (Eds.), *Comparative perspectives on social psychology*. Boston: Little, Brown, 1971.
- BALDWIN, C. P., & BALDWIN, A. L. Children's judgments of kindness. *Child Development*, 1970, 41, 29-47.
- BAREFOOT, J. C., & STRAUB, R. B. Opportunity for information search and the effect of false heart-rate feedback. *Journal of Personality and Social Psychology*, 1971, 17, 154-157.
- BAVELAS, A., HASTORF, A. H., GROSS, A. E., & KITE, W. R. Experiments on the alteration of group structure. *Journal of Experimental Social Psychology*, 1965, 1, 55-70.
- 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.
- BROWN, R. W., & LENNEBERG, E. H. A study in language and cognition. *Journal of Abnormal and Social Psychology*, 1954, 49, 454-462.
- CHAPMAN, L. J., & CHAPMAN, J. P. Genesis of popular but erroneous psychodiagnostic observations. *Journal of Abnormal Psychology*, 1967, 72, 193-204.
- CHAPMAN, L. J., & CHAPMAN, J. P. Illusory correlation as an obstacle to the use of valid psychodiagnostic signs. *Journal of Abnormal Psychology*, 1969, 74, 271-280.
- COOPER, J., JONES, E. E., & TULLER, S. M. Attribution, dissonance, and the illusion of uniqueness. *Journal of Experimental Social Psychology*, 1972, 8, 45-57.
- DECI, E. Effects of externally mediated rewards on intrinsic motivation. *Journal of Personality and Social Psychology*, 1971, 18, 105-115.
- DESOTO, C. B., & KUETHE, J. L. Subjective probabilities of interpersonal relationships. *Journal of Abnormal and Social Psychology*, 1959, 59, 200-204.
- DUNCKER, K. On problem-solving. *Psychological Monographs*, 1945, 58(5, Whole No. 270).
- FESTINGER, L. A theory of social comparison processes. *Human Relations*, 1954, 7, 117-140.
- FRIEZE, I., & WEINER, B. Cue utilization and attributional judgments for success and failure. *Journal of Personality*, 1971, 39, 591-605.
- GERARD, H. B. Emotional uncertainty and social comparison. *Journal of Abnormal and Social Psychology*, 1963, 66, 568-573.
- HARE, A. P. *Handbook of small group research*. Glencoe, Ill.: Free Press, 1962.
- HASTORF, A. H., SCHNEIDER, D. J., & POLEFKA, J. *Person perception*. Reading, Mass.: Addison-Wesley, 1970.
- HEIDER, F. *The psychology of interpersonal relations*. New York: Wiley, 1958.
- HEIDER, F., & SIMMEL, M. An experimental study of apparent behavior. *American Journal of Psychology*, 1944, 57, 243-259.
- IRWIN, F. W., & SMITH, W. A. S. Further tests of theories of decision in an "expanded judgment" situation. *Journal of Experimental Psychology*, 1956, 52, 345-348.
- IRWIN, F. W., & SMITH, W. A. S. Value, cost, and information as determiners of decision. *Journal of Experimental Psychology*, 1957, 54, 229-232.
- IRWIN, F. W., SMITH, W. A. S., & MAYFIELD, J. F. Tests of two theories of decision in an "expanded judgment" situation. *Journal of Experimental Psychology*, 1956, 51, 261-268.
- JONES, E. E., & DAVIS, K. E. From acts to dispositions. 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., & 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*. New York: General Learning Press, 1972.
- JONES, E. E., & NISBETT, R. E. *The actor and the observer: Divergent perceptions of the causes of behavior*. New York: General Learning Press, 1971.

- JONES, E. E., WORCHEL, S., GOETHALS, G. R., & GRUMET, J. F. Prior expectancy and behavioral extremity as determinants of attitude attribution. *Journal of Experimental Social Psychology*, 1971, 7, 59-80.
- KELLEY, H. H. Attribution theory in social psychology. In D. Levine (Ed.), *Nebraska Symposium on Motivation*, 1967. Lincoln: University of Nebraska Press, 1967.
- KELLEY, H. H. *Attribution in social interaction*. New York: General Learning Press, 1971.
- KELLEY, H. H. *Causal schemata and the attribution process*. New York: General Learning Press, 1972.
- KELLEY, H. H., & STAHELSKI, A. J. The inference of intention from moves in the Prisoner's Dilemma Game. *Journal of Experimental Social Psychology*, 1970, 6, 401-419. (a)
- KELLEY, H. H., & STAHELSKI, A. J. Social interaction basis of cooperators' and competitors' beliefs about others. *Journal of Personality and Social Psychology*, 1970, 16, 66-91. (b)
- KELLEY, H. H., & THIBAUT, J. W. Group problem solving. In G. Lindzey & E. Aronson (Eds.), *The handbook of social psychology*. Vol. 4. Reading, Mass.: Addison-Wesley, 1969.
- KRUGLANSKI, A. W., FRIEDMAN, I., & ZEEVI, G. The effect of extrinsic incentive on some qualitative aspects of task performance. *Journal of Personality*, 1971, 39, 606-617.
- LANZETTA, J. T., & HANNAH, T. E. Reinforcing behavior of "naive" trainers. *Journal of Personality and Social Psychology*, 1969, 11, 245-252.
- 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.
- MICHOTTE, A. E. *The perception of causality*. New York: Basic Books, 1963.
- MILLS, J., & JELLISON, J. M. Effect on opinion change of how desirable the communication is to the audience the communicator addressed. *Journal of Personality and Social Psychology*, 1967, 6, 98-101.
- NISBETT, R. E., CAPUTO, C., LEGANT, P., & MARECEK, J. Behavior as seen by the actor and as seen by the observer. *Journal of Personality and Social Psychology*, in press.
- NISBETT, R. E., & VALINS, S. *Perceiving the causes of one's own behavior*. New York: General Learning Press, 1971.
- PETERSON, C. R., & BEACH, L. R. Man as an intuitive statistician. *Psychological Bulletin*, 1967, 68, 29-46.
- PIAGET, J. *The moral judgment of the child*. Glencoe, Ill.: Free Press, 1948.
- RODRIGUES, A. The biasing effect of agreement in balanced and imbalanced triads. *Journal of Personality*, 1968, 36, 138-153.
- SCHACHTER, S. *The psychology of affiliation*. Stanford: Stanford University Press, 1959.
- SCHACHTER, S., & SINGER, J. E. Cognitive, social and physiological determinants of emotional state. *Psychological Review*, 1962, 69, 379-399.
- STEINER, I. D., & FIELD, W. L. Role assignment and interpersonal influence. *Journal of Abnormal and Social Psychology*, 1960, 61, 239-246.
- STORMS, M. D. Videotape and the attribution process: Reversing actors' and observers' points of view. *Journal of Personality and Social Psychology*, in press.
- STRICKLAND, L. H. Surveillance and trust. *Journal of Personality*, 1958, 26, 200-215.
- 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. The belief in the law of small numbers. *Psychological Bulletin*, 1971, 76, 105-110.
- VALINS, S. Cognitive effects of false heart-rate feedback. *Journal of Personality and Social Psychology*, 1966, 4, 400-408.
- WEINER, B., FRIEZE, I., KUKLA, A., REED, L., REST, S., & ROSENBAUM, R. M. *Perceiving the causes of success and failure*. New York: General Learning Press, 1971.
- ZILLIG, M. Einstellung und Aussage. *Zeitschrift für psychologie*, 1928, 106, 58-106.