Part 1

Attitude Theory and Measurement

Chapter 2

Theories
of
Attitude

The preceding chapter provided conceptual definitions of belief, attitude, inten-tion, and behavior, as well as a brief outline of a theoretical network linking these concepts. In the present chapter we will consider alternative formulations by re-viewing some contemporary theories of attitude. We will provide a brief descrip-tion of each theory, using its original terminology, and we will then attempt to identify its basic constructs and their interrelations in terms of the conceptual framework outlined in Chapter 1. We will thus examine the implications of each theory for an understanding of the relations between beliefs, attitudes, intentions, and behavior.

Most contemporary attitude theories have their origins in two major schools of thought that have shaped theory and research in social psychology. Whereas the various learning theories of attitude are based on the stimulus-response ap-proach of behavior theory, most theories of cognitive consistency are influenced by the cognitive approach of field theory. A distinction is therefore usually made between behavior theories of attitude and cognitive consistency theories (e.g., Kiesler, Collins, and Miller, 1969; Fishbein, 1967a; Greenwald, Brock, and Ostrom, 1968). This classification into behavior-versus-consistency theories, how-ever, blurs the distinction between a theory's theoretical origin and the phenomena it deals with. For example, Osgood and Tannenbaum's (1955) congruity principle is typically viewed as a consistency theory (since it deals with attitudinal con-sistency or congruity) although it originated within the behavior-theory tradition. The present review de-emphasizes the distinction between behavior and consis-tency theories in favor of a more unified presentation.

LEARNING THEORIES

Several investigators have used principles taken from the learning theories of Hull (1943, 1951), Spence (1956), and Tolman (1932) to study the acquisition of beliefs and attitudes. Generally speaking, these learning theories are concerned with the processes whereby a given response becomes associated with (or conditioned to) a given stimulus. Most learning is explained in terms of two basic conditioning paradigms: classical conditioning and operant or instrumental conditioning.

Conditioning Principles

An *unconditioned stimulus* (UCS) elicits automatically, without prior learning, one or more overt *unconditioned responses* (UCR). For example, an unexpected loud noise produces a startle response; a bottle in an infant's mouth produces sucking, salivation, and swallowing; an electric shock or other painful stimulus leads to various withdrawal responses. The classical conditioning paradigm starts with an unconditioned stimulus that is always followed by some characteristic unconditioned response. Now consider a new stimulus that does not initially elicit the unconditioned response although it may elicit some other response (R_r). When this new *conditioned stimulus* (CS) is consistently paired with the unconditioned stimulus, it ultimately comes to elicit some of the response characteristics previously produced only by the unconditioned stimulus. That is, the CS by itself now elicits the UCR. When an initially neutral stimulus (the CS) acquires the ability to elicit a response (the UCR) originally elicited only in the presence of another stimulus (the UCS), learning is said to have occurred.[1] Figure 2.1 provides a schematic representation of the classical conditioning paradigm.

The solid lines in Fig. 2.1 represent either innate, nonlearned associations or associations that have been previously learned. That is, once the CS comes to elicit the UCR consistently, it can serve as the UCS in another conditioning situation. This process is known as higher-order conditioning.


CS   = conditioned stimulus
UCS   = unconditioned stimulus
UCR   = unconditioned response
R_x = other response
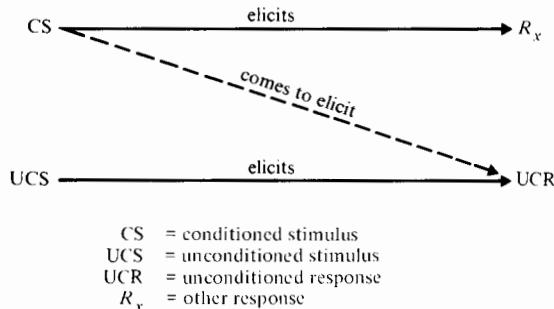
Fig. 2.1 The classical conditioning paradigm.

---

1. The decision to label a given event as a stimulus or a response is somewhat arbitrary, since any response can also serve as a stimulus for some other response.

As an example of classical conditioning, consider a child who always cries (UCR) at the sight of a spider (UCS). Classical conditioning occurs when the child hears the word "spider" (CS) being uttered consistently in the presence of the spider. After several such CS-UCS pairings, the child starts to cry whenever he hears the word "spider," even when no spider is actually present.

Various factors have been found to influence classical conditioning. Among the most important are the frequency with which the CS and UCS are paired (the greater the number of pairings, the more the response to the CS resembles the response to the UCS) and the temporal relation of the CS and UCS. When the UCS precedes the CS, little learning is evidenced; maximal learning seems to occur when the CS precedes the UCS by a short time interval, such as 0.5 second. (For a review of this research, see Kimble, 1961.)

In classical conditioning, then, the response to be learned is initially elicited by the unconditioned stimulus. In contrast, operant conditioning (or trial-and-error learning) involves a situation in which the organism initially emits a variety of different responses. One of these responses (R_x) is reinforced; i.e., the response R_x is *instrumental* to obtaining some reward or avoiding some punishment. The probability of the recurrence of the reinforced response increases with each reinforced trial, and the response is said to be learned when it occurs with high probability. The instrumental conditioning paradigm is illustrated in Fig. 2.2.
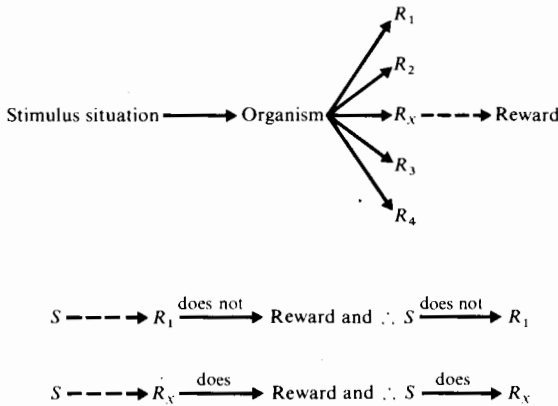
Fig. 2.2 Operant or instrumental conditioning.

As an example of instrumental conditioning, a mother gives her child a piece of chocolate every time he picks up his toys (R_x) but not when he cries, demands chocolate, or throws objects at his brother. The reinforcer (chocolate) will thus strengthen R_x, and the child will learn to pick up his toys.

Factors that have been found to influence instrumental conditioning include frequency of reinforcement (number of times the response is followed by reward), temporal relation between the response and the reinforcement, schedules of reinforcement, and magnitude of the reinforcer (cf. Kimble, 1961).

When a given response reduces the drive state by leading to appropriate reward or by enabling the organism to avoid punishment, the response is said to be reinforced, and the reward is known as a reinforcer. A distinction is made between *primary* and *secondary* (or learned) reinforcers. Primary reinforcers are rewards that are unlearned reducers of drive states (e.g., food, chocolate, water), and secondary reinforcers are previously neutral stimuli that acquire reinforcement properties because they have been associated with primary reinforcers. That is, just as a CS comes to elicit UCR, a stimulus that is consistently paired with a reward will take on some of the reinforcing properties of the reward itself.

Implicit versus explicit responses. In the discussion above, we have considered only observable stimuli and responses. The same processes are assumed to occur with implicit (nonobservable) responses or stimuli. For example, Hull (1951) referred to a "fractional antedating goal reaction," a (usually covert) portion of the overt goal response that antedates, anticipates, or mediates the overt reaction. Osgood, Suci, and Tannenbaum (1957) describe this process with reference to the classical conditioning of a buzzer (CS) to electric shock (UCS):

> Many experiments on the details of the conditioning process combine to support the following conclusion: Components of the total unconditioned reaction vary in their dependence on the unconditioned stimulus and hence in the ease with which they may become conditioned to another stimulus. Typically, the less energy-expending a reaction component (e.g., "light-weight" components like glandular changes and minimal postural adjustments) and the less interfering a reaction component with on-going overt behavior (e.g., components which do not hinder overt approaches, avoidances, manipulations, and the like), the more promptly it appears in the conditioned reaction, and hence the more readily available it is for the mediation function. The argument thus far may be summarized as follows: *Whenever some stimulus other than the [UCS] is contiguous with the [UCS], it will acquire an increment of association with some portion of the total behavior elicited by the [UCS] as a representational mediation process.* As diagrammed [in Fig. 2.3], this stimulus-producing process ($r_m \rightarrow s_m$) is *representational* because it is part of the same behavior (R_T) produced by the [UCS] itself—thus the buzzer becomes a sign [CS] of shock [UCS] rather than a sign of any of a multitude of other things. It is *mediational* because the self-stimulation (s_m) produced by making this short-circuited reaction can now become associated with a variety of instrumental acts (R_x) which "take account of" the [UCS]—the anxiety state generated by the buzzer may serve as a cue for leaping, running, turning a rachet, or some other response sequence which eliminates the signified shock. (Osgood, Suci, and Tannenbaum, 1957. p. 6).

Learning Theories of Attitudes

In one of the first applications of learning theory to the attitude area, Leonard Doob (1947) defined attitude as a learned, implicit anticipatory response. That is, he viewed attitude as an unobservable response to an object that occurs prior to, or in the absence of, any overt response. Osgood, Suci, and Tannenbaum
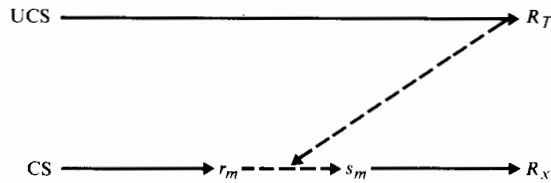
Fig. 2.3 Conditioning of an implicit meaning response. (Adapted from Osgood, Suci, and Tennenbaum, 1957.)

(1957) argued that the implicit mediating response represents the "meaning" of the object, and they suggested that attitude refers only to the *evaluative* part of the total meaning response. Osgood's view of attitude as a mediating evaluative response has met with general acceptance by theorists working within the behavior-theory tradition (e.g., Staats and Staats, 1958; Rhine, 1958; Fishbein, 1967c).

Most learning theories of attitude are concerned with the ways in which attitudes are acquired, i.e., how implicit (evaluative) responses become associated with a given stimulus object. For example, consider a child who has frequently eaten M & M candies. The stimulation involved produces overt responses, such as sucking, salivating, swallowing, etc. In addition, an implicit response with a positive evaluative component has occurred prior to, or in conjunction with, the overt responses. According to the mediational conditioning principle, there will be a tendency for this implicit response to become associated with the candies themselves; i.e., the child develops a favorable attitude toward M & M candies. Further, once this new association has been learned, any other stimulus frequently paired with the M & M candies will also tend to elicit the positive mediating response. Thus, according to the principle of higher-order conditioning, if the M & M's are always dispensed by the child's uncle, a positive attitude toward the uncle should develop.

Primary and higher-order classical conditioning are the basic paradigms in the approach taken by Staats (1968) and Staats and Staats (1958). In addition to classical conditioning, Lott and Lott (1968) have emphasized instrumental conditioning as a basis for attitude formation. That is, "a person who experiences reinforcement or reward for some behavior will react to the reward, i.e., will perform some observable or covert goal response" (p. 68). As indicated above, this covert goal response is often viewed as an attitude. Consistent with the principle of classical conditioning, this implicit response becomes conditioned to all discriminable stimuli present at the time of reinforcement.

Lott (1955) also made use of a second major principle, mediated generalization. According to this principle, once some overt response and the implicit reaction are associated with a given stimulus, any other stimulus that elicits the same mediating reaction will also come to produce the overt response. To return to our previous example, a child who has been rewarded for approaching M & M candies will also tend to approach his uncle, who in the past has given him M & M's.

Doob (1947) emphasized that a person first learns an implicit mediating response (i.e., attitude) to a given stimulus, and he must then also learn to make a specific overt response to the attitude. The first process can be accounted for by classical conditioning, the second by instrumental learning. For other similar stimuli to elicit the same response, principles of generalization are invoked. Thus, according to Doob, the entire mediating response is the attitude. He distinguished between this attitudinal response and other mediating responses (e.g., habit) by reserving the term "attitude" for those implicit responses that are elicited by socially relevant stimuli. In contrast, recall that Osgood, Suci, and Tannenbaum (1957) viewed the total mediating response as representing the meaning of a stimulus and that they defined attitude as only the evaluative part of meaning. Thus Doob argued that people with the same attitude may learn to behave differently, whereas Osgood *et al.* further accounted for low attitude-behavior relations by pointing out that attitude is only a part of the total implicit response, and thus two people with the same attitude toward a given stimulus may differ on other dimensions of meaning vis-à-vis the stimulus.

The discussion above has been concerned primarily with the conditioning of an implicit (evaluative) response to one or more stimuli. The next question to be raised concerns the effects of combining two or more stimuli, each of which elicits a *different* implicit reaction, into a stimulus complex. In fact, most stimulus objects can be considered as representing a complex array of stimuli, and different implicit evaluative reactions may have been conditioned to the different component stimuli. Imagine, for example, that for some person a favorable mediating response has been conditioned to *athlete* and an unfavorable mediating response to' *lazy*. According to most learning approaches, the implicit evaluative response to the stimulus complex *lazy athlete* is some function of the evaluative reactions elicited by the component stimuli.

The Congruity Principle

Although learning theory does not specify the exact combinatorial principle involved, Osgood and Tannenbaum (1955) have proposed the congruity principle to describe this process. According to the congruity principle, whenever two stimuli are combined, the "mediating reaction characteristic of each shifts toward congruence with that characteristic of the other, the magnitude of the shift being inversely proportional to the intensities of the interacting reactions" (Osgood, Suci, and Tannenbaum, 1957, p. 201).

Assume that implicit evaluative reactions can be measured on a 7-point scale ranging from +3 (favorable) through 0 (neutral) to −3 (unfavorable). According to the congruity principle, the more intense or polarized the evaluative reaction (i.e., the greater its distance from the neutral point), the less that evaluation will shift toward the other. Indeed, the amount of shift is assumed to be inversely proportional to the degree of polarization. Algebraically, the point of resolution

(i.e., the evaluative reaction to the complex stimulus) can be predicted from Eq. 2.1 (Osgood, Suci, and Tannenbaum, 1957, p. 207),

$$P_R = \frac{|p_1|}{|p_1| + |p_2|}\, p_1 + \frac{|p_2|}{|p_1| + |p_2|}\, p_2, \qquad (2.1)$$

where P_R is the point of resolution, p_1 and p_2 are the evaluations of the component stimuli and $|p_1|$ and $|p_2|$ are the absolute values of these evaluations, i.e., the polarities of the component stimuli.

For example, if *athlete* had a value of $+1$ and *lazy* a value of -2, the point of resolution for *lazy athlete* would be -1, as shown by the following computation.

$$P_R = \frac{1}{1+2}\, (1) + \frac{2}{1+2}\, (-2) = \frac{1}{3} + \frac{(-4)}{3} = -1.$$

We will show below that the congruity principle has also been applied to other attitudinal phenomena, and in such applications, the principle is quite similar to a number of so-called consistency theories.

Concept Formation

Theories of concept formation are also relevant for an understanding of the ways in which mediating responses become associated with complex stimulus objects. According to a learning-theory approach, concept formation involves the conditioning or learning of a common response to a set of discrete stimuli. For example, the response "vegetable" becomes conditioned to (associated with) a variety of stimuli, such as beets, spinach, carrots, etc. As mentioned above, the implicit evaluative reaction elicited by the concept "vegetable" should be some function of the evaluative reactions associated with the stimuli that elicit this concept, i.e., the evaluative reactions associated with beets, carrots, spinach, etc. Thus, as Fishbein (1967c) has argued, whenever a new concept is learned, an attitude toward that concept is acquired simultaneously. Once a concept has been learned, however, new stimuli may be associated with it, and the mediating evaluative reactions elicited by these new stimuli will also become conditioned to the concept and change the attitude toward it. At any point in time, a given stimulus object (i.e., a concept such as "vegetable") will elicit a large number of responses, some of which correspond to the stimuli that originally defined the concept (e.g., carrots, beets, spinach), and some of which may have been acquired at a later stage (e.g., edible, nourishing).

Recall that each of these associated objects (e.g., beets, nourishing) elicits an implicit evaluative reaction and that the final attitude toward the concept ("vegetable") is some function of all these evaluative reactions. Again, however, learning theory does not specify the exact combinatorial principle involved.

A Model of the Relationship between Beliefs and Attitudes

Fishbein (1963, 1967c) has proposed a model that deals with the ways in which evaluative mediating responses combine to produce the overall attitude. Accord-

ing to this model, a given stimulus object may elicit a variety of responses that refer to the characteristics, attributes, or qualities of the object. It is assumed that these stimulus-response associations are learned through conditioning processes; the strength of an association should thus be a function of the number of conditioning trials. The different responses to the object are viewed as constituting a "habit-family hierarchy," in which the responses are ordered in terms of the probability that they will be elicited by the stimulus object, i.e., in terms of the strength of their association with the stimulus object.

Further, an implicit evaluative reaction is associated with each of the responses in the hierarchy. It is assumed that each evaluative reaction is conditioned to the stimulus object in direct proportion to the strength of the association between the stimulus object and the corresponding responses in the hierarchy. Thus, the lower the position of a response in the habit-family hierarchy, the less the evaluative reaction associated with it will contribute to the overall attitude toward the object. A further assumption is that evaluative mediating responses combine in an additive manner, and the overall attitude toward the object is therefore viewed as a weighted sum of all implicit evaluative reactions conditioned to the object.

Analysis of Learning Theories

Let us now examine the ideas that have been discussed in terms of the conceptual framework presented in Chapter 1. Throughout the discussion above, attitude has been viewed as an implicit, mediating response. Although some investigators leave the exact nature of this implicit response unspecified (e.g., Doob, 1947; Lott and Lott, 1968), most theorists would agree that attitude can best be viewed as an evaluative mediating response. This conception, therefore, is very similar to our definition of attitude as a person's location on a bipolar evaluative or affective dimension with respect to some object. The implicit evaluative reaction, i.e., attitude, is viewed as predisposing the individual to perform various overt behaviors. Thus the individual may be said to hold various behavioral intentions. However, Doob (1947) has made it clear that any particular response will be performed only to the extent that it has been positively reinforced. Thus, consistent with our conceptual framework, two persons may hold the same attitude but learn to perform different responses.

In discussing the acquisition of attitudes, the various learning theories make reference to stimulus-response conditioning processes. It may be argued that stimulus-response bonds established in this manner correspond to what we have called beliefs. We defined beliefs in terms of the probability that a given object is related to some attribute, i.e., to some other object, concept, or goal. If the object is now viewed as a stimulus and the related attribute as a response, a belief about an object corresponds to the probability that the stimulus elicits the response, i.e., to the strength of the stimulus-response association. Indeed, Tolman (1932) explicitly viewed "cognitions" as "expectancies" or subjective probabilities that one event is associated with (or follows from) some other event.

One important implication of these considerations is that, according to a behavior-theory approach, belief formation should follow the laws of learning. Whenever a belief is formed, some of the implicit evaluation associated with the response becomes conditioned to the stimulus object. The implicit evaluation associated with a response constitutes an attitude which may have been formed as the result of prior conditioning. The implication of this conditioning paradigm is that attitude toward an object is related to beliefs about the object.

Fishbein (1963) has made this relationship an explicit part of his theory of attitude, which can be described as follows: (1) An individual holds many beliefs about a given object; i.e., the object may be seen as related to various attributes, such as other objects, characteristics, goals, etc. (2) Associated with each of the attributes is an implicit evaluative response, i.e., an attitude. (3) Through conditioning, the evaluative responses are associated with the attitude object. (4) The conditioned evaluative responses summate, and thus (5) on future occasions the attitude object will elicit this summated evaluative response, i.e., the overall attitude.

According to the theory, a person's attitude toward any object is a function of his beliefs about the object and the implicit evaluative responses associated with those beliefs. The central equation of the theory can be expressed as follows:

$$A_o = \sum_{i=1}^{n} b_i e_i, \tag{1.2}$$

where A_o is the attitude toward some object, O; b_i is the belief i about O, i.e., the subjective probability that O is related to attribute i; e_i is the evaluation of attribute i; and n is the number of beliefs.

Consider, for example, a person's attitude toward the supersonic transport (SST). Assume that he holds the following beliefs: (1) SST is an airplane; (2) SST is noisy; (3) SST is not economical; and (4) SST is a pollutant. According to Fishbein's model, his attitude toward the SST is a function of the strength with which he holds these beliefs (i.e., his subjective probability that the SST is related to the different attributes) and of his evaluations of each attribute. Table 2.1

Table 2.1 Hypothetical Attitude toward SST

Belief	b	e	be
Airplane	.90	+2	1.80
Noisy	.80	−2	−1.60
Not economical	.60	−1	−0.60
Pollutant	.50	−3	−1.50

$A_o = \sum b_i e_i = -1.90$

presents subjective probabilities and evaluations that might have been obtained.[2] Note that this person is predicted to hold a negative attitude toward the SST.

This analysis of the relations between beliefs and attitude is consistent with our conceptual framework, which also indicates that a person's attitude toward some object is a function of his beliefs about the object (see Chapter 1).

EXPECTANCY-VALUE THEORIES

Fishbein's model is concerned with the relations of beliefs to attitudes, and it is of interest to note that other theorists have arrived at similar formulations in attempts to account for overt behavior. The theories presented by Tolman (1932), Rotter (1954), Atkinson (1957), and others may be viewed in this light.[3] We have already seen that, according to Tolman (1932), people learn "expectations," i.e., beliefs that a given response will be followed by some event. Since these "events" could be either positive or negative "reinforcers" (i.e., could have positive or negative valence), his argument, essentially, was that people would learn to perform (or increase their probability of performing) behavior that they "expected" to lead to positively valenced events.[4]

Perhaps the best known expectancy-value model is the subjective expected utility (SEU) model of behavioral decision theory (Edwards, 1954). According to this theory, when a person has to make a behavioral choice, he will select that alternative which has the highest subjective expected utility, i.e., the alternative which is likely to lead to the most favorable outcomes. The subjective expected utility of a given alternative is defined in Eq. 2.3,

$$SEU = \sum_{i=1}^{n} SP_i U_i, \tag{2.3}$$

where SEU is the subjective expected utility associated with a given alternative; SP_i is the subjective probability that the choice of this alternative will lead to some outcome i; U_i is the subjective value or utility of outcome i; and n is the number of relevant outcomes.[5]

In our terminology, this model deals with beliefs about the consequences of performing a given behavior ($SP_i \sim b_i$) and with the evaluations associated with the different outcomes ($U_i \sim e_i$). Thus SEU can be reinterpreted as the per-

---

2. Discussions of appropriate measurement procedures will be found in Chapter 3.

3. See Feather (1959) for a comparison of some of these theories.

4. This is a grossly oversimplified statement of Tolman's position. For a complete discussion, see Tolman (1932).

5. Since n usually refers to a mutually exclusive and exhaustive set of outcomes, $\Sigma SP_i = 1.00$ in most behavioral-decision-theory analyses.

son's attitude toward the behavior (A_B), and Eq. 2.3 can be rewritten as follows:

$$A_B = \sum_{i=1}^{n} b_i e_i. \qquad (2.4)$$

Note that, whereas the SEU model appears to assume a direct link between SEU and behavior, no direct relation between A_B and behavior is assumed. This question will be raised again in a later chapter dealing with the prediction of intentions.

An Instrumentality-Value Model

Rosenberg (1956) was perhaps the first to introduce an explicit expectancy-value model in the attitude area. He defined attitude as a *"relatively stable affective response to an object"* and argued that this attitude is "accompanied by a *cognitive structure* made up of beliefs about the potentialities of that object for attaining or blocking the realization of valued states" (p. 367). According to Rosenberg (1956), the more a given "object" (i.e., an action or policy) was *instrumental* to obtaining positively valued goals (or consequences) and to blocking (or preventing) negatively valued goals, the more favorable the person's attitude toward the object. This hypothesis is expressed in Eq. 2.5,

$$A_o = \sum_{i=1}^{n} I_i V_i, \qquad (2.5)$$

where I_i is instrumentality, i.e., the probability that o would lead to or block the attainment of a goal or value i; V_i is value importance, i.e., the degree of satisfaction or dissatisfaction the person would experience if he obtained value i; and n is the number of goals or value states.

Note that this equation is very similar to Fishbein's (1963) model (see Eq. 2.2) and to the SEU model. That is, Rosenberg's model also deals with beliefs about the object and with associated evaluations or values.

It is interesting that, whereas Fishbein's model was developed within the framework of behavior theory, Rosenberg's formulation was influenced by what today is called a functional approach to attitudes. This approach suggests that attitude formation and change can be understood only in terms of the functions that attitudes serve for the individual. For example, Smith, Bruner, and White (1956) discussed three functions: object appraisal, social adjustment, and externalization. Katz (1960) mentioned the instrumental, adjustive, or utilitarian function, the ego-defensive function, the value-expressive function, and the knowledge function. According to this view, attitudes are necessary because they permit the individual to achieve certain goals or value states (e.g., they allow him to organize knowledge, to maintain his self-esteem, to express his views). Rosenberg's (1956) initial formulation can be viewed as being concerned with the extent to which an object facilitates or hinders the attainment of such valued goals.

In his later theorizing, Rosenberg (1960, 1965a) expanded his definition of attitude by including beliefs within the attitude concept. This expansion was ac-

companied by an explicit statement of affective-cognitive consistency. Specifically, he argued that "humans have a need to achieve and maintain affective-cognitive consistency" (Rosenberg, 1965a, pp. 123–124). It is worth noting that Fishbein's (1963) model accounts for the relation between beliefs and attitude in terms of conditioning processes, whereas Rosenberg's (1956) formulation relies on the assumption of a need for cognitive-affective consistency to account for the same relations. Nevertheless, the two models have considerable structural similarities, and the basic hypothesis of each can be described by the same algebraic expression, such as Eq. 2.2.

Rosenberg's (1960, 1965a) theory of cognitive-affective consistency is one of a number of theories dealing with the effects of inconsistencies among beliefs, attitudes, intentions, and behaviors. The origin of these consistency theories can in large part be traced to Fritz Heider's (1944, 1946, 1958) principle of balance.

BALANCE THEORY

Heider's concern with balanced configurations grew out of his interest in the factors that influence causal attribution of an event to a person. As we shall see below in our discussion of attribution theory, many factors may influence causal attributions. One conclusion arrived at by Heider is that "if the attitudes toward a person and event are similar, the event is easily ascribed to the person." He further argued that "a balanced configuration exists if the attitudes toward the parts of a causal unit are similar" (Heider, 1946, p. 107). That is, a balanced state exists when the two entities composing a unit have the same "dynamic character," in other words, when the person's attitudes or sentiments vis-à-vis the two entities are both positive or both negative.

Consider, for example, a person who attributes responsibility to the President for the fact that his son was drafted. A balanced state exists when the person likes the President and approves of the fact that his son was drafted, or when he dislikes the President and disapproves of this fact. When he has a positive attitude toward one element (e.g., the President) but a negative attitude toward the other (e.g., the drafting of his son), a state of imbalance is said to exist.

According to Heider's model, balance also exists if the person holds different attitudes toward the two elements and perceives that one element has not been caused by the other. Thus, if in the above example the person liked the President, disapproved of the fact that his son was drafted, but perceived that the President was not directly responsible for the drafting of his son, a balanced state would exist.

Note that Heider takes a phenomenological approach; that is, he is concerned with a person's *perceptions* of the relationships between elements. Three basic elements are usually involved: the focal person (p), another person (o), and an object or event (x). In the model discussed above, p likes (L), or dislikes ($\bar{\text{L}}$) o and x, and perceives a causal unit relation (U) between o and x. Heider was aware, however, that two elements may be perceived to form a unit on the

basis of processes other than causal attribution. The Gestalt principles of perception suggested that o and x might also be perceived to form a unit on the basis of similarity, proximity, membership, possession, or belonging. The unit oUx can mean, for instance, o owns x or o made x. The segregation of o and x, that is, $o\overline{U}x$, can mean o does not own x, etc. Other examples of unit relations are: o is familiar with, used, or knows x.

Thus Heider (1946) was able to generalize his balance principle to all unit relations. A balanced state is said to exist if p has similar attitudes toward the two elements of the unit oUx (that is, pLo and pLx or $p\overline{L}o$ and $p\overline{L}x$) and if he holds different attitudes toward two segregated elements $o\overline{U}x$.

Heider (1946) further extended his theory to allow perceived liking relations between o and x (that is, oLx and $o\overline{L}x$). In addition, he included unit relations between p and o (for example, p and o are brothers) and between p and x (for example, p owns x). Finally, he pointed out that more than one relation may exist between two entities. For example, p may own and like x, or p may respect but dislike o.

The basic balance principle can now be stated as follows: "In the case of two entities, a balanced state exists if the relation between them is positive (or negative) in all respects, i.e., in regard to all meanings of L and U. . . . In the case of three entities, a balanced state exists if all three relations are positive in all respects, or if two are negative and one positive" (Heider, 1946, pp. 110–111). Balanced and imbalanced triadic configurations are illustrated in Fig. 2.4.

Although Heider presents the balance principle in this general form, he draws attention to some of the problems that may be involved in generalizing from causal

Balanced states   Imbalanced states
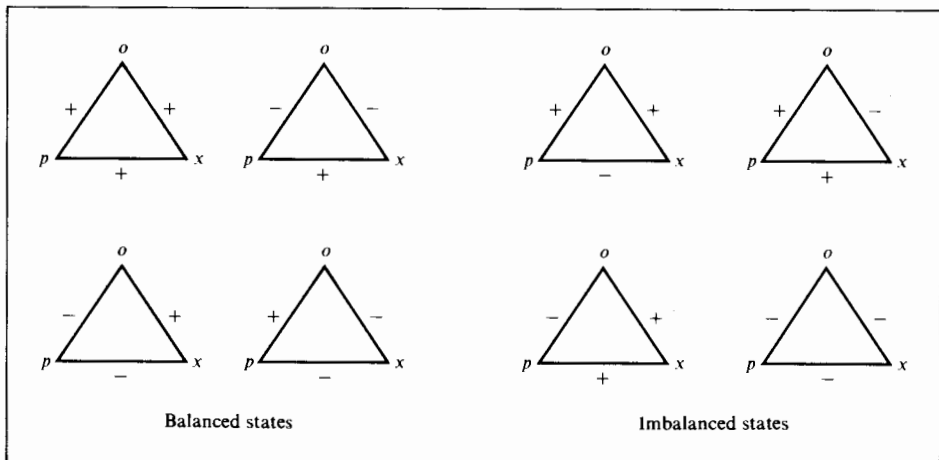
Fig. 2.4 Balanced and imbalanced triads. All triads reflect p's perspective: Lines between elements represent either unit or sentiment relations; the + and − signs stand for positive and negative relations, respectively.

units to other unit relations and to perceived liking relations. A case in point involves the unit relation of ownership. At first glance, the triad pLo, oUx, and pLx appears balanced. However, o owns x sometimes implies that p cannot own x. Since p likes x, the situation is imbalanced (i.e., in a dyad, $p\bar{U}x$ and pLx are imbalanced). Heider refers to this case as an example of envy and provides a similar explanation for jealousy.

A basic dynamic principle underlies balance theory, namely, that liking and unit relations tend toward a balanced state. By a balanced state is meant a situation in which the relations among the entities fit together harmoniously; there is no stress toward change. Further, if a balanced state does not exist, then forces toward such a state will arise. If a change is not possible, the state of imbalance will produce tension (Heider, 1958, p. 201).[6] When changes are possible, either the dynamic characters will change—that is, p's attitudes toward o or x may change—or the unit relations will be changed through action or through cognitive reorganization (Heider, 1946, pp. 107–108).

Analysis of Balance Theory

Having described balance theory using Heider's terminology, we will again examine the major ideas in terms of our conceptual framework. The primary concern of balance theory is with the dynamic interactions among beliefs and attitudes. A person's perception of a relation between o and x corresponds to our definition of a belief about o. Thus the person may believe that o likes movies, that o owns a typewriter, or that o is married to some other person. Unlike our probabilistic conception of belief, however, in Heider's system beliefs are dichotomous; i.e., the person either believes that o is related to x or that o is not related to x.[7]

In addition to dealing with beliefs about an object, balance theory considers a person's beliefs about himself. Thus the person may believe that he is engaged to

6. Although this notion of tension has often been interpreted as a "need" for balance, Zajonc (1968a) has argued that "the dynamic principle of change proposed by Heider does not involve psychological forces of overwhelming strength. They are more akin to *preferences* than to driving forces. There is no anxiety when structures are imbalanced; imbalanced states are not noxious; a compelling need to strive for balance is not assumed. Forces toward balance have the same character as Gestalt forces toward 'good figures' in the perception of forms." (p. 341)

7. Note further that negative associations (that is, $\bar{L}$ and $\bar{U}$) are not permissible in probabilistic or behavior theory terminology. That is, in these frameworks, one has only positive associations. Thus, whereas balance theory views the statement "o is not x" as $(o)\ \bar{U}\ (x)$, behavior theory and probability models view it as (o) is (not x). As we shall see below, assigning a negative sign to the evaluation of (x) is not the same as obtaining an evaluation of (not x), and thus expectancy-value models will lead to different predictions when a statement such as (o is not x) is treated as a negative association between (o) and (x) than when it is treated as a positive association between (o) and (not x).

another person (pUo), that he performed some behavior (pUx), or that he is not responsible for some event $(p\overline{U}x)$.

The balance model also deals with a person's attitudes. Heider's conception of an attitude or sentiment, like our own, concerns positive or negative evaluation. Again, however, whereas we consider a bipolar *dimension* of affect, balance theory is restricted to positive or negative attitudes. That is, a person either likes or dislikes some other entity.

It is of interest to note that when the perceived o-x relation is reinterpreted as a person's belief about o, his attitude toward x is equivalent to the evaluation of the attribute x associated with o. Thus one implication of the balance principle is that a person's attitude toward an object (that is, p's attitude toward o) may be influenced by his belief about the object and by the evaluation of the related attribute. We saw above that the same conclusion [that is, $A_o = f(b,e)$] can be derived from a number of different theoretical approaches. However, balance theory has implications that go beyond this conclusion. For example, p's belief about o may be a function of his attitudes toward o and x, and his attitude toward x may be influenced by his belief about o and his attitude toward o. Further, since all three relations in a triad may be unit relations, the theory has implications for the interactions among beliefs. For instance, consider the case in which p buys a car and then finds out that his wife hates it. This triad can be described as pUo, pUx, and $o\overline{L}x$. Since the configuration is imbalanced, there will be a tendency for at least one of the beliefs to change.

An important and frequently neglected aspect of Heider's balance model is that beliefs may change as a result of some action taken by the person. In the triad above, for example, p may sell his car; that action will lead to the belief that he no longer owns it and thereby restores balance. The notion that a person's actions are represented cognitively in the form of beliefs and may thus influence other beliefs or attitudes plays a crucial role in Festinger's (1957) dissonance theory, which will be discussed below.

In sum, then, balance theory is very rich in its implications. Not only does it deal with the relations between beliefs and attitudes; it has important implications for belief formation and the relations between beliefs. Further, since a person's behaviors may be represented as beliefs, the balance model suggests that a person's beliefs and attitudes may be influenced by his behavior.

One should realize, however, that Heider's original balance theory has certain limitations. As pointed out earlier, the theory deals only with qualitative relations between entities. Thus the relation between two entities is either positive or negative; beliefs and attitudes do not vary in degree. Another limitation is the fact that the theory deals with the relations between a maximum of three entities. A third limitation is that, although Heider discusses the possibility of multiple relations between two entities, he says nothing about the degree of balance that would exist in such complex dyadic or triadic configurations.

A number of theorists have attempted to overcome some of these limitations. For example, Cartwright and Harary (1956) employed principles of mathemati-

cal graph theory to deal with multiple entities and multiple relations between them. Although these relations are not quantified, graph theory provides a means for assessing the degree of balance in a given configuration. Cartwright and Harary applied their model to actual relations within social structures, thereby generalizing the balance principle beyond a given person's perceptions. Similarly, Newcomb (1953) applied Heider's balance model to actual structural relations between two persons. Finally, Abelson and Rosenberg (1958) used concepts of matrix algebra to extend the balance principle beyond three entities and, like Cartwright and Harary (1956), they suggested ways of estimating degrees of balance or imbalance within a given structure. For some other extensions of Heider's balance model, see Feather (1964a, 1971a).

None of these extensions of Heider's balance model has attempted to quantify beliefs or attitudes. Such quantifications, however, are part of other consistency theories. For example, Rosenberg's (1960, 1965a) affective-cognitive consistency model discussed earlier not only quantifies beliefs and attitudes but also deals with multiple beliefs about the attitude object. Similarly, the congruity principle mentioned above has been applied to situations of interest to the balance model. When this is done, the congruity formulation is structurally similar to the balance model, but in addition, it provides quantification of some of the relations involved.

THE CONGRUITY PRINCIPLE

Just as balance theory begins with a consideration of unit formation, Osgood and Tannenbaum's (1955) congruity principle has as its starting point an assertion that links two objects of judgment. The simplest assertion is merely a descriptive statement, such as "athletes are lazy" or "cigarettes contain nicotine." A more complex situation is that in which a source makes an assertion about a concept, e.g., "communists like strong labor unions" or "the President favors medicare" (Osgood and Tannenbaum, 1955). These assertions, or "coupling actions," may be either associative (*favors, are*) or dissociative (*opposes, are not*).[8]

Although positive or negative assertions may appear to be equivalent to positive or negative relations between *o* and *x* in the balance model, there is an important difference. An associative or dissociative assertion is treated as a given in the congruity model, whereas balance theory is concerned with *p*'s *perception* of the relation between *o* and *x*. According to the congruity principle, if *p* encounters an assertion such as "the President favors medicare," then the President and medicare are positively related. In contrast, according to the balance model, even if such an assertion was made, a positive relation between the President and medicare would obtain only if *p* actually believed the assertion. If *p* disbelieved the assertion, i.e., if he believed the President to oppose medicare, a negative relation be-

---

8. Note again that within a behavior theory approach, dissociative assertions would not be permissible.

tween the President and medicare would be entered in the balance model. Thus, whereas balance theory is concerned with *perceived* relations between *o* and *x*, the congruity principle deals with "objective" or stated relations, i.e., assertions.

Osgood and Tannenbaum's (1955) congruity principle is similar to Heider's balance model in that its assertions are qualitative in nature; i.e., they are either associative or dissociative. In contrast, *p*'s attitudes toward, or evaluations of, *o* and *x* are given quantitative values in the congruity principle. Recall that according to the congruity principle, whenever two objects of judgment are related by an assertion, the mediating reaction (i.e., evaluation) characteristic of each shifts toward congruence with the evaluation of the other, the magnitude of the shift being inversely proportional to intensities of the interacting evaluations. A state of congruence exists when the evaluations of two objects are equally intense (i.e., polarized) either in the same direction in the case of associative assertions, or in opposite directions in the case of dissociative assertions (see Osgood, Suci, and Tannenbaum, 1957, pp. 201–203). Thus, although a configuration may be balanced (according to balance theory), it will be incongruous (according to the congruity principle) unless the evaluations of both objects are equally polarized. Balance, therefore, is a necessary but not sufficient condition for congruence.

When a state of incongruity exists, the evaluations of the two objects will tend to change in the direction of congruity.[9] Consider, for example, a person whose evaluation of the President (on a scale ranging from -3 to $+3$) is $+2$ and whose evaluation of medicare is -1. Any assertion linking these two objects, whether associative or dissociative, would result in incongruity. With an associative assertion, congruity would exist if the two objects were evaluated identically. According to the principle of congruity (see Eq. 2.1), in this example both the President and medicare would come to have an evaluation of $+1$. With a dissociative assertion, the exemplified situation is balanced but incongruous. A state of congruence would exist if the evaluations of the two objects were equally polarized but differed in sign. In order to take the direction of the assertion into account, the congruity formula (see Eq. 2.1) can be rewritten as

$$P_R = \frac{|p_1|}{|p_1| + |p_2|} \, p_1 + \frac{|p_2|}{|p_1| + |p_2|} \, p_2 d,$$

where d is the direction of the assertion ($+1$ or -1), and all other terms are as given previously. The predicted values of p_1 and p_2 (that is, $\hat{p}_1$ and $\hat{p}_2$) are then given by Eq. 2.6 and 2.7.

$$\hat{p}_1 = P_R. \tag{2.6}$$

$$\hat{p}_2 = P_R d. \tag{2.7}$$

---

9. This tendency was mentioned in our discussion of congruity theory in the context of word combinations, and it may best be viewed as a compromise response when conflicting evaluations are linked by an assertion (see Osgood, Suci, and Tannenbaum, 1957).

In the present example, the congruity principle predicts that the evaluation of the President would shift to $+1.67$ and the evaluation of medicare to -1.67. That is,

$$P_R = \frac{2}{2+1}(2) + \frac{1}{2+1}(-1)(-1) = \frac{4}{3} + \frac{1}{3} = \frac{5}{3} = 1.67,$$

and $\hat{p}_1 = 1.67, \quad \hat{p}_2 = -1.67.$

Note that the magnitudes of the shifts in evaluations are again inversely proportional to their polarizations. That is, the $+2$ evaluation of the President shifts ⅓ of a unit, and the -1 evaluation of medicare shifts ⅔ of a unit.

One implication of viewing an assertion as a given is that incongruity can be resolved only by shifts in evaluations of the two objects and not by a change in the assertion itself. Consider, for example, a person who places a value of $+2$ on both the President and medicare and who receives a message asserting that the President opposes medicare. According to Heider's balance model, this dilemma is resolved if p rejects the assertion and perceives or believes that the President is in favor of medicare. This resolution is not possible in congruity theory, where the dilemma is resolved only when both evaluations shift to the neutral point.

In their research on the congruity principle, Osgood and Tannenbaum (1955) soon came to realize that a given assertion may not be believed. To enable their model to handle such cases, they introduced a "correction for incredulity." They assumed that some degree of incredulity exists when there is imbalance, that is, when two positively evaluated objects are dissociated or when one positively and one negatively evaluated object are associated. The greater the polarity of evaluations, the greater the degree of incredulity. Osgood and Tannenbaum argued that the congruity principle applies only to the degree that the assertion is perceived as credulous. In the extreme case of complete incredulity, no change is expected since it is assumed that no link betwen the two objects has been established.

An "assertion constant" was also added to the model in order to account for the finding that the object of an assertion (i.e., the x in the o-x link) tends to change more than would be predicted by the congruity formula.

Analysis of the Congruity Principle

Reconsidering the congruity principle in terms of our conceptual framework, we can see that its major concern is with attitudes toward two objects. Here again, attitude is defined in terms of an evaluative or affective dimension. Although the formal model deals with assertions linking one object to another, the concept of belief does not enter into the theory until the notion of incredulity is introduced. It is only at this stage that the person's belief in the assertion, i.e., his belief that o is related to x, is explicitly taken into account. Attitude change is expected to occur only to the extent that the assertion is believed. Unfortunately, Osgood and Tannenbaum's (1955) assessment of incredulity is problematic when it is viewed

as a measure of belief. Not unlike Heider's (1944) notion that belief formation is influenced by the dynamic character of the elements comprising the belief (i.e., by p's attitudes toward o and x), incredulity is assessed by assuming that it varies directly with imbalance and incongruity. This implies that imbalanced assertions are not likely to be believed, but balanced ones are. That is, incredulity is assumed to occur only when imbalance is present; the greater the incongruity of these imbalanced states, the greater the degree of incredulity. Clearly, however, this assumption does not always hold. It is possible for p to believe that a liked o opposes a liked x (for example, p's preferred candidate opposes a policy favored by p), or to disbelieve that a liked o favors a policy approved by p.

To be sure, these configurations constitute imbalanced states, but from a balance-theory viewpoint, incredulity or disbelief need not result. Indeed, it is because p holds the belief that imbalance is present. Further, recall that it is only when the assertion is not rejected that attitude change is expected in congruity theory. Thus the congruity principle again implies that a person's attitude is influenced by his belief that the attitude object is related to some attribute and by the evaluation of that attribute. Although the original congruity principle dealt with only one belief about an object, it was later extended to handle several beliefs simultaneously (Triandis and Fishbein, 1963; Anderson and Fishbein, 1965).

It is worth noting that inconsistency within a triadic configuration may involve beliefs, attitudes, or both. Assume that o and x are perceived as a unit or, in Osgood and Tannenbaum's (1955) terminology, that the assertion is credulous; p is linked to each of the two elements constituting the unit. In balance theory, these links may be either beliefs or attitudes, but in congruity theory only p's attitudes toward o and x are considered. Consequently, in balance theory inconsistency may exist between two beliefs, two attitudes, or a belief and an attitude; in congruity theory, inconsistency always involves two attitudes. In contrast, the consistency theory that has attracted most attention, i.e., dissonance theory, may be viewed as dealing only with inconsistency between beliefs.

A THEORY OF COGNITIVE DISSONANCE

Festinger's (1957) theory of cognitive dissonance begins with a consideration of the relations between two cognitive elements. "These elements refer to . . . the things a person knows about himself, about his behavior, and about his surroundings" (Festinger, 1957, p. 9). The following are examples of cognitive elements: "I know I smoke," "I know that smoking causes cancer," "I know I enjoy smoking," and "I know that George is my brother."

The terms *dissonance, consonance,* and *irrelevance* are used to describe three kinds of relations that may exist between any two cognitive elements. *"Two elements are in a dissonant relation if, considering these two alone, the obverse of one element would follow from the other.* To state it a bit more formally, x and y are dissonant if not-x follows from y." (Festinger, 1957, p. 13) For example, the

element "I know I smoke" would not follow from the element "I know smoking causes cancer," and hence these two cognitive elements are dissonant.

"If, considering a pair of elements, either one *does* follow from the other, then the relation between them is consonant" (Festinger, 1957, p. 15). For instance, since the element "I know I smoke" follows from the element "I know I enjoy smoking," this pair of cognitive elements is in a consonant relation.

Finally, "where one cognitive element implies nothing at all concerning some other element, these two elements are irrelevant to one another" (Festinger, 1957, p. 11). An irrelevant relation is exemplified by the two elements "I know I smoke" and "I know that George is my brother."

The basic hypothesis of dissonance theory was stated as follows:

> The existence of dissonance, being psychologically uncomfortable, will motivate the person to try to reduce the dissonance and achieve consonance. . . . The strength of the pressure to reduce the dissonance is a function of the magnitude of the dissonance. (Festinger, 1957, pp. 3, 18; italics omitted)

Considering a dissonant pair of cognitive elements, the magnitude of dissonance increases with the importance of the elements to the person. However, a given element may have relevant relations to more than one other element. Thus the cognitive element "I know I smoke" is consonant with "I know I enjoy smoking" and is dissonant with "I know that smoking causes cancer." The total amount of dissonance between any given cognitive element (e.g., "I know I smoke") and all other relevant elements is a function of the number of dissonant relations relative to the total number of relevant relations. The magnitude of dissonance will of course also depend on the importance of those relevant elements that exist in consonant or dissonant relations with the one in question.

It is possible to express these ideas more formally, as in Eq. 2.8,

$$D_k = \frac{\sum_{d=1}^{n} I_d}{\sum_{d=1}^{n} I_d + \sum_{c=1}^{m} I_c},\qquad (2.8)$$

where D_k is the magnitude of dissonance associated with cognitive element k; I_d is the importance of dissonant element d; I_c is the importance of consonant element c; n is the number of cognitive elements in a dissonant relation with element k; and m is the number of cognitive elements in a consonant relation with element k.[10]

When two cognitive elements exist in a dissonant relation, psychological tension or discomfort will motivate the person to reduce the dissonance and achieve

---

10. It has also been suggested that the magnitude of dissonance can be defined as a ratio of dissonant to consonant elements, each weighted by its importance (Kiesler, Collins, and Miller, 1969; Brehm and Cohen, 1962). Despite the difference between this definition and Eq. 2.8, where the ratio is computed in terms of the total number of relevant relations, the two formulations have sometimes been discussed as if they were interchangeable (e.g., Brehm and Cohen, 1962).

consonance. The only way to completely eliminate the existing dissonance is to change one of the two elements involved. For example, the dissonant relation between the two cognitive elements "I know I smoke" and "I know that smoking causes cancer" could be changed to a consonant relation in one of two ways. The person might change the first cognitive element to "I know I don't smoke," or he might change the second to "I know that smoking does not cause cancer."

Given that cognitive elements are responsive to "reality" (Festinger, 1957), it may sometimes be impossible to change either of the cognitive elements in a dissonant relation. However, the *magnitude* of dissonance associated with a cognitive element may be reduced in two ways. First, the person may add new cognitive elements that are consonant with the element in question. Thus, in the example above, the cognitive element "I know that I enjoy smoking" may be introduced. Indeed, a person may be expected to actively seek new consonant information and try to avoid exposure to dissonant information. Second, the person may reduce the importance of one or both elements in the dissonant relation.

Festinger (1957) described four basic situations that give rise to cognitive dissonance: decision making, forced compliance, voluntary and involuntary exposure to dissonant information, and disagreement with other persons. For example, whenever a person makes a choice between two or more alternatives, dissonance is assumed to exist. That is, his knowledge that the unchosen alternatives have favorable aspects and his knowledge that the chosen alternative has unfavorable aspects are both dissonant with his knowledge of his choice. In this situation, the theory predicts that dissonance may be reduced by increasing one's evaluation of the chosen alternative and/or decreasing one's evaluation of the unchosen alternatives. The amount of change will be related to the magnitude of dissonance involved. The magnitude of postdecision dissonance is an increasing function of the general importance of the decision and of the degree to which chosen and unchosen alternatives are similar in attractiveness.

In the forced-compliance situation, an individual is induced to perform a behavior that is inconsistent with his beliefs or attitudes. For example, a prisoner of war who believes that communism is a repressive system might be induced (through threat of punishment or promise of reward) to state publicly that communism is not a repressive system. Dissonance would exist between the two cognitive elements "I know that communism is a repressive system" and "I know I publicly said that communism is not a repressive system." The magnitude of dissonance in this situation is inversely related to the amount of threatened punishment or promised reward. That is, the greater the justification for the behavior, the less dissonance it arouses. Here dissonance theory predicts that the person can reduce dissonance by changing his belief about communism such that it becomes consonant with his behavior. The greater the magnitude of dissonance, the greater the expected change in belief.[11]

---

11. Note that dissonance will also arise when the person refuses to comply. Then the dissonance exists between "I know I refused to state publicly that communism

These examples point to another important assumption in dissonance theory, namely, that cognitive elements are responsive to reality; by and large they mirror or map reality. This does not mean, however, that cognitive elements will always correspond to reality. Rather, "the major point to be made is that the reality which impinges on a person will exert pressures in the direction of bringing the appropriate cognitive elements into correspondence with that reality" (Festinger, 1957, p. 11; italics omitted).

Analysis of Dissonance Theory

When dissonance theory is viewed in terms of our conceptual framework, a number of interesting implications can be derived. First, it may be argued that cognitive elements are equivalent to what we have defined as beliefs. That is, a cognitive element refers to a person's *knowledge* that he holds a certain attitude or a certain belief or that he performed a certain behavior and *not* to the attitude, belief, or behavior itself. For example, consider the following three cognitive elements.

1. I know I like Senator Kennedy.
2. I know I believe that Senator Kennedy participated in an antiwar demonstration.
3. I know I participated in an antiwar demonstration.

The first element constitutes a belief about an attitude; the second refers to a belief about a belief; and the third is the person's belief about his own behavior. Not unlike some of the consistency theories we have discussed earlier, dissonance theory deals only with qualitative beliefs. That is, a person either holds or does not hold the belief in question.

From our point of view, beliefs about attitudes are not equivalent to the attitudes themselves, nor are beliefs about behaviors equivalent to the behaviors themselves. It could be argued, however, and indeed Festinger (1957) has argued that beliefs about beliefs are psychologically equivalent to the beliefs themselves.

Similarly, it is possible that measures of beliefs about attitudes will sometimes yield results similar to those of direct measures of attitude. Thus a measure of "I like Kennedy" on a probability dimension may be highly correlated with a measure of the concept "Kennedy" on an evaluative or affective dimension. However, some psychodynamic theories might suggest that a person does not always know what his "true" belief or attitude is, i.e., that the cognitive element may not be a veridical representation of the belief or attitude. For example, according to psychoanalytic theory, defense mechanisms such as repression or reaction formation may affect veridicality by eliminating some threatening belief or attitude from consciousness or by replacing it with its opposite.

---

is a repressive system" and "I know I was punished for my refusal." The magnitude of dissonance in this situation is an *increasing* function of amount of punishment obtained or reward forgone.

It is worth noting that research on dissonance theory has usually obtained measures of beliefs and attitudes, rather than of their cognitive representations. Although this may appear problematic in view of the considerations above, the issue is really more theoretical than practical since beliefs and attitudes are usually assessed by direct self-reports in studies on dissonance theory. In some cases, however, a direct measure of attitude may not be the same as a measure of the belief about the attitude, and then the measure of attitude would be inappropriate for a test of dissonance theory. Further, indirect measures of attitude (see Chapter 3) may also be inappropriate since they may not correspond to the person's knowledge of his attitude.

Also noteworthy is the fact that Festinger's cognitive elements can be viewed as perceived relations in the balance-theory formulation. In the example above, the following triad emerges: "I admire Senator Kennedy" (pLo); "Senator Kennedy participated in an antiwar demonstration" (oUx); and "I participated in an antiwar demonstration" (pUx). However, Festinger's cognitive elements refer to the person's knowledge or belief about each relation in the triad. That is, "I believe (pLo)"; "I believe (oUx)"; and "I believe (pUx)."

Clearly, then, consonance and dissonance refer to relations between beliefs and only between beliefs. Other variables, i.e., attitudes, intentions, or behaviors, are relevant only to the extent that they are represented cognitively. However, relations between beliefs, i.e., between cognitive elements, may influence these other variables (as well as the beliefs themselves). Thus changes in beliefs, attitudes, intentions and behaviors are all viewed as influenced in a similar manner by dissonant relations between beliefs. Although Festinger regarded beliefs about one's own behavior as a special set of cognitive elements, he made little or no distinction between the cognitive representations of beliefs and attitudes. Thus dissonance theory makes no differential predictions about the effects of dissonance on changes in beliefs and attitudes. Moreover, in most applications of dissonance theory, distinctions are usually not made between beliefs, attitudes, and behaviors themselves on the one hand and their cognitive representations on the other.

A typical investigation examines the effects of behaviorally induced dissonance on beliefs and/or attitudes. In a later chapter we will suggest that some of the conflicting findings in research on dissonance phenomena may be due to the failure of dissonance theory to make a distinction between beliefs and attitudes, and to the assumption that a dissonant relation between cognitive elements will have similar effects on these two variables. Perhaps the most important contribution of dissonance theory has been to direct attention to the possibility that different amounts of dissonance may be created by the performance of a given behavior, and therefore that the performance of the behavior may sometimes lead to a considerable change and at other times to little change in any given belief, attitude, or intention. Unfortunately, the assumption of dissonance theory that these effects are mediated by beliefs and the relations between beliefs has often been neglected.

One problem that has arisen with respect to dissonance theory is related to the

difficulty of specifying when two cognitive elements stand in a dissonant relation; that is, what is meant by "the obverse of one element follows from the other"? Aronson (1968) has suggested that "the major source of conceptual ambiguity rests upon the fact that Festinger has not clarified the meaning of the words 'follows from' " (Aronson, 1968, p. 9). He suggested a rule of thumb whereby "the obverse of one element follows from the other" would be defined in terms of the violation of an expectancy.

This interpretation implies a distinction between imbalance and dissonance. A person's knowledge that a liked friend has a negative trait is clearly imbalanced, but it is not necessarily dissonant. It would be dissonant only if the person did not expect his friend to have that negative trait. This distinction between imbalance and dissonance corresponds closely to the distinction between affective and cognitive inconsistency. As we pointed out earlier, balance theory deals with both types of inconsistency whereas dissonance theory, which is concerned with dissonance between cognitions or beliefs, can deal only with cognitive inconsistency.

A second and perhaps more important problem concerning dissonance theory is the specification of the amount of dissonance that is expected to result from the performance of a behavior in a given situation. Most elaborations and extensions of dissonance theory have centered on this problem. The most important factor relates to "volition," or freedom of choice, in performing the behavior. The basic argument is that dissonance is created only to the extent that the person feels he had freedom of choice (Brehm and Cohen, 1962). This notion is closely related to Festinger's (1957) discussion of the forced-compliance situation, where the amount of dissonance resulting from compliance is inversely related to the amount of pressure to comply (e.g., threatened punishment or promised reward). The person's perception of his freedom of choice may be expected to decrease with increased pressure. In other words, dissonance is likely to occur to the extent that there is insufficient justification for the performance of the behavior. It is assumed that promised rewards or threatened punishments, for example, will provide justification for performing the behavior and hence will add consonant elements to the dissonant relation. This should serve to reduce the magnitude of total dissonance.

Finally, one might question the necessity of postulating an aversive motivational state, i.e., dissonance, in order to account for the effects of behavior on beliefs and attitudes in situations such as forced compliance or decision making. Bem (1965, 1968a, 1972) has in fact argued that it is not necessary to make such an assumption, and he has proposed an alternative interpretation based on Skinner's (1957) analysis of verbal behavior. According to this position, beliefs and attitudes are simply self-descriptive, verbal responses. Such responses, like other responses, are under the control of internal or external stimuli. One important source of stimuli for "attitudinal" responses is the person's own behavior, together with the context in which it occurs. Thus "an individual's belief and attitude statements and the beliefs and attitudes that an outside observer would

attribute to him are often functionally equivalent in that both sets of statements are 'inferences' from the same evidence" (Bem, 1965, p. 199).

To see how this might explain dissonance phenomena, consider a person who states publicly that America should cease to support Nationalist China. If he does so for a large reward (e.g., $5,000), other people as well as the person himself will perceive this behavior to be under the control of external stimuli, namely, the reward, and thus little information is provided about the person's stand on the issue. In contrast, when he makes the same statement for little reward (e.g., $1), his behavior will be attributed to internal stimuli, i.e., his personal belief. Thus, in the low-reward condition the inference drawn will be that the person really believes that America should cease to support Nationalist China; this inference will not be drawn in the high-reward condition. Note that this prediction is the same as that derived from dissonance theory (i.e., the smaller the reward, the greater the dissonance, and thus the more change).

Bem's analysis concerns the attribution of dispositions (i.e., beliefs, attitudes, etc.) to one's self and to others. Although Bem's analysis has been applied primarily to the explanation of dissonance phenomena, theories of attribution deal with questions of a more general nature, namely, with the formation and change of beliefs.

THEORIES OF ATTRIBUTION

Recall that Heider's (1946, 1958) development of the balance model grew out of his interest in causal attribution, i.e., the formation of causal units. His first question concerned the degree to which a given action or event would be attributed to some person or object. Imagine, for example, that a violent crime has been committed and that two persons were in the vicinity at the time of the crime. As we have seen earlier, according to Heider, a causal unit will easily be formed if the dynamic characters of two entities are similar. Thus, if one of the persons involved is perceived as brutal and the other as gentle, the violent crime is likely to be attributed to the first. If only the gentle person had been found in the vicinity of the crime, the event might still not have been attributed to him but to an external factor, e.g., a third person (Heider, 1944).

Heider (1958) distinguished five levels of causal attribution in reference to the attribution of responsibility for the outcomes of an action: association, commission, foreseeability, intentionality, and intentionality with justification. At the first level, the actor is held responsible for any effect that is in some way associated with him. At the second level, he is held responsible only when the effect is seen as a direct result of his behavior. Attribution of responsibility at the third level requires that the effect was foreseeable, even if not intended. Intentionality is the prerequisite for attribution of responsibility at the next level; that is, here the actor is held responsible only for effects that he foresaw and intended. Finally, if his action is perceived as justified, i.e., caused by factors beyond his control, he will be held

less responsible, even though he may have intended to produce the observed effects.

Later, Heider (1958) went beyond the formation of causal units and considered attribution of stable dispositions to an actor. "Dispositional properties are the invariances that make possible a more or less stable, predictable, and controllable world. They refer to the relatively unchanging structures and processes that characterize or underlie phenomena" (Heider, 1958; p. 80). Indeed, for the most part, attribution theory is concerned with inferences about stable dispositions of people based on information about or direct observation of their actions. Specifically, attribution theory deals with specifying the conditions under which attributions to a person will or will not be made.

A distinction has been made between internal and external attribution; that is to say, a person's behavior may be attributed to some disposition of the person himself or to some external factor. In the latter case, no inferences can be made about the person's stable characteristics. For example, if a person is observed to succeed at some task, his performance may be attributed to his ability or motivation (i.e., internal attribution) or to an external cause, such as good luck or the low difficulty level of the task.

Heider (1958) distinguishes between personal and impersonal causality. By personal causality he refers to levels four and five in the attribution of responsibility, i.e., to instances in which a person is perceived to have caused a certain event intentionally. According to Heider, internal attributions are made only under conditions of personal causality, i.e., where the action is perceived to be purposive. Jones and Davis's (1965) analysis of attribution processes centers on such personal causality. According to this analysis, attribution of a disposition to an actor is based on the observation of his action and its consequences or effects. The perceiver's basic problem is to decide which of these effects, if any, were intended by the actor. Two factors are assumed to influence the degree to which the actor will be perceived to have intended a given action. The first condition in the inference process is the assumption of knowledge (or foreseeability) on the part of the actor; that is, "in order to perceive that at least some of the effects achieved by an action were intended, the perceiver must first believe that the actor was aware his actions would have the observed effects" (Jones and Davis, 1965, p. 220). The second condition in the inference process is the assumption of ability on the part of the actor; that is, the observed effects may have been achieved as a result of luck or chance rather than ability. Thus, for intention to be inferred, the actor must be perceived to have had (a) the knowledge that the effects would result from his action, and (b) the ability to produce the effects. These ideas are summarized in Fig. 2.5.
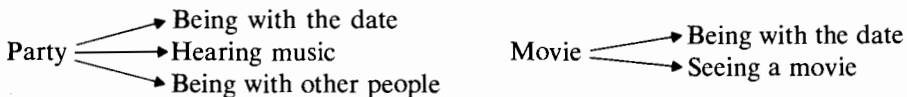
Two questions, then, are central to dispositional attribution: (1) What dispositions will be attributed to an actor? (2) What factors influence the confidence with which such attributions are made? Most theorizing on attribution processes is concerned with the latter question rather than the former.
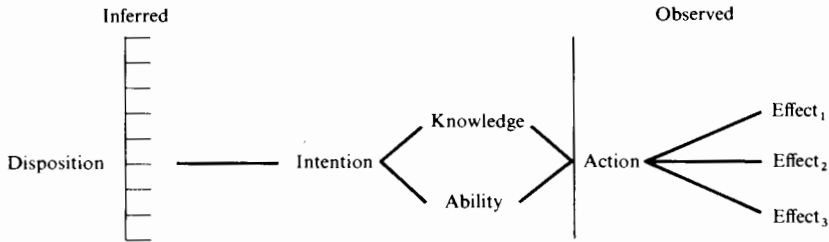
Fig. 2.5 The action-attribute paradigm. (Reproduced from Jones and Davis, 1965, p. 222.)

According to Jones and Davis (1965) the certainty of attributions (or as they call it, correspondent inference) depends on two factors: the desirability of the effects produced by the action, and the degree to which these effects are common to other behavioral alternatives that were available to the actor. Assume that a person is observed to perform one of two possible behaviors, and that these behaviors would produce some common and some noncommon effects. The lower the perceived desirability of the effects produced by the behavior that is performed and the fewer the unique effects (i.e., noncommon effects) produced by that behavior, the more confident an observer will be in attributing a disposition to the actor. The particular disposition that is attributed will correspond to the unique effects produced by the behavior.

For example, consider a person who can take his date either to a party or to a movie, and assume that the relevant effects of these alternative behaviors are as follows:

Party ⟨ Being with the date / Hearing music / Being with other people   Movie ⟨ Being with the date / Seeing a movie

First, note that "being with the date" is common to both alternatives. The other effects are unique. If the person went to the party, two attributions could be made. According to the theory, the actor might be perceived to like music or to like being with other people. Thus the observer would not be certain as to the actor's disposition. If the actor went to the movie, however, the attribution would be made that he likes movies. Further, according to the theory, the less attractive or desirable that particular movie is perceived to be, the more certain the observer will be that the actor likes movies.

On the basis of Heider's (1958) analysis, Kelley (1967, 1971, 1972, 1973) has also discussed some of the ways in which effects produced by an action are attributed to the various factors present in the situation. He suggested two major principles of causal attribution. The first is the *principle of covariation,* which states, "The effect is attributed to that condition (factor) which is present when the effect is present and which is absent when the effect is absent" (Kelley, 1967,

p. 194). Not unlike a scientist, the naive observer is assumed to examine the presence or absence of effects under different conditions, at different points in time, and for different actors. One of Kelley's major concerns is with attributions to the actor (internal) as opposed to other factors (external). Internal attribution requires that the actor behave consistently under different conditions and at different times, and also that his behavior differ from that of other actors. For example, if a person always assists others in need of help, even under adverse conditions, and in this behavior he is not like other actors in similar situations, he is likely to be perceived as altruistic. When the actor's behavior varies across situations and is similar to the behaviors of others in those situations, some factor in the situation is assumed to be responsible for his behavior; that is, an external attribution is made.

Specifically, Kelley (1967, 1972) identified three major factors that influence attribution. (1) *Consistency*—the degree to which the actor performs the same behavior toward an object on different occasions. The more consistent his behavior, the more likely it is that an internal attribution will be made. (2) *Distinctiveness*—the degree to which the actor performs different behaviors with respect to different objects. The lower the distinctiveness (i.e., the more the actor performs the same behavior with respect to different objects) the more likely it is that an internal attribution will be made. (3) *Consensus*—the degree to which other actors perform the same behavior with respect to a given object. Internal attribution decreases with consensus.

The second principle refers to *multiple plausible causes* for a behavior performed at a given point in time. Clearly, the greater the number of plausible explanations, the lower the certainty of any given attribution.

Other factors, such as the actor's perceived decision freedom and the prior probability of his behavior, have also been suggested as determinants of the likelihood that an attribution will be made. For example, Steiner (1970) argued that confidence in an internal attribution should increase with the actor's perceived decision freedom. Steiner's notion of perceived decision freedom is related to the desirability of the effects associated with the actor's behavioral alternatives. According to Steiner (1970, p. 195) "whenever an individual must choose between two or more available alternatives, decision freedom should be a negative function of the discrepancies between the gains offered by the available options." Consider, for example, a person who performs a behavior that results in desirable consequences. If the other alternatives available to him would also have led to desirable consequences, the actor would be perceived to have had a high degree of decision freedom; since he was free to choose other attractive alternatives, his actual behavioral choice can be used to make inferences about his dispositions. However, if the other available alternatives would have led to undesirable consequences, he would be perceived to have had little decision freedom; thus little information about his dispositions would be gained.

It is worth noting that all the factors above that have been assumed to influence attribution involve the utilization of information obtained about an actor's

behavior and about the circumstances under which it occurred. However, as Heider (1958) and others have noted, factors residing in the observer may also affect his attributions. Thus, as indicated earlier, a person's attitudes may influence his perception of causal units. Jones and Davis (1965) argued that when an actor's behavior directly affects the observer, attributions to the actor are made with greater confidence than when the same behavior is directed toward some other person. Another possibility that has been mentioned (e.g., Heider, 1944; Kelley, 1971; Jones and Nisbett, 1971) is that an observer may tend to underemphasize the importance of the environment and to exaggerate the importance of the personal factor.

Analysis of Attribution Theories

From the point of view of our conceptual framework, attribution theory deals with the formation and change of beliefs.[12] These beliefs are *inferences* about the causes of observed events or about an actor's stable dispositions. All the principles discussed above concern the degree to which observing a person's behavior influences beliefs about him or about the environment. Since a person also serves as an observer of his own behaviors, attribution theory has implications for the formation of beliefs about one's self. Although it has been argued that the principles governing attribution of dispositions to others are similar to those governing attributions of dispositions to one's self (e.g., Bem, 1965; Steiner, 1970), some theorists (e.g., Jones and Nisbett, 1971) have suggested that different processes are involved. This question will be discussed in Chapter 5.

Irrespective of whether self-attributions or attributions to others are considered, a central concern of attribution theory is with the subjective probability that a given trait or disposition is associated with a given actor. However, it must also be realized that, prior to observing the actor's behavior, the observer will have had some subjective probability that the actor has the disposition in question (even if that probability is .50, or complete uncertainty).[13] One may therefore argue that the degree to which an attribution has been made in a given situation is reflected in the *change* or shift in the belief (i.e., a change in subjective probability), rather than in the absolute level of the belief (see Ajzen, 1971).

This conceptualization of the attribution process as a revision in subjective probability suggests the possibility of applying Bayes's theorem as a model of the attribution process. Bayes's theorem and other probability models describing relations between beliefs (cf. McGuire, 1960a,b,c; Warr and Smith, 1970) will be discussed in Chapter 5, which considers the formation of beliefs.

12. As in dissonance theory, these may be beliefs about attitudes, intentions, or behaviors.

13. Such prior beliefs may be based on familiarity with the actor, with his social background, education, etc. Indeed, one difference between attributions to self and others may be due to differences in prior beliefs held by actor and observer.

CONCLUSION

This chapter has reviewed some of the contemporary theories in the attitude area. Although we have repeatedly pointed to similarities, we have also shown that these theories differ in many respects. They differ in terms of the variables that play a central role in the theory, in the ways that the variables are interrelated, and in terms of focusing on processes of formation and/or change of the variables. First, a distinction should be made between the constructs of a theory and the operations that have been employed to manipulate or assess those constructs. For instance, in our discussion of dissonance theory we pointed out that at the conceptual level the theory deals only with cognitions, i.e., beliefs. Operationally, however, these cognitions are often manipulated by changing behavior and assessed by measuring attitude.

From our point of view, most theories can be classified at the conceptual level as dealing either with beliefs or attitudes or with both beliefs and attitudes. With the exception of a few learning theories (e.g., Doob, 1947; Lott and Lott, 1968) and Edwards's (1954) decision theory, behavioral intentions and behavior are dealt with only at the operational level, not at the conceptual level. A simple classical conditioning model deals only with attitudes (e.g., Staats and Staats, 1958) whereas dissonance and attribution theories deal only with beliefs. The more sophisticated learning theories, most expectancy-value models, congruity theory, and balance theory deal with both beliefs and attitudes at the conceptual level.

It should be clear that since different variables are involved, the theories described also deal with different relationships. Thus some theories are concerned only with the relations between attitudes (e.g., Staats and Staats, 1958), but others are concerned with relations between beliefs (e.g., dissonance theory). Most theories, however, deal with the relations between beliefs and attitudes. Further, some theories specify causal relations (e.g., expectancy-value models, attribution theories), and others are concerned with the dynamic interrelations among variables (e.g., balance theory, dissonance theory). For example, according to expectancy-value models, beliefs and associated evaluations are the determinants of attitude. In contrast, balance theory is concerned with the effects of imbalance on the relationships in a given configuration.

Stated somewhat differently, some theories postulate what can best be described as an information-processing model; information about an object or issue or about one's self leads to the formation of beliefs or attitudes. Other theories postulate a dynamic process where information affects beliefs or attitudes only to the extent that it introduces some inconsistency or instability among these variables.

This discussion suggests that virtually every theory is in some way concerned with information. The information may be about an object or about one's own beliefs, attitudes, intentions, or behaviors with respect to the object. Both types of information may be gained either through direct observation or by means of some communication. Although all theories deal with information about the object,

relatively few (e.g., balance theory, dissonance theory, attribution theory) also deal with information about the self.

Some theorists might argue with the use of the term "information" in this context. For example, classical conditioning is usually assumed to occur without awareness. In contrast, the present analysis assumes that pairing an object with an unconditioned stimulus provides information about the object; that is, the subject is aware of the contingency and thus forms a new belief. This information contributes to his attitude. The question of awareness in classical conditioning will be discussed more fully in Chapters 4 and 6.

At the operational level, many investigators fail to distinguish between beliefs, attitudes, and intentions. Thus, although the conceptual variable in a theory may be attitude, the operation utilized may be assessing beliefs, intentions, or even behavior. A second problem discussed previously regarding dissonance theory is the possibility that measures of beliefs, attitudes, intentions, or behaviors may not correspond directly to their cognitive representations. Thus, for example, a person's belief that he holds a certain attitude may not be adequately assessed by a direct measure of the attitude. It is interesting to note that although dissonance theory does suggest that beliefs about an attitude and the attitude per se may differ, at least one self-attribution theory (Bem, 1968a) holds that these two concepts are identical in that they are both self-descriptive responses. From our point of view, however, attribution theories deal only with beliefs, and tests of these theories should therefore employ only measures of beliefs.

Another difficulty of particular concern to dissonance theory is its failure to specify exactly the cognitive elements of relevance in a given situation. It is thus left to the investigator's intuition to select the relevant elements. Some investigators may employ measures of beliefs, and others may employ measures of attitude, etc.

Finally, the theories discussed in this chapter can be compared in terms of their focus on formation and/or change of beliefs and attitudes. Generally speaking, theories based on notions of information processing deal directly with processes of formation and therefore have immediate implications for change. Indeed, it is often difficult to tell whether some new information leads to the formation of a belief or attitude, or whether it changes an existing belief or attitude. The different learning theories, as well as expectancy-value models, are concerned with the effects of information on attitude whereas attribution theories are concerned with the effects of information on beliefs.

In contrast, theories postulating a dynamic process tend to focus on change, without great concern for the formation of beliefs or attitudes. Thus the congruity model is concerned with changes in attitudes produced by incongruity, and dissonance theory deals with the effects of inconsistent beliefs. Attempts to reduce dissonance may produce changes in beliefs, attitudes, intentions, or behaviors. Similarly, balance theory deals with changes in both beliefs and attitudes, and it can also be applied to changes in intentions and behaviors.

Our comparison of attitude theories discussed in this chapter is summarized in Table 2.2. One can see that theories may differ in terms of their conceptual variables and thus the kinds of relations they can deal with, in terms of the processes underlying these relations (informational or dynamic), and in terms of the operational variables that have been considered in empirical research. Finally—not shown in the table—information-processing theories focus on formation and change, whereas dynamic theories tend to deal with change only.

Table 2.2 Comparison of Contemporary Attitude Theories

	Conceptual variables	Type of theory	Operational variables
Learning theories			
Staats and Staats	A	Informational	A
Lott	b,A,B	Informational	b,A,I,B
Doob	b,A,B	Informational	—
Expectancy-value theories			
Fishbein	b,A	Informational	b,A
Edwards	b,A,B	Informational	B
Rosenberg	b,A	Dynamic	b,A
Consistency theories			
Balance (Heider)	b,A	Dynamic	b,A,I,B
Congruity (Osgood and Tannenbaum)	b,A	Dynamic	A
Dissonance (Festinger)	b	Dynamic	b,A,I,B
Attribution theories			
Self-attribution (Bem)	b	Informational	b,A,I
Attribution to others	b	Informational	b,A,I

Abbreviations: b = belief, A = attitude, I = intention, B = behavior