Psychological Bulletin
1994, Vol. 115, No. 3, 401–423

0033-2909/94/$3.00

Relationship Between Attitudes and Evaluative Space: A Critical Review, With Emphasis on the Separability of Positive and Negative Substrates

John T. Cacioppo and Gary G. Berntson

Evaluative processes refer to the operations by which organisms discriminate threatening from nurturant environments. Low activation of positive and negative evaluative processes by a stimulus reflects neutrality, whereas high activation of such processes reflects maximal conflict. Attitudes, an important class of manifestations of evaluative processes, have traditionally been conceptualized as falling along a bipolar dimension, and the positive and negative evaluative processes underlying attitudes have been conceptualized as being reciprocally activated, making the bipolar rating scale the measure of choice. Research is reviewed suggesting that this bipolar dimension is insufficient to portray comprehensively positive and negative evaluative processes and that the question is not whether such processes are reciprocally activated but under what conditions they are reciprocally, nonreciprocally, or independently activated.

In *The Philosophy of Physical Science*, Sir Arthur Eddington (1939; cited in R. F. Green & Goldfried, 1965) described a hypothetical scientist who sought to determine the size of the various fish in the sea. The scientist began by weaving a 2-in. mesh net and setting sail across the seas, repeatedly sampling catches and carefully measuring, recording, and analyzing the results of each catch. After extensive sampling, the scientist concluded that there were no fish smaller than 2 in. in the sea. The moral of Sir Eddington's analogy is twofold. First, scientific instruments are shaped by a developer's implicit theory of the phenomena to be investigated. The scientist in Sir Eddington's analogy began with a notion that fish in the sea were sizable, and the construction of the 2-in. mesh net reflected this notion. Second and relatedly, scientific instruments, once developed, shape data and theory in ways more fundamental than are often appreciated. Thus, the scientist in Sir Eddington's analogy found support for the initial notion of what constituted a fish, and the scientist's observations predisposed subsequent investigators to think about fish in the same way.

Like Sir Eddington's hypothetical scientist, social scientists have carefully investigated factors and modeled processes by which attitudes are formed and changed.[1] By attitudes, we mean general and enduring favorable or unfavorable feelings about, evaluative categorizations of, and action predispositions toward stimuli (e.g., see Cacioppo, Petty, & Berntson, 1991; Thurstone, 1928; Zanna & Rempel, 1988). The attitude measures most often used—bipolar rating scales—reflect an assumption underlying the study of attitudes since Thurstone (1928, 1931): An attitude is reducible to the net difference between the positive and negative valent processes aroused by a stimulus (Allport, 1935; R. F. Green & Goldfried, 1965; Thurstone, 1931). This assumption can be expressed as three key principles: (a) An attitude is a joint function of positively and negatively valent activation functions (principle of evaluative activation); (b) positively and negatively valent activation functions have generally opposing effects on an attitude (principle of opposing evaluative actions); and (c) positively and negatively valent activation functions are reciprocally controlled (principle of reciprocal evaluative activation). The third of these principles was an important starting point that simplified the measurement of attitudes. People's feelings about an object, person, or issue can range anywhere between two endpoints: maximally positive (and minimally negative) to maximally negative (and minimally positive). By assuming that positive and negative evaluative processes are reciprocally controlled, these endpoints can be placed on a single (bipolar) continuum, and attitudes can be measured along this continuum or vector. The conceptualization of attitudes as falling along a positive–negative continuum also fits well with the behavioral dispositions of approach–withdrawal. Accordingly, some early attitude theorists working during the heyday of behaviorism defined attitudes as nothing more than a bivalent action disposition toward a stimulus (e.g., Doob, 1947). With the advent of the cognitive paradigm in psychology, attitude theorists began distinguishing evaluative categorizations of stimuli

John T. Cacioppo and Gary G. Berntson, Department of Psychology, Ohio State University.

This research was supported by National Science Foundation Grant DBS-9211483. Portions of this article were presented by John T. Cacioppo as an invited address at the meeting of the Midwestern Psychological Association in Chicago, May 1991. We would like to thank Anthony Greenwald, Stephen Crites, Mary Snydersmith, and Joseph Priester for comments on this research.

Correspondence concerning this article should be addressed to John T. Cacioppo, Department of Psychology, Ohio State University, 1885 Neil Avenue, Columbus, Ohio 43210-1222. Electronic mail may be sent to Cacioppo.1@osu.edu.

[1] In Sir Eddington's analogy, the limitation of the 2-in. mesh net was its measurement sensitivity, a fairly obvious measurement limitation that careful investigators would probably not overlook when drawing inferences from their measurements (e.g., see Dawes & Smith, 1985; Ostrom, 1973). However, some of the measurement limits of bipolar attitude rating scales are less intuitive, and therefore their implications for understanding attitude phenomena less obvious, than are the issue and implications of limits in measurement sensitivity.

from bivalent action dispositions toward stimuli (e.g., Cacioppo, Crites, Berntson, & Coles, 1993; Fazio, 1990). The notion that the evaluative processes underlying attitudes were reciprocally activated, however, continued largely unchallenged (cf. R. F. Green & Goldfried, 1965).

Interestingly, even though physical constraints may restrict behavioral manifestations to bivalent actions (approach–withdrawal), early behavioral theorists recognized that approach and withdrawal were behavioral manifestations that could be thought of as derived from separable motivational substrates (Miller, 1951, 1961). Conflict theory was enriched by conceptualizing approach and withdrawal separately, investigating their unique antecedents and consequences, and examining the psychological and physiological constraints that led typically to the reciprocal activation of approach and withdrawal tendencies (Lang, Bradley, & Cuthbert, 1992). The major thesis of the present article is that attitude theory and measurement may similarly benefit from expanding the principle of reciprocal evaluative activation to accommodate (a) the separable activation of positive and negative evaluative processes, (b) investigation of their unique antecedents and consequences, and (c) examination of the psychological and physiological constraints that produce their reciprocal activation. Thus, the third principle outlined earlier can be replaced by the following: Positively and negatively valent activation functions are not necessarily reciprocally activated but can be activated reciprocally (e.g., mutually exclusive and incompatible), uncoupled (e.g., singularly activated), or nonreciprocally (e.g., coactivational or coinhibitory; principle of bivalent modes of evaluative activation).

The introduction of bivalent modes of evaluative activation requires a two-dimensional representation of positive and negative evaluative activation. The bivariate evaluative plane depicted in Figure 1 accommodates all possible combinations of positive and negative evaluative activation: (a) The reciprocal mode of evaluative activation is represented as one diagonal vector that ranges from maximal positivity–minimal negativity to maximal negativity–minimal positivity; (b) the nonreciprocal mode of evaluative activation is represented as the alternate diagonal that ranges from minimal positivity and negativity to maximal positivity and negativity; and (c) the uncoupled modes of evaluative activation are represented as vectors lying along the axes. The family of vectors parallel to those just mentioned represents the general categories, or modes, of evaluative activation expressed from varying starting points within the two-dimensional plane depicted in Figure 1.[2]

The reciprocal activation of positive and negative evaluative processes by an attitude object may be commonplace, just as most of the fish in the sea may be longer than 2 in. However, just as the use of a 2-in. mesh net by Sir Eddington's (1939) hypothetical scientist allowed certain species of fish to slip through unnoticed, the theory that attitude objects necessarily evoke the reciprocal activation of positive and negative evaluative processes may hinder theoretical and empirical investigation of certain classes of attitudinal phenomena. The bivariate structure depicted in Figure 1 provides an interesting alternative perspective. Its utility, however, stems from the questions it stimulates and the relative ease rather than the exclusivity with which interesting attitude phenomena can be represented. For instance, the low activation of positive and negative evaluative

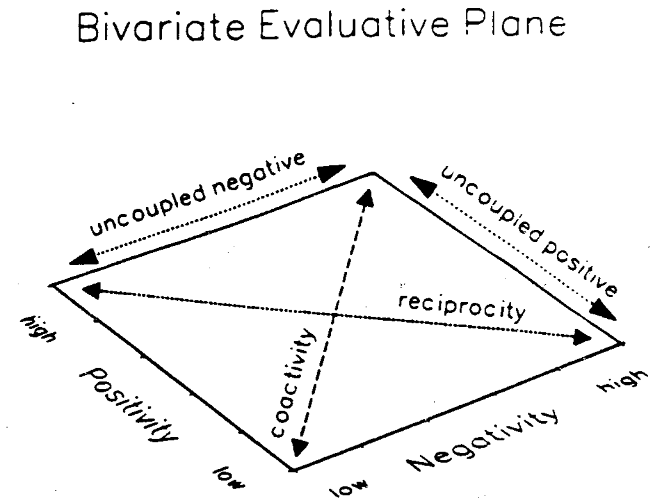

Figure 1. The bivariate evaluative plane. The left axis represents the level of excitatory activation of positive evaluative processes (labeled positivity), and the right axis represents the level of excitatory activation of negative evaluative processes (labeled negativity). Along each axis, the level of activity increases with movement away from the front axis intersection. The dotted diagonal extending from the left to the right axis intersections represents the diagonal of reciprocal control (labeled reciprocity). The dashed diagonal extending from the back to the front axis intersections depicts the diagonal of nonreciprocal control (labeled coactivity). The arrows alongside the axes represent uncoupled changes in positive or in negative evaluative processing. These diagonals and axes, and vectors parallel to them, illustrate the major modes of evaluative activation.

processes by a stimulus is depicted at the front corner of Figure 1 and reflects neutrality, whereas high activation of positive and negative evaluative processes by the stimulus is depicted at the rear corner of Figure 1 and reflects maximal conflict. The projection of both these points onto the reciprocal diagonal, or onto a bipolar rating scale, places them at the midpoint. The bipolar conceptualization of attitudes could also accommodate these attitude phenomena in at least two ways. Attitude conflict could be depicted as maximal at the midpoint because the attitude object evokes moderate and approximately equal feelings of positive and negative affect. Attitude conflict could also be depicted as maximal at the midpoint because that is the central tendency of oscillating positive and negative evaluative reactions to the stimulus. Note that attitude neutrality is also maximal at the midpoint because the attitude object evokes neither significant feelings of positive affect nor significant feelings of negative affect. Thus, there can be an ambiguity in the psychological significance of bipolar ratings that is not inherent in the bivariate conceptualization because all possible combinations of positive and negative evaluative activation are represented uniquely.

The issue addressed in this article is not whether positive and negative evaluative processes are stochastically independent but

[2] Vectors that are not parallel to those just mentioned represent intermediate modes of evaluative activation (e.g., negatively weighted reciprocal mode).

whether attitude theory and measurement might be fostered by adopting a framework and measurement strategies in which the antecedents and consequences of positive and negative evaluative processes could be investigated separately. Stochastic independence is "the relation between two events in which the probability of their joint occurrence is equal to the product of the probabilities of the occurrence of each event alone" (Tulving, 1985, p. 393). Its occurrence requires no systematic manipulation of independent variables. Howe (in press) and Hintzman (1990) have noted that stochastic independence between measures of two processes actually represents a failure to reject the null hypothesis. There are a variety of reasons one might fail to reject the null hypothesis in addition to stochastic independence between the theoretical processes underlying two empirical indices, including insufficient statistical power (Howe, in press), an extremity response style (Bentler, 1969), asymmetrical bipolar rating scales (Meddis, 1972), rating instructions that asked subjects to rate how they felt over extended periods of time (Russell, 1979), and nonrandom error covariation stemming, for instance, from similarly worded items that are placed close together (see D. P. Green, Goldman, & Salovey, 1993). It may be useful to frame the question in terms of the separability rather than the stochastic independence of positive and negative evaluative processes.

Functional independence can exist even when stochastic independence in a given measurement context does not. Functional independence refers to "the relation between two dependent variables in a situation in which one variable does and the other one does not vary as a *function* of an independent variable" (Tulving, 1985, p. 394). The possibility that positive and negative evaluative processes (or, more specifically, two empirical indices of these theoretical processes) can be influenced differentially by some antecedent condition is inherent in the bivariate framework illustrated in Figure 1 but not in a bipolar perspective wherein effects are represented in terms of the difference between the activation of positive and the activation of negative valent processes. Establishing a functional independence between positive and negative evaluative processes does not establish definitively that these processes are distinct but only that they are separable. There is no necessary logical connection between these two kinds of independence (Tulving, 1985). However, functional independence can provide important data bearing on the mechanism(s) governing these separable processes. Two theoretical processes may have two differentiable classes of antecedents (i.e., functional independence may be established); however, if the two classes of antecedent conditions are mutually exclusive, then bipolar indices of these processes may not be particularly problematic. If, on the other hand, these two classes of antecedents are not mutually exclusive (i.e., if antecedents from each could operate simultaneously), then bipolar indices of these processes could be misleading when, for instance, trying to discriminate between a condition in which neither process was activated and a condition in which both processes were activated equivalently. Even if two different classes of antecedents were to underlay positive versus negative evaluative processes and these classes of antecedents were found to be mutually exclusive, understanding of the genesis of the resulting bipolar representations would benefit from conceptualizing evaluative processes as represented within a bivariate plane.

Stochastic independence between positive and negative evaluative processes is not likely because, as Tulving (1985) and Howe (in press) have noted, two theoretical processes are characterized by stochastic dependence as long as there is any overlap in the operating components (e.g., information, stages, processes, and mechanisms). Functional independence, on the other hand, can be established as long as the operating components are not identical, and the specific nature of the functional independence helps to illuminate what specific components differ between the theoretical processes. The bipolar representation of evaluative processes, with its emphasis on the difference between positive and negative processes, assumes that the operating components for positive and negative evaluative processes are identical or, at best, does not inspire investigation of the differences in these operating components. Thus, our focus here is on the *separability* (i.e., imperfect dependence and functional independence) of positive and negative evaluative processes rather than on their stochastic independence and, correspondingly, on the possible differences in the operating components of positive and negative evaluative processes rather than on their classification as two entirely different types or systems of evaluative processing. Specifically, we review evidence suggesting that the biological mechanisms underlying the activation of positive and negative valent classes of behavior are distinguishable, we review several areas of attitudes (e.g., racial prejudice) in which evidence for the functional independence of positive and negative evaluative processes has been reported, and we survey classic and contemporary studies of psychological states such as approach–avoidance conflict and attitude ambivalence that can be represented unambiguously in terms of the coactivation of positive and negative evaluative processes.

In light of the existence of these literatures and the potential generative utility of a bivariate conceptualization of positive and negative evaluative processes, why might the principle of reciprocal evaluative activation have proven so durable? One reason is that the bipolar attitude scale has shaped data and theory in fundamental ways. A second possible reason is that psychological or physiological processes in some conditions tend to constrain the full expression of a bivariate plane because they are incompatible or unstable. Examples of the latter are psychological states, such as conflict, ambivalence, and inconsistency among beliefs about an attitude object, which tend to be unexpected, nonharmonious, and unstable (see Abelson et al., 1968; Petty & Cacioppo, 1981, chap. 5). In a classic experiment, Brehm (1956) reported physiological evidence that choosing between two similarly attractive alternatives produced more sympathetic activation than did choosing between an attractive and an unattractive alternative. Brehm (1956) also demonstrated that choosing between two similarly attractive alternatives resulted in individuals exaggerating the positive features and discounting the negative features of the chosen alternative while discounting the positive features and exaggerating the negative features of the unchosen alternative. That is, the coactivation of positive and negative evaluative processes was evident but transient, and the subjects' subsequent representations of the attitude stimuli were organized more along the lines of bipolarity than along the lines of orthogonality. Thompson, Zanna, and Griffin (in press) have further demonstrated that ratings of positive and negative reactions to an attitude stimulus are more negatively correlated (i.e., organized in a bipolar fashion) when relevant questions are close together (enhancing the salience of subjects' inconsistent feelings) rather than separated within a questionnaire. Research on people's conceptual organization of evaluative processes and affective states may thus tend toward a bipolar structure because of the operation of motives to maintain a simple and psychologically consistent representation of the world.[3]

In sum, the important question for attitude theory at this juncture may not be whether positive and negative evaluative processes are organized in terms of a bivariate or a bipolar structure but what factors constrain a bivariate plane to approximate a bipolar vector. Even if bipolar (positive–negative) representations of the world tend to be the most stable, expected, and harmonious, understanding the genesis of these bipolar representations and the interpretation of attitudinal phenomena in general (e.g., distinguishing neutral from ambivalent attitudes) should benefit from conceptualizing evaluative processes as represented within a bivariate plane. We begin by reviewing research on the origins of bipolar attitude rating scales because understanding of attitudinal phenomena is shaped dramatically by the instruments through which they are studied. We then review evidence suggesting that a single bipolar dimension is nonoptimal to represent the positive and negative evaluative processes (i.e., valent activation functions) underlying some attitudes.

Origins of Bipolar Attitude Rating Scales

More than 60 years have passed since Thurstone (1928, 1931; Thurstone & Chave, 1929) extended the experimental methods and logic of psychophysics beyond the field of sensory discrimination to the study of affect and attitudes. Psychophysics was the first aspect of scientific psychology to develop, and research dating back to Weber in the middle of the 19th century established significant differences between the physical description of a stimulus and a person's conscious experience of that stimulus. The physical properties of a stimulus and the judgment of these properties are typically measured on quite different scales, so psychophysics has been concerned primarily with the relations between the two types of scales, one physical and one based on human judgment. Thurstone's (1928; Thurstone & Chave, 1929) important contribution to the scientific study of attitudes was not a psychophysical law that related the physical properties of a stimulus to the judged attitudinal intensity of the stimulus but his adaptation of the methods of psychophysics to measure attitudes.

In keeping with the tradition of psychophysics, Thurstone (1931) conceptualized the net difference in the affect favoring and disfavoring a stimulus as producing an affective experience or evaluative perception that established a "potential action" toward the stimulus and related stimuli in its class:

> Attitude is the affect for or against a psychological object. Affect in its primitive form is described as appetition or aversion. Appetition is the positive form of affect which in more sophisticated situations appears as liking the psychological object, defending it, favoring it in various ways. Aversion is the negative form of affect which is described as hating the psychological object, disliking it, destroying it, or otherwise reacting against it. Attitude is here used to describe the *potential action* toward the object with regard only to the question whether the potential action will be favorable or unfavorable toward the object. When we say that a man's attitude toward prohibition is negative, we have merely indicated the affective direction of his potential action toward the object. We have not said anything about the particular detailed manner in which he might act. (p. 261)

Accordingly, the attitude measure developed by Thurstone (e.g., Thurstone, 1931; Thurstone & Chave, 1929) was based on the method of equal appearing intervals, wherein a judging group sorts statements of opinion along a bipolar continuum ranging from favorable through neutral to unfavorable (cf. McGuire, 1985). As long as opinion statements and attitude objects reciprocally activate positive and negative evaluative processes, sorting the statements of opinion along this bipolar continuum is straightforward. However, this continuum may not be applicable to all statements of opinion. For instance, some opinion statements regarding cigarettes may be evaluated both positively and negatively by smokers who want to quit smoking, some opinion statements about alcohol may be regarded positively and negatively by undergraduates, and so on (e.g., see Cacioppo & Gardner, 1993; Fishbein, 1980). Kaplan (1972) suggested that individuals faced with placing these opinion statements on a bipolar continuum are compelled to redefine the scale endpoints to something like *much more favorable than unfavorable* through neutrality/ambivalence/indifference to *much more unfavorable than favorable* (see also Dawes & Smith, 1985). Thus, Thurstone's (1931) notion of what constitutes an attitude and his adaptation of the methods of psychophysics to assess attitudes are captured by the first three principles outlined in the preceding section, and exceptions to the principle of reciprocal evaluative activation produce apparent limitations in this scaling procedure.[4]

In his monograph titled "A Technique for the Measurement of Attitudes," Likert (1932) proposed a simpler bipolar scaling procedure that became known as the method of summated rat-

---

[3] Interestingly, in a recent review of the circumplex (Pleasant–Unpleasant × Low–High Activation) model of emotion, Larsen and Diener (1992) concluded that "the circumplex structure has been found with both self-report ratings and with judgments of facial expressions . . . [but] the circumplex structure needs to be further explored in non-self-report data. For example, preliminary evidence suggests that a circumplex may not emerge with EMG [electromyographic] recordings of facial muscles or with cognitive tasks that are sensitive to affective states" (p. 52).

[4] Perceptual psychophysics classically dealt with single perceptual dimensions (e.g., luminosity–brightness, wavelength–hue, and frequency–pitch) in which both the physical and psychological measures were on interval or ratio scales. Thurstone's application of this psychophysical model to the measurement of attitudes resulted in a bipolar, unidimensional construct ranging from favorable to unfavorable. This conceptualization is not inherent in psychophysical analyses, however. The field of psychophysics includes multidimensional scalings, and psychophysical investigations of perceptual phenomena such as saturation have produced bivariate scalings of frequency composition versus intensity. Thus, the present reexamination of the long shadows cast by Thurstone's conceptualization and measurements of attitudes is not at odds with a psychophysical approach to attitude measurement.

ings (Spector, 1992). In Likert's procedure, subjects are requested to check *strongly agree, agree, undecided, disagree,* or *strongly disagree* in response to each of a large number of items about an attitude stimulus. In early studies, Likert calculated the percentage responding to each alternative in an item and used the tables from the normal curve to assign scale values to each response alternative. However, results revealed a high correlation between scores obtained with this laborious procedure and scores obtained by assigning the arbitrary weightings of 1, 2, 3, 4, and 5 to the alternatives; therefore, the arbitrary weightings became more commonplace. From the large initial set of items, as many as 25 were then selected with a criterion of internal consistency, and the summed responses to these 25 items served as the attitude measure. Thus, Likert's attitude measure is not dependent on a judging group to construct the attitude scale. It is nevertheless apparent that Thurstone's conceptualization of attitudes as net affective perceptions and his scaling of attitudes using bipolar judgments laid the groundwork for Likert's method of attitude measurement as well as the subsequent development and use of bipolar rating scales.

Thurstone and Likert scales differ in their sensitivity over the continuum from *maximally negative/not at all positive* to *maximally positive/not at all negative,* but exceptions to the principle of reciprocal evaluative activation are problematic for both scaling procedures. For example, when items included in a Likert scale are rescaled by the method of equal appearing intervals, few items fall over the middle of the scale continuum, a result that appears attributable to the criterion used for selecting items in the Likert method (A. I. Edwards, 1946). Criterion groups, in terms of low and high total scores, are selected, and items are deleted if they do not differentiate these groups. A. I. Edwards (1946) demonstrated that Thurstone's "neutral" items were nondifferentiating; thus, these items were simply not included in attitude scales developed by the method of summated ratings. In addition, items that do not adhere to the principle of reciprocal evaluative activation of positive and negative evaluative processes tend to be excluded from Thurstone scales because they fail to meet the ambiguity or irrelevance criteria in the method of equal appearing intervals, and they tend to be excluded from Likert scales because they fail to meet the internal consistency criterion in the method of summated ratings. When items that reflect attitude ambivalence or indifference are included, they are scaled as if they reflect neutral attitudes. Thus, the item "There is much to be said on both sides of the censorship problem" carried a scale value of 5.5 (i.e., the midpoint) on the Thurstone scale designed to measure attitudes toward censorship (Thurstone & Chave, 1929).

Klopfer and Madden (Klopfer & Madden, 1980; Madden & Klopfer, 1978) have also found that the middlemost options on Thurstone and Likert attitude scales are ambiguous in terms of their psychological significance, despite the common practice of interpreting these options as reflecting a neutral attitude. Klopfer and Madden (1980) studied this issue by testing four different groups of subjects, with the middlemost option explicitly defined as appropriate for sentiments of ambivalence to one group, for sentiments of neutrality to another group, for sentiments of uncertainty to a third group, and for cannot-decide responses to a fourth group. Subjects completed two scales devised by Thurstone: the Attitude Toward Capital Punishment and the Attitude Toward Sunday Observance scales. Frequency of the use of the middlemost option was tabulated for each subject within each group. No differences were found for the topic of Sunday observance, which Klopfer and Madden attributed to the topic being dated. Results for the topic of capital punishment, however, indicated that the ambivalence definition of the middlemost option evoked significantly more use of this option than the uncertainty definition, and the neutrality and cannot-decide definitions produced frequencies of usage that were intermediate, indicating that the middlemost option could reflect various psychological reactions to the attitude stimulus ranging from neutrality to ambivalence. This result also is consistent with the notion that positive and negative evaluative processes are not always activated reciprocally by attitude objects.

The methods of constructing Thurstone and Likert scales and the subsequent development and validation of simpler bipolar rating scales (McGuire, 1985; Summers, 1970) resulted in the widespread use of measures sensitive to variations in attitudes arising from the reciprocal activation of positive and negative evaluative processes but insensitive to variations in attitudes arising from nonreciprocal modes of evaluative activation. An early experiment conducted by J. D. Edwards and Ostrom (1971) is illustrative. These researchers investigated attitude ratings of "neutral attitudes" that derived from (a) all neutral experiences with an attitude object or (b) a wide range of positive, neutral, and negative experiences. In the present framework, the former experiences would produce a neutral attitude and a middling rating on a bipolar attitude scale because neither positive nor negative evaluative processes were activated, whereas the latter, more varied experiences would produce a more ambivalent attitude and a middling rating on a bipolar attitude scale because of the coactivation of positive and negative evaluative processes. J. D. Edwards and Ostrom (1971) found that although the different substrates of neutral attitude ratings were apparent in their measure of cognitive structure, these attitude differences were not evident on Fishbein and Raven's (1962) five semantic differential scales (*bad–good, harmful–beneficial, wise–foolish, sick–healthy,* and *dirty–clean*) or an 8-point self-rating scale ranging from *highly unfavorable* to *highly favorable* (Anderson, 1965). Thus, the principle of reciprocal evaluative activation may not be sufficient to simply and comprehensively capture the patterns of evaluative activation underlying attitudes.

Attitude rating scales have not been held to be flawless, of course, and this early research underscored some of the limitations of bipolar rating scales. The fragility of self-report measures has been trumpeted every decade or so (e.g., Cook & Selltiz, 1964; Gutek, 1978; Turner & Martin, 1984), and recent research on the cognitive psychology of rating scales (e.g., survey questions) has emphasized the measurement reactivity of self-report attitude scales. For instance, subtle differences in the wording of questions can affect an individual's response to attitudinal scales (e.g., Fischoff, 1991; Loftus, Fienberg, & Tanur, 1985; Poulton, 1989; Tourangeau & Rasinski, 1988; Turner, 1984) and can have unintended effects on the individual's attitudes (e.g., Turner & Krauss, 1978; Wilson & Schooler, 1991). Alternative methods for exploring attitudes have been considered, including nonverbal, projective, and indirect behavioral measures (Campbell, 1950; Cook & Selltiz, 1964; Rosenthal &

Rosnow, 1969); autonomic assessments (see reviews by Cacioppo & Sandman, 1981; Zanna, Detweiler, & Olson, 1984); event-related brain potentials (Cacioppo, Crites, Berntson, & Coles, 1993); and the bogus pipeline procedure (Jones & Sigall, 1971). However, the bipolarity inherent in Thurstone and Likert attitude scales has characterized most of these assessments as well. Autonomic assessments, which are more sensitive to attitude extremity than attitude valence, are an exception to this rule (Cacioppo, Petty, & Geen, 1989; Zanna et al., 1984). They are also the exception that makes the rule, because they have been criticized over the years for failing to differentiate positive from negative attitudes along a single response vector (e.g., Summers, 1970).

Among the appeals of the ubiquitous bipolar rating scales in attitude research are their economy and practicality, the ethics of their application, and the concordance between common sense and attitudes as indexed by these rating scales (Dawes & Smith, 1985; Ostrom, 1973). The literature on attitude–behavior correspondence (e.g., Cacioppo, Petty, Kao, & Rodriguez, 1986; Fazio, Powell, & Williams, 1989; Fishbein & Ajzen, 1975) further supports the scientific value of conceptualizing attitudes as an individual's net positive or negative feelings about or evaluative perception of a stimulus or category of stimuli. Thus, a voluminous empirical literature and a field rich in theory exist in large part because of the use of bipolar, self-report attitude scales (e.g., see McGuire, 1985; Schuman & Kalton, 1985). Does the existence of this copious literature argue against the notion that bipolar rating scales have artificially bounded attitude theory and research? We suggest not, at least no more than the multitudinous fish cataloged by Sir Eddington's hypothetical scientist argue against the 2-in. mesh net distorting theory and research on what constitutes fish in the sea.

The Principles of Evaluative Activation, Opposing Evaluative Actions, and Bivalent Modes of Evaluative Activation

One line of evidence for the potential utility of a bivariate conceptualization of positive and negative evaluative processes relies neither on analyses of self-ratings nor on analyses of the language of emotion; instead, it comes from research in animal learning and psychobiology. All organisms have one or more mechanisms for differentiating hostile from hospitable environments (Martin & Levey, 1978). The unconditioned reflexes represent one such set of innate mechanisms. Konorski (1948, 1967) classified unconditioned reflexes as falling into one of two classes: appetitive (e.g., ingestion, copulation, and nurturance of progeny) and aversive (e.g., withdrawal from or rejection of noxious agents). Konorski further noted that both appetitive and aversive unconditioned reflexes could vary in their level of activation and, thus, were orthogonal to the activity or arousal dimension underlying behavior (see also Mackintosh, 1983). Masterson and Crawford (1982) and Schneirla (1959) similarly summarized behavioral evidence for distinguishing between approach and withdrawal reflexes or behaviors, with negatively valent behaviors organized in terms of an "aversive" or "defensive motivational system" and positively valent behaviors organized in terms of an "appetitive motivational system":

> In general, what we shall term the A-type of mechanism, underlying approach, favors adjustments such as food-getting, shelter-getting, and mating; the W-type, underlying withdrawal, favors adjustments such as defense, huddling, flight, and other protective reactions. Also, through evolution, higher psychological levels have arisen in which through ontogeny such mechanisms can produce new and qualitatively advanced types of adjustment to environmental conditions. (Schneirla, 1959, p. 4)

Despite these distinctions, an intuitive attraction of bipolar conceptualizations of attitudes and rating scales is the apparent bipolarity that characterizes much of people's behavior. Individuals approach, acquire, or ingest certain classes of stimuli and withdraw from, avoid, or reject others. Many behaviors that do not fall at these extremes can nevertheless be placed along an approach–withdrawal continuum. Individuals may give a small or large contribution to support a political candidate or may perform few or many tasks that are beyond the call of duty because of the way they feel about (i.e., their attitude toward) a person or post. Thus, behavior is often constrained by the mutual exclusivity of the various options.

Although behavioral constraints may also be manifest in bipolar attitudinal expressions and behavioral dispositions (Fishbein & Ajzen, 1975), these constraints do not have the same force at the level of underlying mechanism. Studies of behavioral conflict and attitude ambivalence, for instance, have shown that an individual's position on this approach–withdrawal continuum can be insufficient to identify important differences in psychological mechanisms and behavioral predisposition. Recall our discussion in the previous section of J. D. Edwards and Ostrom's (1971) study of neutral attitude ratings that derived either from neutral experiences with an attitude object or from a wide range of positive, neutral, and negative experiences. Despite the similarity in subjects' responses on bipolar scales to these attitude stimuli, very different substrates were found to underlay these attitudes when belief structures pertaining to the attitude target were examined. Indeed, neutral, indifferent, and ambivalent attitudes may all have similar manifestations on bipolar attitude scales despite the differences in the evaluative substrates and behavioral ramifications of each.

An especially important type of data arguing for a bivariate framework is evidence that both positive and negative valent forces can be activated by the same manipulation. The classic research conducted by Miller (1948, 1959, 1961, 1966; see also Amsel, 1990) on conflict provides evidence not only that positive and negative evaluative processes are separable in their influence on approach–withdrawal behavior but that environmental contingencies can produce a strong coactivation of each. Miller's conflict theory is illustrated in Figure 2. In an early test of the basic assumptions underlying this theory, Brown (1948) trained one group of hungry albino rats to run down a short alley to secure food, and he trained a second group of rats in the same alley to avoid the site at which they received an electric shock. Each animal wore a small harness connected to a recording device to determine the strength of pull when the animal was stopped in the alley. Brown found that (a) the animals that were stopped near the food pulled harder toward the food than those stopped farther from it, thereby establishing an approach (positivity) gradient; (b) the animals that were stopped near the place they had been shocked pulled harder away from the shock

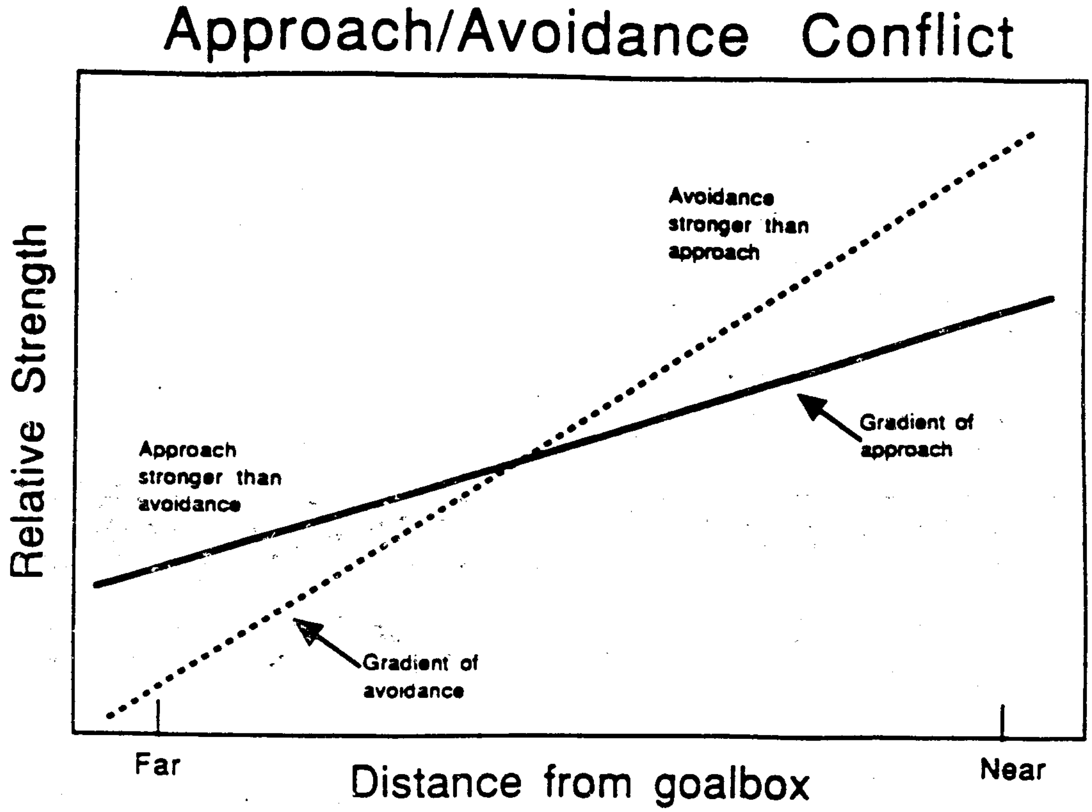

Figure 2. Graphic representation of the dynamics of an approach–avoidance conflict. From "Liberalization of Basic S-R Concepts: Extensions to Conflict Behavior, Motivation and Social Learning" by N. E. Miller, 1959. In S. Koch (Ed.), *Psychology: A Study of a Science. Study 1* (p. 206). New York: McGraw-Hill. Copyright 1959 by McGraw-Hill. Adapted by permission. Miller (1959) depicted the gradients as straight lines for didactic reasons but noted that similar predictions could be made with a variety of curvilinear gradients.

than those stopped farther from it, thereby establishing an avoidance (negativity) gradient; (c) the strength of pull for avoidance increased more rapidly with nearness than did that for approach, indicating that the avoidance gradient was steeper than the approach gradient; (d) increasing the hours of food deprivation increased the strength of the pull toward the goal box both near and far from it, indicating that the approach gradient was elevated as a whole rather than made more steep; (e) increasing the strength of the electric shock increased the pull away from the goal box near and far from it, indicating that the avoidance gradient was also elevated as a whole; and (f) the manipulations of approach and avoidance did not interact in determining the strength of the pull. Miller (e.g., Miller, 1959) further posited and found that when two incompatible responses were in conflict, the stronger dominated (although vacillation was also evident). Thus, the manipulation of positive and negative environmental contingencies in the same situation consistently had opposing effects on approach–withdrawal behaviors. However, the fact that the actions of positive and negative evaluative processes are distinguishable and opposing does not necessarily imply that these processes are reciprocally activated. Therefore, Miller's key insight, that environmental contingencies can arouse approach–avoidance conflict, is particularly important because it suggests that positive and negative substrates underlying behavior can be nonreciprocally activated. Indeed, the gradients in Figure 2 indicate that both the strength of approach and the strength of avoidance (or withdrawal) increase as the distance from the goal decreases; that is, there is a coactivation of positive and negative valent processes as distance from the goal decreases.

In addition to behavioral data, evidence from the neurosciences is increasingly in accord with the partial independence of positive and negative evaluative mechanisms or systems (Berntson, Boysen, & Cacioppo, in press; Gray, 1987, 1991).[5] The notion dates back at least to the experimental studies of Olds (1958; Olds & Milner, 1954), who spearheaded a literature identifying separate neural mechanisms to be related to the subjective states of pleasure and pain. Gray (e.g., 1990, 1991) has proposed that the septohippocampal system, together with closely related structures such as central noradrenergic and serotonergic fibers ascending from the locus coeruleus and raphe nuclei, respectively, plays a central role in fear and anxiety in animals. Reiman, Raichle, Butler, Hersovitch, and Robbins (1984) used

[5] As Gray (1990) noted, the idea that there exist separable subsystems of the brain mediating particular classes of affect and emotion is inherently open to criticism because "any part of the brain put into one subsystem is in fact also connected to myriad others" (p. 274). But the same criticism applies to the concept of a visual system, even though few now would contest the scientific utility of this concept.

positron emission tomography (PET) scanning to compare patients who suffered from spontaneous panic attacks and normal controls. Consistent with Gray's theory, the PET scans from the two groups differed in only one brain region, the area containing the major neocortical inputs to the septohippocampal system (the entorhinal area) and its major output (the subicular area). Gray has termed this brain subsystem the behavioral inhibition system, which, at the behavioral level, organizes reactions to conditioned stimuli associated with punishment or the termination or omission of reward.

Gray (1991) also posited the existence of what he termed the "approach system," so named because it is responsive to conditioned appetitive stimuli. In Gray's (1991) theory, the specific response to a reward differs dramatically depending on the particular reward (e.g., food, sex partner, or omission of shock),

> but the successful gaining of rewards would not normally differ in this way as a function of the nature of the reward concerned. Thus, it is reasonable to suppose the existence of a single system for approach to rewards, whatever these may happen to be. (p. 282)

Although the neurological basis of this system is not yet fully delineated, an important role appears to be played by the dopaminergic fibers that ascend from the substantia nigra and the ventral tegmental area and innervate regions in the basal ganglia, neocortex, and limbic system; the dopaminergic projections to the nucleus accumbens appear to be particularly important in mediating incentive motivation (see also Hoebel, 1985, 1988).

Davidson (1992, in press) reviewed evidence that approach- and withdrawal-related systems are localized in different cerebral hemispheres, with the left frontal region implicated in approach-related emotional behavior and the right frontal region implicated in withdrawal-related emotional behavior. For instance, Davidson and his colleagues have found that individual differences in the relative activation of the right and left frontal electroencephalogram (EEG) are related to dispositional differences in positive and negative moods (Tomarken, Davidson, Wheeler, & Doss, 1992) and that relative right versus left frontal EEG activity varies as a function of positive and negative events (e.g., Ahern & Schwartz, 1985; Davidson, Ekman, Saron, Senulis, & Friesen, 1990; Fox, 1991). More important in the present context, Davidson and his colleagues have reported electrocortical data bearing on the separability of positive–approach and negative–withdrawal substrates, including that (a) the experimental arousal of disgust and fear is associated with an elevation of right frontal activation relative to either a nonemotional baseline or a positive emotional state (Davidson et al., 1990), (b) left frontal EEG activity is diminished in depression (Henriques & Davidson, 1990, 1991; Schaffer, Davidson, & Saron, 1983), (c) toddlers whose temperamental style is characterized by reticence about approaching novel and unfamiliar people and objects show lower left frontal activation than toddlers who display an uninhibited temperamental style (Davidson, Finman, Straus, & Kagan, 1992), and (d) the administration of diazepam, a benzodiazepine that increases approach-related behavior in novel and unfamiliar settings, increases left frontal activation in rhesus monkeys (Davidson, Kalin, & Shelton, in press).

Although the physiological bases of positive and negative evaluative processes appear to be distinguishable, interactions reminiscent of reciprocal activation have also been identified. Lang, Bradley, and Cuthbert (1990) posited, for instance, that an individual's affective state could modulate the magnitude of appetitive and aversive (defensive) reflexes. Lang et al. (1990) focused on reflexes because they are obligatory responses to environmental events whose strength varies with psychological factors. The notion that a prevailing affective state would augment the strength of evaluatively congruent reflexes and would dampen the strength of evaluatively incongruent reflexes followed from the premise that affective states represented action dispositions. In examinations of their hypothesis, the magnitude of the startle eyeblink (a defensive reflex) has been measured during positive, neutral, and negative affective states. Results from several laboratories have shown that the startle eyeblink is augmented during the presentation of negative (e.g., fear-arousing) stimuli and is diminished by the presentation of positive stimuli (e.g., Bradley, Cuthbert, & Lang, 1990; Hamm, Stark, & Vaitl, 1990; Vrana, Spence, & Lang, 1988). Results using this paradigm have not been singular, however. McManis and Lang (1992), for instance, recently investigated the utility of Tellegen's (1985) Positive and Negative Affect Schedule (PANAS) dimensions in organizing psychophysiological responses (including the startle eyeblink) to emotional pictures. Pictures were selected from the International Affective Picture Show (Lang, Ohman, & Vaitl, 1988) such that the set of experimental stimuli varied in terms of positivity and negativity. Each slide was presented for 6 s while psychophysiological recordings were measured, and a startle probe (50 ms, 95-dB white noise burst) was presented randomly during the 6-s slide presentations. McManis and Lang (1992) found that positive affect diminished the magnitude of the startle blink, whereas negative affect was unrelated to blink magnitude (although directionally the effect of negative affect was to inhibit the startle blink, an effect opposite in direction to what would be expected on the basis of reciprocal activation). Additional research is needed to determine precisely the conditions that produce an apparent bipolar structure of data (e.g., fear–anxiety rather than negative affect per se may be important) and the conditions that foster bivariate manifestations.

Another distinct line of evidence for the separability of positive and negative evaluative processes comes from research on racial prejudice. Patchen, Hofman, and Davidson (1976) report a factor-analytic study of the racial attitudes of more than 4,000 desegregated Black and White high school students. Their results indicated that positive and negative racial attitudes were not merely opposite ends of a continuum. Rather, one factor was characterized as a general positive evaluation of Blacks and Whites, and an orthogonal factor was characterized by a general negative evaluation of Blacks and Whites.

The separability of positive and negative evaluative processes underlying racial attitudes appears to be attributable to the activation of positive and negative valent processes by different social and environmental factors. Patchen, Davidson, Hofman, and Brown (1977) found that negative interracial behaviors (e.g., avoidance and unfriendly interaction) were predicted best by variables such as personal aggressiveness, initial racial attitudes, and racial attitudes of family and peers, whereas positive interracial behaviors were predicted best by opportunities for interracial contact before entering and within high school. Lipset and Schneider (1978) further posited that positive sentiments toward Blacks among the White majority were related to the value of egalitarianism, whereas negative sentiments toward Blacks were related to the value of individualism. Inconsistencies in the White majority's reactions to busing and related civil rights issues, according to Lipset and Schneider (1978), reflected the conflict between these two strongly held and opposing values. Consistent with this reasoning, Lipset and Schneider (1978) found that public attitudes were supportive of civil rights when egalitarian values were at stake (e.g., equal opportunity and antidiscrimination laws) but were more mixed when individualism was threatened (e.g., busing and affirmative action statutes).

Katz, Wackenhut, and Hass (1986; Katz & Hass, 1988) have also reported research suggesting that positive and negative feelings toward Blacks by the White majority are separable and can be activated reciprocally, separately, or nonreciprocally. Katz and Hass (1988), for instance, developed a scale to assess positive attitudes toward Blacks (Pro-Black attitude scale) and a scale to assess negative attitudes toward Blacks (Anti-Black attitude scale). The Pro-Black scale contains items measuring friendly sentiments toward Blacks as a minority underdog (e.g., "Many Whites show a real lack of understanding of the problems that Blacks face"). The Anti-Black scale, in contrast, contains items measuring criticisms of Blacks (e.g., "One of the biggest problems of a lot of Blacks is their lack of self-respect"). Hass, Katz, Rizzo, Bailey, and Eisenstadt (1991; see also Hass, Katz, Rizzo, Bailey, & Moore, 1992) provided evidence that not only do some individuals score low on one and high on the other (consistent with the notion of reciprocal evaluative activation), but some individuals score low on both scales and others score high on both scales (consistent with the notion of the separability of positive and negative evaluative processes and bivalent modes of evaluative activation).

D. P. Green and his colleagues (e.g., Green, 1988; Green & Citrin, in press; Green et al., 1993) have argued that greater attention needs to be given to measurement issues such as response sets because they can contribute artifactually to the appearance of independence between the positive and negative substrates of racial attitudes. In an interesting analysis of data on the public perceptions of Hispanics, Green and Citrin (in press) demonstrated that nonrandom measurement error in survey questions (e.g., error attributable to the common format and response options) can diminish the correlation between positively and negatively valenced statements. Among the methodological aids for detecting nonrandom error recommended by Green and Citrin (in press) is a multimethod approach to the construction of attitude measures (Campbell & Fiske, 1959).

Although these methodological issues are important and bear on how one might measure positivity and negativity, starting with the assumption that the bivalent processes underlying attitudes are invariably bipolar can also be costly in terms of the fertile avenues of research it precludes. On the basis of research suggesting that positive and negative interracial attitudes and behaviors are related to different factors, for instance, Schofield (1989, 1991) reasoned that practices designed to increase positive interaction among Black and White students may not substantially decrease negative interaction. Schofield's hypothesis is provocative because it suggests that theory and research on promoting positive interracial attitudes and behaviors may never completely illuminate the origins of negative interracial attitudes and behaviors. Thus, the importance of questioning the guidance provided by the principle of reciprocal evaluative activation is not merely academic and is not limited to attitude measurements; whether or not the bivalent processes underlying racial attitudes are invariably bipolar also has significant implications for how to best achieve attitude and behavior change. If the bivalent processes underlying racial attitudes are characterized by bipolarity, the proposed bivariate conceptualization reduces to a bipolar model and nothing is lost. If, however, these processes are not captured fully by bipolarity, the bivariate framework provides a superior heuristic for representing the processes and for guiding research on their unique antecedents and consequences.

Prior research on the predictors of people who donate blood or organs or join a bone-marrow registry, along with the reasons they do so, is also consistent with the notions that (a) positive and negative evaluative processes have distinguishable antecedents and consequences and (b) both positive and negative evaluative processes can be coactivated (e.g., Parsi & Katz, 1986; see Cacioppo & Gardner, 1993). Social influences (e.g., friends and peer pressure) are relatively important for initial donors, whereas internalized mechanisms (e.g., past experience and perception of oneself as a donor) become important for repeat donors (e.g., Callero & Piliavin, 1983; Charng, Piliavin, & Callero, 1988; P. W. Edwards & Zeichner, 1985; McCombie, 1991; Paulhus, Shaffer, & Downing, 1977; see also Callero, Howard, & Piliavin, 1987; Sarason, Sarason, Pierce, Shearin, & Sayers, 1991). Moreover, there appears to be a developmental process in which the perception of oneself as a donor occurs in some individuals as donor behaviors unfold. Thus, Charng et al. (1988) found that behavioral intentions were predicted by subjective norms and attitudes in novice donors but that the prediction of the behavioral intentions of repeat donors was improved by adding measures of "role identity." Of course, most people never become donors, and most first-time donors never become experienced donors. This is so obvious that it is easy to forget that many nondonors hold positive attitudes toward donor behaviors and realize the need for and value of blood and organ donations (Pomazal & Jaccard, 1976). Why are so few people with positive attitudes toward donor behaviors themselves donors? The answer may lie in the findings that the positive and negative evaluative processes underlying donor attitudes and behaviors are separable and that the negative substrate tends to be an overlooked impediment.

Briggs, Piliavin, Lorentzen, and Becker (1986), for instance, found that the greater an individual's perceptions of personal risk in donating bone marrow, the less willing he or she was to join a bone-marrow registry. In addition, several studies have shown that beliefs regarding the negative (rather than positive) consequences of donating differentiate individuals who intend versus do not intend to become (or continue as) blood donors (e.g., Condie, Warner, & Gillman, 1976; P. W. Edwards & Zeichner, 1985; Obone & Bradley, 1975; Oswalt & Napoliello, 1974; Pomazal & Jaccard, 1976). Relatedly, Parsi and Katz (1986) developed an attitude toward posthumous organ donation scale that consisted of two subscales, Positive (prodonation) and Negative (antidonation). Responses to the scale by a sample of 110 participants, most from lower to middle management positions in several financial corporations in the Wall Street area of New York, revealed that the Positive and Negative subscales were uncorrelated. Parsi and Katz also found that subjects who had strong positive attitudes and weak negative attitudes were the most likely to sign donor cards but that only when the score on the Negative subscale was low did it make a difference whether the score on the Positive subscale was high or low.

The evidence for the separability of positive and negative evaluative processes becomes more controversial when one turns to the literature on the conceptual organization of moods, affect, and emotion. Among the best-known research bearing on the centrality of people's net positive and negative feelings is Osgood, Suci, and Tannenbaum's (1957) classic work on the measurement of meaning. In multiple studies and cultures, evaluative bipolar word pairs (e.g., *pleasant–unpleasant, good–bad, positive–negative,* and *harmful–beneficial*) were found to constitute a fundamental dimension underlying people's understanding of the world. Conceptually similar results have been found in cross-cultural, multidimensional scaling studies of people's conceptual organization of emotional feelings (e.g., Russell, 1980; Russell, Lewicka, & Niit, 1989) and in cross-cultural ratings of emotionally evocative pictures (e.g., Greenwald, Cook, & Lang, 1989; Lang, Greenwald, Bradley, & Hamm, 1993; Lang et al., 1988). Thus, the two-dimensional representation that best represents people's conceptual organization of affect and emotion may be Positive–Negative × Active–Inactive rather than Positive–Nonpositive × Negative–Nonnegative. As noted earlier, however, one possible reason for this organization is that psychological states such as conflict, ambivalence, and inconsistency among beliefs about an attitude object tend to be unexpected, nonharmonious, and unstable (Abelson et al., 1968; Brehm, 1956). Research on people's conceptual organization of evaluative processes and affective states (e.g., moods) may thus tend toward a bipolar structure because of the operation of motives to maintain a simple and psychologically consistent representation of the world.

A second reason, articulated by R. F. Green and Goldfried (1965), is that the bipolar attitude scale constrains the expression of the separable manifestations of positive and negative evaluative processes:

> The assumption that semantic space is bipolar implies that an individual's mediational responses to signs occur along a series of dimensions which have reciprocally antagonistic verbal opposites at their polar points. The assumption of reciprocal antagonism, like that of bipolarity, is built right into the measuring instrument itself and consequently leaves no room for disconfirmation. (p. 2)

R. F. Green and Goldfried (1965) constructed 58 single-adjective unipolar scales (e.g., *pleasant, good, heavy, sharp,* and *unpleasant*) and asked subjects to describe various concepts using each scale. The adjectives were analyzed in two subsets: the 15 pairs from Osgood's semantic differential that were judged most likely to be bipolar and 14 from Osgood's semantic differential that were judged the least likely to be bipolar. R. F. Green and Goldfried's methodology was not optimal by current standards, but, to the extent that problems such as asymmetrical scaling, nonrandom error, or extremity response biases were operating in this study, these artifacts should have attenuated all of the correlations between all logical opposites. Results, however, revealed that some of the logical opposites rated on these unipolar scales showed consistent and strong negative correlations (e.g., *strong–weak, clean–dirty,* and *fast–slow*),[6] whereas others (e.g., *pleasant–unpleasant*) were not significantly negatively (or positively) correlated with each other across attitude stimuli. Results of factor analyses also suggested that when subjects were not forced to make judgments on bipolar scales, they tended to use the scales as if they reflected unipolar continua rather than simply one end of a bipolar continuum. R. F. Green and Goldfried called into question Osgood et al.'s (1957) assumption that bipolar continua along which people organize the connotative meaning of concepts are intrinsic to their use of language (see also de Mille, 1970).

Another line of evidence bearing on the conceptual organization of affect and emotion comes from recent studies of self-reports of moods and emotions (e.g., Bradburn, 1969; Bradburn & Caplovitz, 1965; Diener & Emmons, 1985; Diener & Iran-Nejad, 1986; D. P. Green et al., 1993; Larsen & Diener, 1992; Tellegen, 1985; Watson, 1988; Zveon & Tellegen, 1982), research on preferences related to gains and losses (Fischoff, Slovic, & Lichtenstein, 1980; Kahneman & Tversky, 1979), and factor analyses of affect terms (e.g., Gotlib & Meyer, 1986). Studies of the language of moods and emotions, for instance, have led to the claim of stochastic independence between positive and negative moods; terms such as *delighted, excited, alert,* and *determined* load on a high positivity factor, whereas terms such as *nervous, afraid, angry, guilty, contemptuous,* and *disgusted* load on a high negativity factor (see reviews by Watson, Clark, & Tellegen, 1988; Watson & Tellegen, 1985).

Whether or not these data actually support the notion of stochastic independence between people's positive and negative conceptual organizations of moods and emotions is still being debated (e.g., D. P. Green, 1988; D. P. Green et al., 1993; Larsen & Diener, 1992; Watson, 1988). As D. P. Green et al. (1993) and D. P. Green and Citrin (in press) have recently demonstrated, evidence purportedly showing the independence of opposite mood states—low correlations between positive and negative moods—can result from random and nonrandom response error. D. P. Green et al. (1993) further demonstrated that when multiple methods of mood assessment were used and response errors were considered, a bipolar (positive–negative) structure underlying moods emerged. Zanna and Thompson (1991) have also suggested that methodological artifacts (e.g., carryover between unipolar positive and negative rating scales) can inflate the negative correlation between positive and negative evaluative processes. These methodological considerations will need to be considered in future work on the stochastic independence

[6] R. F. Green and Goldfried's (1965) studies do not call into question the centrality of evaluative processes in people's conceptual organization of their world as much as they call into question the specific dimensional structure of evaluative processes advanced by Osgood, Suci, and Tannenbaum (1957). For instance, some of the words examined by R. F. Green and Goldfried (1965) were found to be organized in terms of bipolarity, suggesting that their failure to find bipolarity for many of the evaluative words is not easily explained as an artifact of their psychometric procedures.

between positive and negative moods.[7] A bivariate model can accommodate any range of outcomes, whereas a bipolar model can accommodate only one. It will be equally illuminating to determine the conditions under which a potentially bivariate plane manifests as a bipolar (e.g., pleasantness–unpleasantness) continuum and to identify the overlapping and nonoverlapping operating components (e.g., information, stages, processes, and mechanisms) that are responsible for the (in)dependence of positive and negative evaluative processes.

In an intriguing study that bears on functional rather than stochastic independence, Goldstein and Strube (in press) demonstrated the separability of positive and negative affect within a specific situation and time and the uncoupled activation of positive and negative processes after success and failure feedback, respectively. Briefly, Goldstein and Strube (in press) administered Watson et al.'s (1988) PANAS to students at the beginning and end of three successive class meetings. Students were led to believe that they were participating in a large-scale longitudinal study of affect, and they were instructed to indicate on the PANAS how they felt "right now (that is, at the present moment)." However, these class meetings were selected by the investigators such that, on the second day of the study, students received performance feedback on a class examination. Students were informed of their performance and of the class average on the exam. Analyses of the responses to the PANAS indicated that students who scored above the mean (i.e., success feedback) showed elevations in positive affect but no change in negative affect, whereas students who scored below the mean (i.e., failure feedback) showed elevations in negative affect but no change in positive affect. Goldstein and Strube (in press) also found nonsignificant correlations within each measurement period between the measures of positivity and negativity. Thus, not only were the momentary feelings of positivity and negativity independent and the activation of positive and negative evaluative processes uncoupled, but different events (success and failure feedback, respectively) were associated with the uncoupled activation of positive and negative evaluative processes. These results are consistent with the notion that the principle of reciprocal evaluative activation can be usefully subsumed by the more general principle of bivalent modes of evaluative activation.

In summary, evaluative processes have historically been depicted in attitude measurement and theory in terms of a bipolar vector extending from maximally negative (and minimally positive) to maximally positive (and minimally negative). If the activation of positive and negative evaluative processes is inevitably reciprocally coupled, then detailing separable positive and negative substrates would not enhance prediction of the resulting evaluative response dispositions. However, theory and research from areas ranging from the conceptual organization of emotion through racial prejudice to unconditioned responses indicate that positive and negative evaluative processes have some nonoverlapping operating components (e.g., functional independence), are opposing in their effects on attitudes or behavior, and are capable of being differentially activated. Accordingly, we have suggested that the bivariate framework may be preferable to the bipolar framework, at least at this juncture, because the former has generative utility. If attitudes are truly bipolar, then the bivariate and bipolar models are formally equivalent, and the attitude dimension lies along the reciprocal diagonal of the bivariate plane. To whatever extent attitudes deviate from a strict bipolarity, the bivariate representation is more accurate and comprehensive.[8]

The theoretical utility of a bivariate framework is manifest in the social psychological literature on racial prejudice and on blood–organ donations, in which positive and negative attitude factors have been proposed, and in the psychobiological literature on motivation, in which some of the neural mechanisms underlying appetitive and defensive actions have been differentiated. In the next section, we introduce a computational model of an evaluative space. The purpose of the model is not to celebrate mathematical formalisms per se but to define relations among theoretical constructs and to illustrate the generative value of the bivariate framework. Given the metrics commonly used in attitude research, the point predictions made by the model are not as critical as are the relative effects predicted. In developing and illustrating the model, we draw especially on the research on behavioral conflict and on attitude ambivalence, phenomena that underscore the limitations of the principle of reciprocal evaluative activation and the heuristic value of the principle of bivalent modes of evaluative activation.

Attitudes and Evaluative Space

The conceptualization of an attitude represented by the principles of evaluative activation, oppositive evaluative actions, and

---

[7] The research on the separability of positivity and negativity using the PANAS has focused on moods (e.g., Watson, 1988) and dispositions (e.g., Marshall, Wortman, Kusulas, Hervig, & Vickers, 1992) rather than on attitudes. Are the findings from this area of research relevant to the study of attitudes? As a preliminary inquiry, 30 undergraduate women living in campus residence halls were interviewed to assess the positivity and negativity evoked by their roommate (Snydersmith & Cacioppo, 1992). The PANAS was administered with instructions that subjects should respond to the items in terms of how they felt about their roommate. The responses to the nine items constituting the positivity scale were summed to yield an index of the positivity evoked by the roommate, and a comparable procedure was used to index the negativity evoked by the roommate. The Cronbach alphas for both scales were high (.77 for negativity and .90 for positivity). Analyses further revealed that (a) roommates evoked stronger feelings of positivity (M = 27.3) than negativity (M = 15.63) and (b) positivity and negativity toward the roommate were uncorrelated ($r = -.08$, *ns*). Research on the separability of positivity and negativity in moods and investigations of the conditions under which stochastic independence between positivity and negativity occurs may therefore bear on the operating components of positive and negative evaluative processes, at least for some attitudes.

[8] Previous multidimensional models of attitudes have emphasized components such as affect, behavior, and cognition, with each scaled along a bipolar evaluative continuum (e.g., Breckler, 1984; see Cacioppo, Petty, & Geen, 1989). The present model is not incompatible with these perspectives; it implies only that to understand the attitudinal significance of any such set of components, the effects of each on positive and on negative evaluative processes should be considered to ensure that the overall level of positivity and negativity activated by the attitude stimulus is gauged accurately. It is interesting to note in this regard that Zanna and Thompson (1991) gauged positivity and negativity within specific attitudinal components (e.g., affective and cognitive) and subsequently aggregated across components as a result of the similarity in data patterns.

bivalent modes of evaluative activation can be expressed quantitatively as follows:

$$\text{Attitude} = \frac{W_p}{W_p + W_n} Pi - \frac{W_n}{W_p + W_n} Nj + Iij, \qquad (1)$$

where P is the activation function for positive evaluative processes (abbreviated hereafter as positivity), Pi represents the level of positivity activated by an attitude object, and W_p is a weighting factor representing the relative attitudinal effect of variations in positivity; N is the activation function for negative evaluative processes (abbreviated hereafter as negativity), N_j represents the level of negativity activated by an attitude object, and W_n is a weighting factor representing the relative attitudinal effect of variations in negativity; and Iij represents nonadditive effects (e.g., the potentially mutually inhibitory interaction between positivity and negativity as their mutual activation increases). The weighting factors are scaled such that the weighting of positivity and of negativity can vary from zero to one, and the weights of each sum to one. The different modes of evaluative activation manifest as changes in the activation indices, i and j. Reciprocal activation of positivity and negativity occurs, for instance, when an attitude object produces an increase in Pi and a decrease in Nj; uncoupled activation occurs when the attitude object affects only Pi or Nj; and nonreciprocal activation occurs when the attitude object evokes increases (or decreases) in both Pi and Nj (Cacioppo, 1991).

Equation 1 makes it possible to derive an attitudinal surface for all combinations of Pi and Nj. The exact shape of the attitudinal surface is dependent on the form of the activation functions, the relative value of the weighting coefficients, and the scaling of the axes. For purposes of illustration, the weights in Equation 1 were assumed to be equal, the activation functions for positivity and for negativity were assumed to have an exponent of .5 (e.g., power functions representative of psychophysical functions), and the interaction and error terms were assumed to be zero.[9] The resulting attitudinal surface overlying the evaluative plane is depicted in Figure 3. This overlying surface represents the attitudinal disposition predicted for each coordinate within the evaluative plane, where the z-axis represents the attitudinal manifestation on a bipolar continuum.

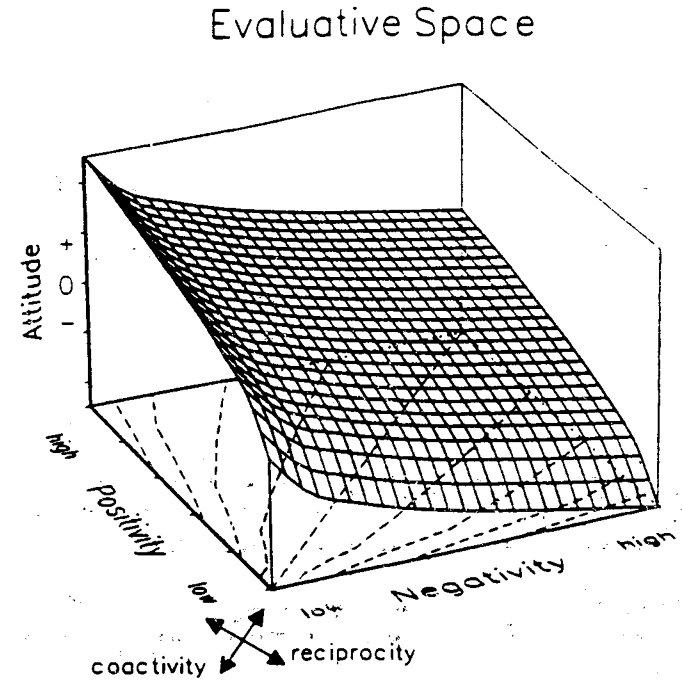

Figure 3. Bivariate evaluative plane and its associated attitudinal surface. This surface represents the attitudinal disposition of an individual toward (+) or away from (−) the target stimulus. The attitudinal disposition is expressed in relative units, as derived from Equation 1 with W_p = .5, W_n = .5, and Iij = 0. The axis dimensions are in relative units of activation. The point on the surface overlying the left axis intersection represents a maximally positive attitude, and the point on the surface overlying the right axis intersection represents a maximally negative attitude. In addition, each of the points overlying the dashed diagonal extending from the back to the front axis intersections depicts the same midpoint attitude. Thus, the nonreciprocal diagonal on the evaluative plane, which represents different evaluative processes (e.g., neutral to ambivalence), yields the same middling expression on the attitudinal surface. Dashed lines (including the coactivity diagonal) represent isoattitudinal contours on the evaluative plane. These isoattitudinal contours are illustrative rather than exhaustive.

A number of features and properties of this evaluative plane and attitudinal surface are worth noting here because they depend only on the positivity and negativity functions being monotonic. First, inspection of the attitudinal surface in Figure 3 confirms the empirical finding (e.g., J. D. Edwards & Ostrom, 1971) that changes in the evaluative processes underlying an attitude may not result in measurable changes using bipolar scales (i.e., values on the attitudinal surface). In general the relationship between loci on the evaluative plane and points on the attitudinal surface is characterized by a many-to-one mapping.[10] That is, each coordinate on an evaluative plane dictates a specific value on the attitudinal surface; however, each value on the attitudinal surface does not translate unequivocally into a specific coordinate on the evaluative plane. This is most obvious along the coactivity diagonal, which extends from the front to the back corners of the evaluative plane. As is apparent from inspecting Figure 3, all of the points along this diagonal map into the same value on the attitudinal surface. At the front corner of the evaluative plane, the activation level of both positivity and negativity equals zero, the resulting positivity and negativity equal zero, and the combination of these outputs is zero. The eliciting stimulus in this case evokes neutrality (i.e., no affectivity, as reflected by the locus on the evaluative plane), and the attitudinal response would be a middling position on a bipolar attitude scale (as reflected by the locus on the attitudinal surface). At the far corner, in contrast, the activation level of both positivity and negativity is maximal and their outputs are asymptotic, but the combination of these outputs still equals zero. The eliciting stimulus in this case evokes a great deal of affectivity and ambivalence (as reflected by the locus on the evaluative plane), but the attitudinal response on a bipolar scale is

[9] The simplifying assumptions implemented in Equation 1 when generating Figure 3 were made for didactic reasons. We consider the parameters in Equation 1 in greater detail shortly.

[10] See Cacioppo and Tassinary (1990) for a discussion of one-to-one, one-to-many, many-to-one, and many-to-many mappings between conceptual and empirical domains.

again a middling position (as reflected by the locus on the attitudinal surface). This coactivity diagonal can therefore be thought of as an *isoattitudinal contour.*

Many isoattitudinal contours exist on the evaluative plane (e.g., see the dashed lines in Figure 3). Measurements obtained from the attitudinal surface that fall above an isoattitudinal contour do not allow unambiguous inferences to be drawn about a person's attitudinal disposition with a bipolar index because any such measurement has multiple potential origins in the evaluative plane, each of which has quite different psychological implications. Thus, isoattitudinal contours represent *regions of indeterminism* in the mapping from the attitudinal surface (e.g., bipolar attitude scale responses) onto the underlying evaluative plane.

We do not mean to imply that measurements of the attitudinal surface have no bearing on underlying origins of evaluative processes, as represented in the bivariate evaluative plane. Note, for instance, that as the locus on the evaluative plane moves toward either the left or the right corner in Figure 3, the region of indeterminism becomes smaller until, at the extreme left and right corners, one-to-one mappings are achieved between the points on the evaluative plane and the attitudinal surface. The diagonal connecting these two corners, of course, reflects the reciprocal modes of evaluative activation, the implicit assumption underlying the ubiquitous bipolar scale (and, consequently, many theories) in the area of attitudes. Given that the only reasonably good mapping between surface measures and their underlying psychological origin is found in regions near the corners of the reciprocal diagonal, it perhaps should not be surprising that the past 60 years of research on attitudes can be interpreted as falling along this diagonal. Factors that produce nonreciprocal movements along the evaluative plane (e.g., the coactivation of positivity and negativity) produce attitudinal manifestations that are largely indistinguishable with bipolar scales. In the absence of distinguishable manifestations, these factors would appear to have little or no effect on attitudes and would be assigned little or no significance in attitude theories. Thus, bipolar attitude scales may have shaped and constrained attitude theories in more fundamental and substantial ways than generally recognized. Other regions of the evaluative plane may be at least as important to understand, but the interpretive limitations of traditional attitude conceptualizations and measures may have limited investigation of these regions.

Several aspects of Miller's conflict theory and data may bear on Equation 1. Recall that the manipulation of positive and negative environmental contingencies in the same situation consistently had opposing effects on approach–withdrawal behaviors. However, the fact that the behavioral effects of positive and negative environmental contingencies were opposite does not necessarily imply that positive and negative valent processes were reciprocally activated. Indeed, the gradients in Figure 2 indicate that both the strength of approach and the strength of avoidance (or withdrawal) increase as the distance from the goal decreases; in Equation 1, this is characterized as the coactivation of positivity and negativity.

Second, the avoidance and approach gradients depicted in Miller's research varied in slopes, but the avoidance gradient was typically considerably steeper. This steeper slope for negative input represents a *negativity bias* and suggests that bipolar attitudinal and behavioral expressions may be more strongly affected by negative than positive events. Interestingly, Skowronski and Carlston (1989) recently reviewed research on impression formation and found support for a negativity bias in attitude formation toward people, and Cacioppo and Gardner (1993) reviewed evidence for this bias in attitudes toward blood and organ donations. The negativity bias can be modeled in the present framework, for instance, by weighting negativity higher than positivity (i.e., $W_n > .5 > W_p$) in Equation 1.

Research on approach–withdrawal conflict also reveals what might be termed a *positivity offset* that is particularly apparent at low levels of evaluative activation. At distances far from the goal, the motivation to approach is higher than the motivation to avoid, even though the slope is lower in the former (see Figure 2). Thus, the positivity offset, evident in the higher intercept for the approach gradient, represents a higher starting point for the positive than negative activation functions. The negativity bias, evident in the steeper slope for the avoidance gradient, denotes a higher first derivative for the negativity than positivity activation functions.

The notion of a positivity offset is consistent with early research on word frequencies and meaning. For instance, the Thorndike-Lorge count (1944) revealed that the word *happiness* occurs more than 15 times as often as the word *unhappiness*, *good* occurs 5 times more often than *bad*, *love* occurs almost 7 times as often as *hate*, *sweet* is found almost 7 times more often than *sour*, *better* is used about 5 times more often than *worse*, and *best* occurs about 6 times as often as *worst*. In reviewing this and related work, Zajonc (1968) observed that

> the strength and pervasiveness of the relationship between word frequency and meaning—the *evaluative* aspect of meaning, in particular—is truly remarkable. For, if there is any correspondence between the frequency with which words are used and the actual preponderance of the things and events for which these words stand, then we may congratulate ourselves on living in a most happy world. (p. 2)

The positivity offset may, in fact, bear on Zajonc's (1968) mere exposure effect, that is, the tendency for unfamiliar neutral stimuli to become liked as a function of repeated (i.e., unreinforced) exposure per se.

It is interesting to speculate about the possible evolutionary significance of the positivity offset and negativity bias. The positivity offset would foster approach or exploratory behavior at low levels of evaluative activation, a tendency that has obvious survival value. Indeed, it is this natural curiosity that drives the scientific enterprise. In addition, a socially useful consequence of the positivity offset is that cohesion (i.e., positive attitudes) would develop among individuals who are simply exposed repeatedly to one another (Zajonc, 1968). The negativity bias, in contrast, fosters stronger responsiveness to proximate negative than proximate positive or neutral events. To the extent that it is more difficult to reverse the consequences of an injurious or fatal assault than an opportunity unpursued, a propensity to react more strongly to negative than positive stimuli may have developed through the process of natural selection. This is not to suggest that negative events foster learning over positive events across contexts, but only that psychological, physiological, and behavioral responses are more strongly influenced by proximate negative than positive stimuli within a given context.

By virtue of the negativity bias, individuals should be increasingly inclined toward withdrawal or avoidance at high levels of coactivation.[11] Consistent with this reasoning, Taylor (1991) has recently summarized a wide range of evidence showing that negative events in a context evoke stronger and more rapid physiological, cognitive, emotional, and social responses than neutral or positive events.

Third, Miller's theory and research on conflict indicate that animals approach the goal when distant from it, withdraw from the goal when near it, and hesitate when at the distance from the goal at which the approach and avoidance gradients intersect. These predictions are in complete accord with the computational model outlined earlier, in which the difference between the two gradients at any given distance from the goal reflects the strength and valence of the bipolar disposition arising as a result of the bivalent components.[12] Finally, Miller's theory concerns behavioral conflict originating from complex learning histories that included both positive and negative events. As such, it can be subsumed within the present framework as falling along the coactivity diagonal with the following revisions to Equation 1:

$$\text{Attitude} = .4(Pi + c) - .6Nj + Iij, \qquad (2)$$

where P is the activation function for positivity and is approximated by the function $i^{.5}$; i represents the level of positivity activated by an attitude object; c is a constant that contributes to a higher intercept for positivity than negativity (i.e., positivity offset); .4 is a conservative estimate of the weighting factor for positivity, approximated on the basis of the research on approach and avoidance gradients, that reflects the lower weighting for positivity relative to negativity (i.e., negativity bias); N is the activation function for negativity and is approximated by the function $j^{.5}$; j represents the level of negativity activated by an attitude object; .6 is the corresponding weighting factor for negativity; and Iij represents nonadditive effects.[13]

As before, the different modes of evaluative activation manifest as changes in the activation indices, i and j. This is a feature that allows attitude and conflict theory to be examined within a common framework. For example, Equation 2 yields the evaluative space displayed in Figure 4. The differential weighting of positivity and negativity produces functions with different slopes, depicting the relative impact and dynamic ranges of these evaluative processes on bipolar attitudes. The inset superimposes the separate marginals for the activation of positivity and negativity, providing the same graphical representation of these opposing forces used by Miller (1959) and depicted in Figure 2. Note that the bipolar attitudinal expressions overlying the reciprocal diagonal, particularly at the endpoints of the surface vector, continue to map into movements on the evaluative plane with the least indeterminism. In contrast, the bipolar attitudinal expressions overlying the coactivity diagonal continue to map in a one-to-many fashion onto the evaluative plane. Unlike Figure 2, however, the attitudinal manifestations over the coactivity diagonal in Figure 4 vary somewhat across this diagonal. Specifically, the attitudinal expression at each point on the diagonal equals the difference between the corresponding point on the positivity and negativity functions (i.e., marginals) depicted in Figure 2. Thus, the coactivity diagonal that moves through the rear axis intersection captures much of Miller's theory of conflict and is an isoattitudinal contour that corresponds to Miller's crossover point.

Miller (1951, 1966) also provided evidence that increasing or decreasing the strength of the "approach motivation" (e.g., by subjecting the animals to food deprivation before the test) affected the offset but not the slope of the approach gradient. Similarly, the intercept of the avoidance gradient but not its slope was affected when the strength of the avoidance motivation was manipulated (e.g., by increasing the electric shock). Given the stability of the slopes and the sensitivity of the offsets to the underlying motivational state of the organism before testing, we have proposed numeric coefficients for positivity and negativity, but we have not enumerated the constant. For illustrative purposes only, we specified $c = .1$ when generating Figure 4. We have not added a second constant to negativity because the offset can be thought of as the relative difference in intercept between the approach and avoidance gradients, and the offsets for the approach gradients generally were higher than those for the avoidance gradients.[14]

---

[11] The bias favoring approach at distances far from the goal may be an artifact of the methodology used; both gradients can be raised or lowered by increasing or decreasing the strength of the reinforcers. However, studies of the effects of stimulus similarity on the approach and avoidance gradients also revealed a positivity bias that cannot be explained away as easily (e.g., see Miller, 1951, 1959).

[12] Miller's conflict theory was developed in experiments on rats in which the notion of an attitudinal disposition underlying behavior was not considered. Miller also applied his theory to mental phenomena and problems in humans, however, as he wrote about its implications for personality (e.g., Miller, 1951) and psychotherapy (e.g., Miller, 1966). In a recent review, we found evidence that the difference in the bivalent organization of behavior across the phylogenetic scale is not that rudimentary mechanisms are absent in their effects on human behavior but that these mechanisms are modulated or overshadowed by the operation of more complex (e.g., relationally sensitive) mechanisms (Berntson, Boysen, & Cacioppo, in press). Thus, the separability and apparent coactivation of positive and negative motives in the animal research on conflict may generalize directly to human motives and attitudes (Epstein, 1980; Thompson, Zanna, & Griffin, in press).

[13] The most common empirical measures of positivity and negativity have been based on self-reports such as unipolar attitude scales (e.g., Bargh, Chaiken, Govender, & Pratto, 1992; Kaplan, 1972) and Watson, Clark, and Tellegen's (1988) PANAS. Tests of Miller's (1948, 1959) conflict theory relied on behavioral measures such as the strength of the pull toward or away from the goal box (Brown, 1948). Psychophysiological research has shown that electromyographic measures of expressive facial responses to attitudinal stimuli (e.g., Cacioppo, Petty, Losch, & Kim, 1986; Dimberg, 1988) and a late positive potential of the event-related brain response to attitudinal stimuli (Cacioppo, Crites, Berntson, & Coles, 1993; Cacioppo, Crites, Gardner, & Berntson, 1993) may be used to differentiate attitudinal valence and extremity. Each class of measures has potential limitations. Until the requisite research has been conducted on the reliability, validity, relative sensitivity, and practicality of verbal, behavioral, or physiological measures of positivity, negativity, and ambivalence, measurement by multiple operationalizations is suggested (see also D. P. Green, Goldman, & Salovey, 1993).

[14] An additional constant, β, can be incorporated into Equation 2 to quantify general individual differences in affective predisposition. This constant would simply change the scaling on the z-axis, however, and would not change the form of the gradients or overlying attitudinal surface.

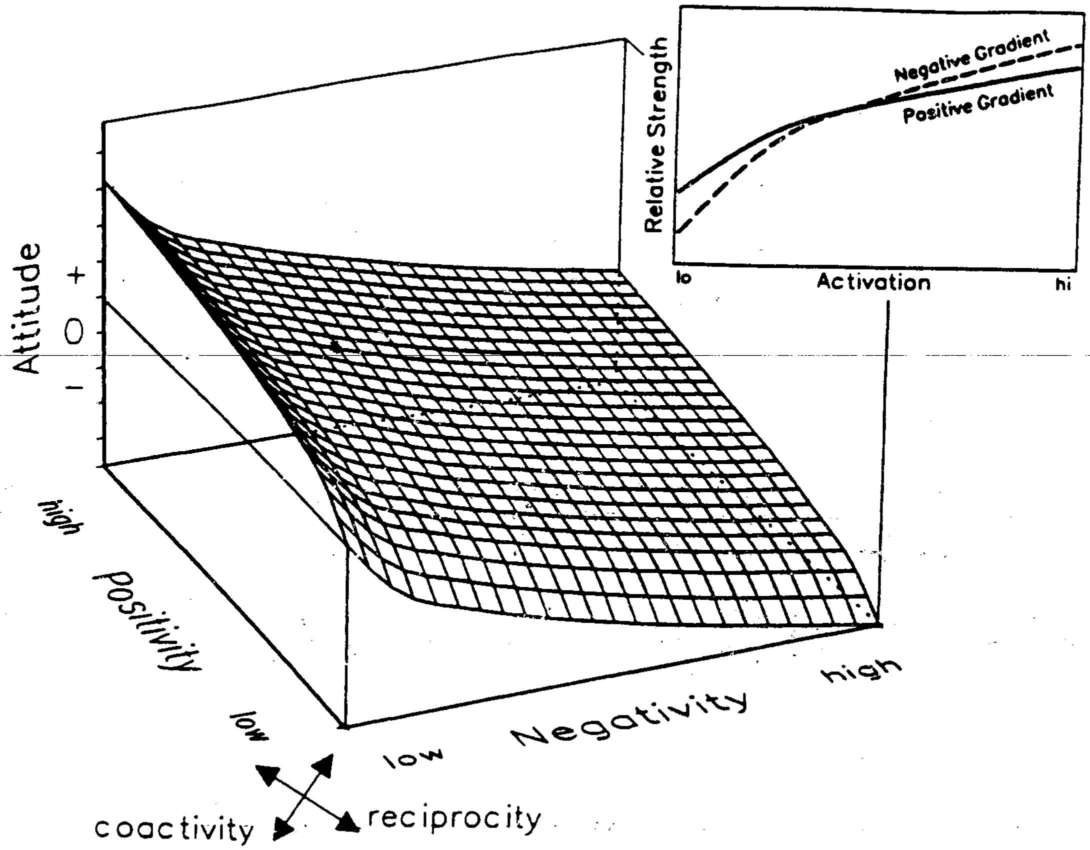

Figure 4. The bivariate evaluative plane and its associated attitudinal surface. The surface represents the attitudinal disposition of an individual toward (+) or away from (−) the target stimulus. The attitude is expressed in relative units, as derived from Equation 2. The point on the surface overlying the left axis intersection represents the maximally positive attitude evoked by the target stimulus, and the point on the surface overlying the right axis intersection represents the maximally negative attitude toward the target stimulus. The inset superimposes the strength of the activation of positive and negative valent forces as a function of movements along the coactivity diagonal. Note that the predictions depicted in this inset mirror those for approach–avoidance conflict in Miller's (1959) theory of conflict (see Figure 2).

In summary, we have reviewed evidence that the reciprocal central control over opposing neural mechanisms and exteroceptive reflexes is by no means a universal feature of their activation. We have also reviewed research suggesting that the positivity function begins at a higher point (e.g., fostering exploratory behavior) but has a slope that is less steep than the negativity function (e.g., fostering escape from hostile stimuli). These findings were used to further specify the parameters in the computational model (see Equation 2) and the corresponding attitudinal surface (see Figure 4). Although the positivity offset and negativity bias were derived from the research on conflict, Zajonc (1968) reviewed evidence from word frequency studies consistent with the existence of a positivity offset, Skowronski and Carlston (1989) reviewed evidence from the domain of impression formation (i.e., attitude formation toward people) for the existence of a negativity bias, and Taylor (1991) summarized evidence for the generally larger social, psychological, and physiological impact of negative than neutral and positive events. Thus, although we drew on research on behavioral conflict to estimate the parameters in Equation 2 that depict a positivity offset and a negativity bias, both effects have been observed in the area of attitudes as well.

Finally, interest in attitude ambivalence appears to be on the rise as the field has turned to questions about attitude structure (e.g., Hass et al., 1991; Parsi & Katz, 1986), automatic attitude activation (e.g., Bargh, Chaiken, Govender, & Pratto, 1992), and attitude strength (e.g., Thompson et al., in press). The bivariate framework and computational model specified in Equation 2 complement and extend existing approaches to attitude ambivalence (e.g., Abelson, Kinder, Peters, & Fiske, 1982; Breckler, in press; Crano, Gorenflo, & Shackelford, 1988; Crano & Sivacek, 1984; Kaplan, 1972; Scott, 1966, 1968; Thompson et al., in press). Two features common to most ambivalence formulations are that (a) ambivalence is predicted to increase as the distance from the origin increases along the coactivity diagonal and (b) ambivalence is predicted to be symmetrical about the coactivity diagonal (e.g., see Thompson et al., in press). However, the higher weighting of negativity than positivity in Equation 2 suggests that, *ceteris paribus*, ambivalence should be greater when negativity is the weaker of the two evaluative substrates. Unlike prior ambivalence formulations, therefore, a small displacement of the contour of peak ambivalence is predicted, representing an exaggerated effect of negativity on attitude ambivalence. A second difference between prior ambivalence formulations and theoretical expectations from the present model originates in the constant, c, which reflects the positivity offset. When positivity is zero, the constant implies that ambivalence generally decreases as negativity is increasingly activated. The absence of a comparable constant in the negativity term results in a different prediction for the effects of increasing positivity on ambivalence. Thus, both the negativity bias and the positivity offset suggest that attitude ambivalence may be asymmetrically rather than symmetrically distributed about the coactivity diagonal in Figure 3.[15]

Although the preliminary parameter estimates for deriving attitudinal manifestations (Equation 2) may require revision, the principles of evaluative activation, opposing evaluative actions, and bivalent modes of evaluative activation, along with Equation 1, may offer a useful framework for the study of an array of attitudinal phenomena. In the final section, we outline three additional attitudinal properties illuminated by the bivariate framework and explain how these properties depend only on the positivity and the negativity functions being monotonic. Given monotonic activation functions for positivity and negativity, we also outline several novel predictions about attitude change that follow from the bivariate computational model.

Attitudinal Properties as a Function of the Bivalent Mode of Evaluative Activation

The three modes of evaluative activation (i.e., reciprocal, uncoupled, and nonreciprocal) and the consequent attitude differ in terms of at least three properties: directional stability, dynamic range, and response lability (see Table 1: Berntson, Cacioppo, & Quigley, 1991). In reciprocal evaluative activation, changes in attitude are fueled consistently by the changes in the positive and negative evaluative processes evoked by the stimulus. For instance, an individual's attitude toward a stimulus would become more positive if the stimulus evoked positive evaluative processes and diminished negative evaluative processes. Variations in the magnitude of their reciprocal activation, although affecting the magnitude of the change in attitude, would not be expected to alter the basic direction of the change. Thus, whether positive and negative evaluative processes were reciprocally activated by a small or a large amount would not alter the direction of the movement across the evaluative plane and the consequent movement on the overlying attitudinal surface. Of course, positive and negative evaluative processes could be reciprocally activated with the negativity evoked by a stimulus increasing and the positivity evoked by the stimulus decreasing rather than vice versa. This form of reciprocal activation would also produce a directionally stable change in attitude toward the stimulus. To distinguish between these two forms of reciprocal activation, we offer the convention of naming both the mode (i.e., reciprocal) and the evaluative dimension that is increasing (i.e., positive or negative activation). Thus, reciprocal positive activation manifests in directionally stable and increasingly positive attitudes, and reciprocal negative activation results in directionally stable and increasingly negative attitudes.

The uncoupled modes of evaluative activation are characterized by changes in positive or negative evaluative processes but not both (e.g., Goldstein & Strube, in press). Thus, uncoupled modes of evaluative activation exhibit stable, although somewhat smaller, effects on attitudes than reciprocal modes.

In contrast, nonreciprocal modes (coactivation or coinhibition) have an inherently variable or unstable effect on the direction of the attitude change. In coactivation, for instance, both positive and negative evaluative processing are increasing; however, changes in the activation of positivity and negativity have functionally opposite effects on the consequent attitude. Minor variations in the magnitude of the relative activation of positivity or negativity, therefore, can lead to a more positive attitude, a more negative attitude, or no manifest change in attitude toward the stimulus, depending on which evaluative substrate predominates.

The properties of dynamic range and response lability are also inherent consequences of the mode of evaluative activation. Because of the opposing functional consequences of the activation of positivity and negativity on attitudes, (a) reciprocal activation of these underlying evaluative processes produces synergistic effects, thereby maximizing the dynamic range and the reactive lability (i.e., the response per unit of movement across the evaluative plane) in evaluative disposition; (b) nonreciprocal activation produces opposing effects on attitudes, which constrains the dynamic range and reactive lability; and (c) uncoupled modes of activation represent cases in which only one of the two valent classes of determinants is active, which yields attitudes with intermediate dynamic range and reactive lability. Thus, reciprocal activation maximizes directional stability and the range of change in attitudes, whereas nonreciprocal activation (e.g., coactivation and coinhibition) potentiates directional changes in attitudes but tends to restrict their range. Interestingly, the directional instability of ambivalent attitudes has been noted (e.g., Katz et al., 1986), but the mechanism depicted here underlying this directional stability is a formal property of the coactivation of opposing valent activation functions.

It is of interest to note that, especially when the activation functions for positivity and negativity are nonlinear, the rate of manifest change in attitudes (e.g., reactive lability) can diverge rather dramatically from the rate of change in underlying psychological (i.e., evaluative) processes. The gradients of a surface are the mathematical equivalent of the first derivative of a vector. The gradient of a surface, therefore, is directionally dependent (consider, for instance, a skier's acceleration when traveling across or up vs. down the surface of a mountain). Thus, there are different gradients, or rates of manifest change, for move-

[15] The conceptual differences among the ambivalence formulations, in addition to the anomalies in the ambivalence predictions made by some of these formulations, seem to call for research in which the comparative ability of these formulas to predict reported attitude ambivalence is investigated empirically. However, all of the formulations predict relatively equivalent outcomes for some combinations of positivity and negativity. For instance, all of these formulations predict that ambivalence increases as the distance from the origin increases along the coactivity diagonal. The most discriminating tests are the ambivalence ratings observed when either positivity or negativity equals zero. Further discussion of these issues is beyond the scope of this article, however.

Table 1
Modes of Evaluative Activation and Their Attitudinal Properties

Negativity	Positivity: Decrease	Positivity: No change	Positivity: Increase
Decrease	Coinhibition	Uncoupled negative reduction	Reciprocal positive activation
Dynamic range	Low	Intermediate	High
Reactive lability	Low	Intermediate	High
Directional stability	Low	High	High
No change	Uncoupled positive reduction	Baseline	Uncoupled positive activation
Dynamic range	Intermediate		Intermediate
Reactive lability	Intermediate		Intermediate
Directional stability	High		High
Increase	Reciprocal negative activation	Uncoupled negative activation	Coactivation
Dynamic range	High	Intermediate	Low
Reactive lability	High	Intermediate	Low
Directional stability	High	High	Low

Note. The possible modes of evaluative activation are reciprocal (reciprocal positive activation and reciprocal negative activation), uncoupled (uncoupled negative reduction, uncoupled negative activation, uncoupled positive reduction, and uncoupled positive activation), and nonreciprocal (coinhibition and coactivation). The naming convention in reciprocal modes is to specify the evaluative substrate that is increasing rather than the one that is decreasing. Thus, reciprocal positive activation signifies that positivity is increasing and negativity is decreasing. The properties of dynamic range, response lability, and directional stability are inherent consequences of the mode of evaluative activation. Because of the opposing functional consequences of the activation of positivity and negativity on attitudes, (a) reciprocal activation of these underlying evaluative processes produces cumulative effects, thereby maximizing the dynamic range, reactive lability, and directional stability of the consequent attitudes; (b) nonreciprocal activation produces competing effects on attitudes, which constrains the dynamic range and reactive lability and potentiates directional instability; and (c) uncoupled modes of activation represent cases in which only one of the two valent classes of determinants is active or inhibited, which produces attitudes with intermediate dynamic range and reactive lability and high directional stability. Thus, reciprocal activation maximizes directional stability and the range of change in attitudes, whereas nonreciprocal activation (e.g., coactivation) potentiates directional changes in attitudes but tends to restrict their range.

ments of a given distance across the evaluative plane depending on the direction of the movement (i.e., the mode of evaluative activation). Figure 5 illustrates the gradients for uncoupled positive activation (upper left panel), uncoupled negative reduction (upper right panel), coinhibition (lower left panel), and reciprocal positive activation (lower right panel).[16] The inset above each panel depicts the attitudinal surface, and the arrow in each inset indicates the direction of the movement (i.e., the mode of evaluative activation) across the attitudinal surface that yields the surface gradient in that panel. Movements in the opposite direction, as depicted by the arrow in each inset, yield surface gradients that are of the same magnitude as but opposite in sign to the gradient depicted in the corresponding panel.

The attitudinal property of directional stability is visually evident in the four gradient surfaces depicted in Figure 5. Note, for instance, that only the surface gradient in the lower left panel (coinhibition) depicts both positive and negative attitude changes as originating in movements in one direction across the evaluative plane. This directional instability is a hallmark feature of the nonreciprocal mode of evaluative activation.

Inspection of the gradient surface associated with the reciprocal mode of evaluative activation (lower right panel) reveals that it is elevated more above zero than any of the other gradient surfaces, consistent with this mode producing the greatest change in attitudes. The gradient surface is not flat, however, and its topography reveals that, *ceteris paribus*, the rate of attitude change should be greater for a neutral attitude (see the front axis intersection) than for an ambivalent attitude (see the rear axis intersection). Because the initial manifestations of these two attitudes would be similar on a bipolar attitude scale (see the front and rear axis intersections in Figure 5), these predictable differences in attitude change as a function of the initial location on the evaluative plane and the mode of evaluative activation would typically be relegated to the error term.

Investigations of the general locus of an initial attitude on the evaluative plane and the mode of evaluative activation evoked by a stimulus (e.g., persuasive communication or mere exposure) may contribute considerably to an understanding of attitude change. Consider an individual whose precommunica-

[16] The mappings inherent in these quantitative formulations are relative, and the depictions described here and elsewhere in this article adhere to the operational rather than to the representational approach to measurement (Michell, 1986). For instance, the predictions represented by the relative height and topography of these surface gradients require only that (a) the activation functions for positivity and negativity are negatively accelerating (i.e., marginal utility functions), (b) there is a positivity offset, and (c) there is a negativity bias.

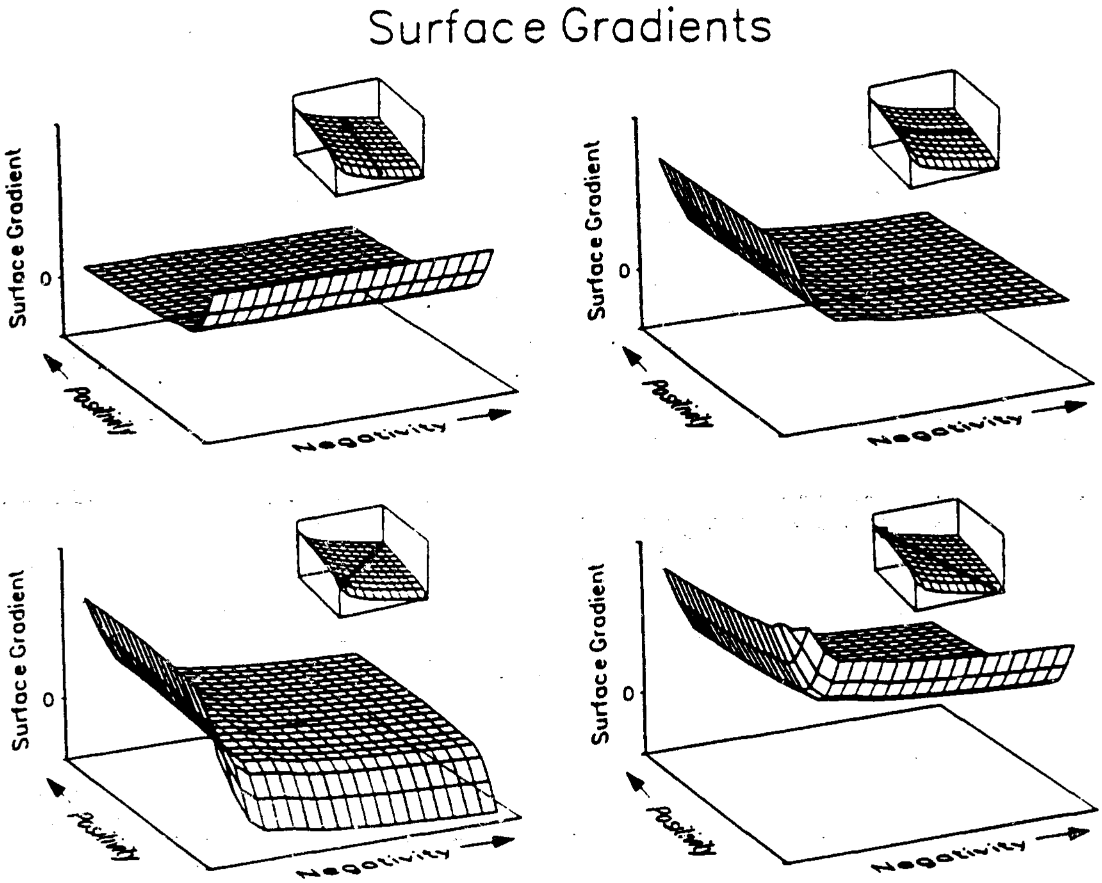

Figure 5. Surface gradients as a function of four directions of movement (modes of activation) within evaluative space. The bivalent modes of evaluative activation are uncoupled positive activation (upper left panel), uncoupled negative reduction (upper right panel), coinhibition (lower left panel), and reciprocal positive activation (lower right panel). The small insets show the functional surface, and the arrow in each inset indicates the direction of the movement in evaluative space. The larger plot in each panel depicts the surface gradient resulting from movements across the attitudinal surface in the depicted direction for all loci within the evaluative space. Movements in the direction opposite to that depicted by the arrow in an inset yield surface gradients identical in topography but opposite in sign.

tion attitude is middling and whose location on the evaluative plane is the center. Inspection of the upper left panel shows that movement of a given distance from the center across the evaluative plane in the uncoupled positive activation mode produces only modest attitude change because the movements on the attitudinal surface are approaching the asymptote for the positivity activation function. Specifically, the middle of the arrow in the inset in the upper left panel overlays the center of the evaluative plane. Uncoupled positive activation from the center of the evaluative plane means that there is a movement from the centerpoint of the arrow toward the end of the arrow. Inspection of the positivity marginal in this inset confirms that this movement on the evaluative plane corresponds to relatively (and increasingly) small positive slopes on the attitudinal surface. Thus, the surface gradient in the upper left panel is positive but approaches zero as the activation of positivity approaches its asymptote.

Now consider the surface gradient depicted in the upper right panel of Figure 5. As earlier, assume that an individual's initial attitude is middling and the location on the evaluative plane is the center. The overlying surface gradient indicates that, if a stimulus (e.g., persuasive communication) were to activate the uncoupled negative reduction mode, bipolar attitude measures would show that the communication produced more positive attitudes. Moreover, the attitude change that would result from a movement of a given distance from the center by virtue of this mode would be considerably larger than would result from a movement of the same distance from the center by uncoupled positive activation. This difference is especially apparent as the movement approaches the edge of the evaluative plane.

The lower left panel shows that movement of the same distance from the center across the evaluative plane in the coinhibition mode would produce little or no manifest attitude change on the bipolar measure, whereas the lower right panel shows that movement of the same distance from the center in the reciprocal positive activation mode would yield considerable attitude change (i.e., reflecting the synergistic effects on attitude change of both increased positivity and decreased negativity).

Examination of the topographies of the surface gradients in Figure 5 indicates other features that may be of some interest.

For instance, attitude changes resulting from either uncoupled mode of evaluative activation are predicted to be relatively unaffected by the tonic (i.e., initial) level of the other evaluative substrate. This is evident in the generally level surface gradients illustrated in the upper left and upper right panels of Figure 5. The amount of attitude change predicted to result from the reciprocal or nonreciprocal mode of evaluative activation, in contrast, is affected rather dramatically in some regions by the initial level of both evaluative substrates.

Although research is needed to test the formal properties outlined in this section, these properties illustrate the generative utility of thinking qualitatively and quantitatively about the positive and negative evaluative processes underlying attitudes as separable rather than as reciprocal. Consideration of factors such as the general location of an initial attitude on the evaluative plane and the mode of evaluative activation evoked by a stimulus, for instance, may allow explanation of variance that is presently ignored by attitude change theories and is relegated to the error term in attitude change research.

Summary

Thurstone's (1928, 1931; Thurstone & Chave, 1929) scaling of attitudes along a single continuum originated in his conceptualization of an attitude as the net difference between feelings of appetition and feelings of aversion toward a stimulus. When separable component operations for positive and negative valent activation functions were posited after demonstrations of approach–avoidance conflict, these developments were not assimilated into the attitude literature (cf. Kelman, 1962). Instead, the activation of positive and negative evaluative (e.g., motivational and cognitive response) processes was reduced to a simple difference that could be represented by a single bipolar vector or scale. In the present article, an alternative bivariate formulation of evaluative processes has been outlined, and evidence reviewed, indicating that attitude theory and measurement may benefit from expanding the principle of reciprocal evaluative activation to accommodate (a) the separable activation of positive and negative evaluative processes, (b) investigation of their unique antecedents and consequences, and (c) examination of the psychological and physiological constraints that produce an apparent bipolar structure of data and that foster bivariate manifestations. In view of the ubiquity of positive and negative valent processes in behavior, quantification of their separable and interactive dynamics may have broad significance for the study of attitudes, preferences, ambivalence, and conflict.

References

Abelson, R. P., Aronson, E., McGuire, W. J., Newcomb, T. M., Rosenberg, M. J., & Tannenbaum, P. H. (1968). *Theories of cognitive consistency: A sourcebook*. Chicago: Rand McNally.

Abelson, R. P., Kinder, D. R., Peters, M. D., & Fiske, S. T. (1982). Affective and semantic components in political person perception. *Journal of Personality and Social Psychology, 42*, 619–630.

Ahern, G. L., & Schwartz, G. E. (1985). Differential lateralization for positive and negative emotion in the human brain: EEG spectral analysis. *Neuropsychologia, 23*, 745–756.

Allport, G. W. (1935). Attitudes. In C. Murchison (Ed.), *Handbook of social psychology* (pp. 798–844). Worcester, MA: Clark University Press.

Amsel, A. (1990). Arousal, suppression, and persistence: Frustration theory, attention, and its disorders. In J. A. Gray (Ed.), *Psychobiological aspects of relationships between emotion and cognition* (pp. 239–268). Hillsdale, NJ: Erlbaum.

Anderson, N. H. (1965). Primacy effects in personality impression formation using generalized order effects paradigm. *Journal of Personality and Social Psychology, 2*, 1–9.

Bargh, J. A., Chaiken, S., Govender, R., & Pratto, F. (1992). The generality of the automatic attitude activation effect. *Journal of Personality and Social Psychology, 62*, 893–912.

Bentler, P. M. (1969). Semantic space is (approximately) bipolar. *Journal of Psychology, 71*, 33–40.

Berntson, G. G., Boysen, S. T., & Cacioppo, J. T. (in press). Neurobehavioral organization and the cardinal principle of evaluative bivalence. In F. M. Crinella & J. Yu (Eds.), *Brain mechanisms. 1990: Papers in memory of Robert Thompson*. New York: New York Academy of Sciences.

Berntson, G. G., Cacioppo, J. T., & Quigley, K. S. (1991). Autonomic determinism: The modes of autonomic control, the doctrine of autonomic space, and the laws of autonomic constraint. *Psychological Review, 98*, 459–487.

Bradburn, N. M. (1969). *The structure of psychological well-being*. Chicago: Aldine.

Bradburn, N. M., & Caplovitz, D. (1965). *Reports on happiness*. Chicago: Aldine.

Bradley, M. M., Cuthbert, B. N., & Lang, P. J. (1990). Startle reflex modification: Attention or emotion? *Psychophysiology, 27*, 513–527.

Breckler, S. J. (1984). Empirical validation of affect, behavior, and cognition as distinct attitude components. *Journal of Personality and Social Psychology, 47*, 1191–1205.

Breckler, S. J. (in press). A comparison of numerical indexes for measuring attitude ambivalence. *Educational and Psychological Measurement*.

Brehm, J. (1956). Post-decision changes in desirability of alternatives. *Journal of Abnormal and Social Psychology, 52*, 384–389.

Briggs, N. C., Piliavin, J. A., Lorentzen, D., & Becker, G. A. (1986). On willingness to be a bone marrow donor. *Transfusion, 26*, 324–330.

Brown, J. S. (1948). Gradients of approach and avoidance responses and their relation to level of motivation. *Journal of Comparative and Physiological Psychology, 41*, 450–465.

Cacioppo, J. T. (1991, May). *Attitudes*. Invited address presented at the annual meeting of the Midwestern Psychological Association, Chicago.

Cacioppo, J. T., Crites, S. L., Berntson, G. G., & Coles, M. G. H. (1993). If attitudes affect how stimuli are processed, should not they affect the event-related brain potential? *Psychological Science, 4*, 108–112.

Cacioppo, J. T., Crites, S. L., Gardner, W. L., & Berntson, G. G. (1993). *Bioelectrical echoes from evaluative categorization: I. A late positive brain potential that varies as a function of trait negativity and extremity*. Manuscript submitted for publication.

Cacioppo, J. T., & Gardner, W. L. (1993). What is underlying medical donor attitudes and behavior? *Health Psychology, 12*, 269–271.

Cacioppo, J. T., Petty, R. E., & Berntson, G. G. (1991). Persuasion. In R. Dulbeco (Ed.), *Encyclopedia of human biology* (pp. 799–810). San Diego, CA: Academic Press.

Cacioppo, J. T., Petty, R. E., & Geen, T. R. (1989). Attitude structure and function: From the tripartite to the homeostasis model of attitudes. In A. R. Pratkanis, S. J. Breckler, & A. G. Greenwald (Eds.), *Attitude structure and function* (pp. 275–310). Hillsdale, NJ: Erlbaum.

Cacioppo, J. T., Petty, R. E., Kao, C. F., & Rodriguez, R. (1986). Central and peripheral routes to persuasion: An individual difference perspective. *Journal of Personality and Social Psychology, 51,* 1032–1043.

Cacioppo, J. T., Petty, R. E., Losch, M. E., & Kim, H. S. (1986). Electromyographic activity over facial muscle regions can differentiate the valence and intensity of affective reactions. *Journal of Personality and Social Psychology, 50,* 260–268.

Cacioppo, J. T., & Sandman, C. A. (1981). Psychophysiological functioning, cognitive responding, and attitudes. In R. E. Petty, T. M. Ostrom, & T. C. Brock (Eds.), *Cognitive responses in persuasion* (pp. 81–104). Hillsdale, NJ: Erlbaum.

Cacioppo, J. T., & Tassinary, L. G. (1990). Inferring psychological significance from physiological signals. *American Psychologist, 45,* 16–28.

Callero, P. L., Howard, J. A., & Piliavin, J. A. (1987). Helping behavior as role behavior: Disclosing social structure and history in analysis of prosocial action. *Social Psychology Quarterly, 50,* 247–256.

Callero, P. L., & Piliavin, J. A. (1983). Developing a commitment to blood donation: The impact of one's first experience. *Journal of Applied Psychology, 13,* 1–16.

Campbell, D. T. (1950). The indirect assessment of social attitudes. *Psychological Bulletin, 47,* 15–38.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56,* 81–105.

Charng, H., Piliavin, J. A., & Callero, P. L. (1988). Role identity and reasoned action in the prediction of repeated behavior. *Social Psychology Quarterly, 51,* 303–317.

Condie, S. J., Warner, W. K., & Gillman, D. C. (1976). Getting blood from collective turnips: Volunteer donation in mass blood drives. *Journal of Applied Psychology, 61,* 290–294.

Cook, S. W., & Selltiz, C. A. (1964). A multiple-indicator approach to attitude measurement. *Psychological Bulletin, 62,* 36–55.

Crano, W. D., Gorenflo, D. W., & Shackelford, S. L. (1988). Overjustification, assumed consensus, and attitude change: Further investigation of the incentive-aroused ambivalence hypothesis. *Journal of Personality and Social Psychology, 55,* 12–22.

Crano, W. D., & Sivacek, J. (1984). The influence of incentive-aroused ambivalence on overjustification effects in attitude change. *Journal of Experimental Social Psychology, 20,* 137–158.

Davidson, R. J. (1992). Emotion and affective style: Hemispheric substrates. *Psychological Science, 3,* 39–43.

Davidson, R. J. (in press). Parsing affective space: Perspectives from neuropsychology and psychophysiology. *Neuropsychology.*

Davidson, R. J., Ekman, P., Saron, C., Senulis, J., & Friesen, W. V. (1990). Approach/withdrawal and cerebral asymmetry: Emotional expression and brain physiology, I. *Journal of Personality and Social Psychology, 58,* 330–341.

Davidson, R. J., Finman, R., Straus, A., & Kagan, J. (1992). *Patterns of frontal asymmetry differentiate between wary and outgoing children.* Manuscript submitted for publication.

Davidson, R. J., Kalin, N. H., & Shelton, S. E. (in press). Lateralized effects of diazepam on frontal brain electrical asymmetries in rhesus monkeys. *Biological Psychology.*

Dawes, R. M., & Smith, T. L. (1985). Attitude and opinion measurement. In G. Lindzey & E. Aronson (Eds.), *The handbook of social psychology* (pp. 509–566). New York: Random House.

de Mille, R. (1970). Logical and empirical oppositeness in value responses. *Psychological Reports, 26,* 143–154.

Diener, E., & Emmons, R. A. (1985). The independence of positive and negative affect. *Journal of Personality and Social Psychology, 47,* 1105–1117.

Diener, E., & Iran-Nejad, A. (1986). The relationship in experience between various types of affect. *Journal of Personality and Social Psychology, 50,* 1031–1038.

Dimberg, U. (1988). Facial electromyography and the experience of emotion. *Journal of Psychophysiology, 2,* 277–282.

Doob, L. W. (1947). The behavior of attitudes. *Psychological Review, 54,* 135–156.

Eddington, A. (1939). *The philosophy of physical science.* New York: Macmillan.

Edwards, A. I. (1946). A critique of 'neutral' items in attitude scales constructed by the method of equal appearing intervals. *Psychological Review, 53,* 159–169.

Edwards, J. D., & Ostrom, T. M. (1971). Cognitive structure and neutral attitudes. *Journal of Experimental Social Psychology, 7,* 36–47.

Edwards, P. W., & Zeichner, A. (1985). Blood donor development: Effects of personality, motivational, and situational variables. *Personality and Individual Differences, 6,* 743–751.

Epstein, S. (1980). Conflict and stress. In L. Goldberger & S. Breznitz (Eds.), *Handbook of stress* (pp. 49–68). New York: Free Press.

Fazio, R. H. (1990). Multiple processes by which attitudes guide behavior: The MODE model as an integrative framework. *Advances in Experimental Social Psychology, 23,* 75–109.

Fazio, R. H., Powell, M. C., & Williams, C. J. (1989). The role of attitude accessibility in the attitude-to-behavior processes. *Journal of Consumer Research, 16,* 280–288.

Fischoff, B. (1991). Value elicitation: Is there anything in there? *American Psychologist, 64,* 835–847.

Fischoff, B., Slovic, P., & Lichtenstein, S. (1980). Knowing what you want: Measuring labile values. In T. Wallsten (Ed.), *Cognitive processes in choice and decision behavior* (pp. 117–141). Hillsdale, NJ: Erlbaum.

Fishbein, M. (1980). A theory of reasoned action: Some applications and implications. In H. Howe & M. Page (Eds.), *Nebraska symposium on motivation, 1979* (pp. 65–116). Lincoln: University of Nebraska Press.

Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intention, and behavior.* Reading, MA: Addison-Wesley.

Fishbein, M., & Raven, B. H. (1962). The AB scales. *Human Relations, 15,* 35–43.

Fox, N. A. (1991). If it's not left, it's right. *American Psychologist, 46,* 863–872.

Goldstein, M. D., & Strube, M. J. (in press). Independence revisited: The relation between positive and negative affect in a naturalistic setting. *Personality and Social Psychology Bulletin.*

Gotlib, I. H., & Meyer, J. P. (1986). Factor analysis of the multiple affect adjective check list: A separation of positive and negative affect. *Journal of Personality and Social Psychology, 50,* 1161–1165.

Gray, J. A. (1987). *The psychology of fear and stress* (2nd ed.). Cambridge, England: Cambridge University Press.

Gray, J. A. (1990). Brain systems that mediate both emotion and cognition. In J. A. Gray (Ed.), *Psychobiological aspects of relationships between emotion and cognition* (pp. 239–288). Hillsdale, NJ: Erlbaum.

Gray, J. A. (1991). Neural systems, emotion and personality. In J. Madden (Ed.), *Neurobiology of learning, emotion and affect* (pp. 273–306). New York: Raven Press.

Green, D. P. (1988). On the dimensionality of public sentiment toward partisan and ideological groups. *American Journal of Political Science, 32,* 758–780.

Green, D. P., & Citrin, J. (in press). Measurement error and the structure of attitudes: Are positive and negative judgments opposites? *American Journal of Political Science.*

Green, D. P., Goldman, S. L., & Salovey, P. (1993). Measurement error masks bipolarity in affect ratings. *Journal of Personality and Social Psychology, 64,* 1029–1041.

Green, R. F., & Goldfried, M. R. (1965). On the bipolarity of semantic space. *Psychological Monographs, 79*(6, Whole No. 599).

Greenwald, M. K., Cook, E. W. III, & Lang, P. J. (1989). Affective judgment and psychophysiological response: Dimensional covariation in the evaluation of pictorial stimuli. *Journal of Psychophysiology, 3,* 51–64.

Gutek, B. A. (1978). On the accuracy of retrospective attitudinal data. *Public Opinion Quarterly, 42,* 390–401.

Hamm, A., Stark, R., & Vaitl, D. (1990). Startle reflex potentiation and electrodermal response differentiation: Two indicators of two different processes in Pavlovian conditioning (Abstract). *Psychophysiology, 27,* S37.

Hass, R. G., Katz, I., Rizzo, N., Bailey, J., & Eisenstadt, D. (1991). Cross-racial appraisal as related to attitude ambivalence and cognitive complexity. *Personality and Social Psychology Bulletin, 17,* 83–92.

Hass, R. G., Katz, I., Rizzo, N., Bailey, J., & Moore, L. (1992). When racial ambivalence evokes negative affect, using a disguised measure of mood. *Personality and Social Psychology Bulletin, 18,* 786–797.

Henriques, J. B., & Davidson, R. J. (1990). Regional brain electrical asymmetries discriminate between previously depressed subjects and healthy controls. *Journal of Abnormal Psychology, 99,* 22–31.

Henriques, J. B., & Davidson, R. J. (1991). Left frontal hypoactivation in depression. *Journal of Abnormal Psychology, 100,* 535–545.

Hintzman, D. L. (1990). Human learning and memory: Connections and dissociations. *Annual Review of Psychology, 41,* 109–139.

Hoebel, B. G. (1985). Brain neurotransmitters in food and drug reward. *American Journal of Clinical Nutrition, 42,* 1133–1150.

Hoebel, B. G. (1988). Neuroscience and motivation: Pathways and peptides that define motivational systems. In R. C. Atkinson, R. J. Herrnstein, G. Lindzey, & R. D. Luce (Eds.), *Stevens handbook of experimental psychology* (2nd ed., pp. 547–625). New York: Wiley.

Howe, M. L. (in press). On measuring (in)dependence of cognitive processes. *Psychological Review.*

Jones, E. E., & Sigall, H. (1971). The bogus pipeline: A new paradigm for measuring affect and attitude. *Psychological Bulletin, 76,* 349–364.

Kahneman, D., & Tversky, A. (1979). Prospect theory. *Econometrica, 47,* 263–292.

Kaplan, K. J. (1972). On the ambivalence-indifference problem in attitude theory and measurement: A suggested modification of the semantic differential technique. *Psychological Bulletin, 77,* 361–372.

Katz, I., & Hass, R. G. (1988). Racial ambivalence and American value conflict: Correlational and priming studies of dual cognitive structures. *Journal of Personality and Social Psychology, 55,* 893–905.

Katz, I., Wackenhut, J., & Hass, R. G. (1986). Racial ambivalence, value duality and behavior. In J. F. Dovidio & S. L. Gaetner (Eds.), *Prejudice, discrimination, and racism* (pp. 35–59). San Diego, CA: Academic Press.

Kelman, H. C. (1962). The induction of action and attitude change. *Proceedings of the XIVth International Congress of Applied Psychology: Personality Research, 2,* 81–110.

Klopfer, F. J., & Madden, T. M. (1980). The middle-most choice on attitude items: Ambivalence, neutrality, or uncertainty? *Personality and Social Psychology Bulletin, 6,* 97–101.

Konorski, J. (1948). *Conditioned reflexes and neuron organization.* Cambridge, England: Cambridge University Press.

Konorski, J. (1967). *Integrative activity of the brain: An interdisciplinary approach.* Chicago: University of Chicago Press.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1990). Emotion, attention, and the startle reflex. *Psychological Review, 97,* 377–395.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1992). A motivational analysis of emotion: Reflec-cortical connections. *Psychological Science, 3,* 44–49.

Lang, P. J., Greenwald, M. K., Bradley, M. M., & Hamm, A. O. (1993). Looking at pictures: Affective, facial, visceral, and behavioral manifestations. *Psychophysiology, 30,* 261–273.

Lang, P. J., Ohman, A., & Vaitl, D. (1988). *The international affective picture system* [Photographic slides]. Gainesville: Center for Research in Psychophysiology, University of Florida.

Larsen, R. J., & Diener, E. (1992). Promises and problems with the circumplex model of emotion. *Review of Personality and Social Psychology, 13,* 25–59.

Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology, 140,* 5–53.

Lipset, S. M., & Schneider, W. (1978). The Bakke case: How would it be decided at the bar of public opinion. *Public Opinion, 1,* 38–44.

Loftus, E. F., Fienberg, S. E., & Tanur, J. M. (1985). Cognitive psychology meets the National Survey. *American Psychologist, 40,* 175–180.

Mackintosh, N. J. (1983). *Conditioning and associative learning.* New York: Oxford University Press.

Madden, T. M., & Klopfer, F. J. (1978). The 'cannot decide' option in Thurstone-type attitude scales. *Educational and Psychological Measurement, 38,* 259–264.

Marshall, G. N., Wortman, C. B., Kusulas, J. W., Hervig, L. K., & Vickers, R. R., Jr. (1992). Distinguishing optimism from pessimism: Relations to fundamental dimensions of mood and personality. *Journal of Personality and Social Psychology, 62,* 1067–1074.

Martin, I., & Levey, A. B. (1978). Evaluative conditioning. *Advances in Behavior Research and Therapy, 1,* 57–102.

Masterson, F. A., & Crawford, M. (1982). The defense motivation system: A theory of avoidance behavior. *Behavioral and Brain Sciences, 5,* 661–696.

McCombie, R. P. (1991). Blood donation patterns of undergraduate students: Family and friendship correlates. *Journal of Community Psychology, 19,* 161–165.

McGuire, W. J. (1985). Attitudes and attitude change. In G. Lindzey & E. Aronson (Eds.), *The handbook of social psychology* (3rd ed., Vol. 2, pp. 233–346). New York: Random House.

McManis, M. H., & Lang, P. J. (1992, October). *PANAS picture rating.* Paper presented at the annual meeting of the Society for Psychophysiological Research, San Diego, CA.

Meddis, R. (1972). Bipolar factors in mood adjective checklists. *British Journal of Social and Clinical Psychology, 11,* 178–184.

Michell, J. (1986). Measurement scales and statistics: A clash of paradigms. *Psychological Bulletin, 100,* 398–407.

Miller, N. E. (1948). Theory and experiment relating psychoanalytic displacement to stimulus-response generalization. *Journal of Abnormal and Social Psychology, 43,* 155–178.

Miller, N. E. (1951). Comments on theoretical models illustrated by the development of a theory of conflict behavior. *Journal of Personality, 20,* 82–100.

Miller, N. E. (1959). Liberalization of basic S-R concepts: Extensions to conflict behavior, motivation and social learning. In S. Koch (Ed.), *Psychology: A study of a science, Study 1* (pp. 198–292). New York: McGraw-Hill.

Miller, N. E. (1961). Some recent studies on conflict behavior and drugs. *American Psychologist, 16,* 12–24.

Miller, N. E. (1966). Some animal experiments pertinent to the problem of combining psychotherapy with drug therapy. *Comprehensive Psychiatry, 7,* 1–12.

Obone, D. J., & Bradley, S. (1975). Blood donor and non-donor motivation: A transnational replication. *Journal of Applied Psychology, 60,* 409–410.

Olds, J. (1958). Self-stimulation in the brain. *Science, 127,* 315–324.

Olds, J., & Milner, P. (1954). Positive reinforcement produced by electrical stimulation of septal area and other regions. *Journal of Comparative and Physiological Psychology, 47,* 419–427.

Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. (1957). *The measurement of meaning*. Urbana: University of Illinois Press.

Ostrom, T. M. (1973). The bogus pipeline: A new *ignis fatuus? Psychological Bulletin, 79*, 252–259.

Oswalt, R. M., & Napoliello, M. (1974). Motivations of blood donors and non-donors. *Journal of Applied Psychology, 59*, 122–124.

Parsi, N., & Katz, I. (1986). Attitudes toward posthumous organ donation and commitment to donate. *Health Psychology, 5*, 565–580.

Patchen, M., Davidson, J. D., Hofman, G., & Brown, W. R. (1977). Determinants of students' interracial behavior and opinion change. *Sociology of Education, 50*, 55–75.

Patchen, M., Hofman, G., & Davidson, J. D. (1976). Interracial perceptions among high school students. *Sociometry, 39*, 341–354.

Paulhus, D. L., Shaffer, D. R., & Downing, L. L. (1977). Effects of making blood donor motives salient upon donor retention: A field study. *Personality and Social Psychology Bulletin, 3*, 99–102.

Petty, R. E., & Cacioppo, J. T. (1981). *Attitudes and persuasion*. Dubuque, IA: William C. Brown.

Pomazal, R. J., & Jaccard, J. J. (1976). An informational approach to altruistic behavior. *Journal of Personality and Social Psychology, 33*, 317–326.

Poulton, E. C. (1989). *Bias in quantifying judgments*. Hillsdale, NJ: Erlbaum.

Reiman, E. M., Raichle, M. E., Butler, F. K., Herscovitch, P., & Robbins, E. (1984). A focal brain abnormality in panic disorder, a severe form of anxiety. *Nature, 310*, 683–684.

Rosenthal, R., & Rosnow, R. L. (1969). *Artifact in behavioral research*. San Diego, CA: Academic Press.

Russell, J. A. (1979). Affective space is bipolar. *Journal of Personality and Social Psychology, 37*, 345–356.

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology, 39*, 1161–1178.

Russell, J. A., Lewicka, M., & Niit, T. (1989). A cross-cultural study of a circumplex model of affect. *Journal of Personality and Social Psychology, 57*, 848–856.

Sarason, I. G., Sarason, B. R., Pierce, G. R., Shearin, E. N., & Sayers, M. H. (1991). A social learning approach to increasing blood donations. *Journal of Applied Social Psychology, 21*, 896–918.

Schaffer, C. E., Davidson, R. J., & Saron, C. (1983). Frontal and parietal electroencephalogram asymmetry in depressed and nondepressed subjects. *Biological Psychiatry, 18*, 753–762.

Schneirla, T. C. (1959). An evolutionary and developmental theory of biphasic processes underlying approach and withdrawal. In *Nebraska symposium on motivation: 1959* (pp. 1–42). Lincoln: University of Nebraska Press.

Schofield, J. W. (1989). *Black and White in school: Trust, tension, or tolerance?* New York: Teachers College Press.

Schofield, J. W. (1991). School desegregation and intergroup relations. In G. Grant (Ed.), *Review of research in education* (Vol. 17, pp. 335–409). Washington, DC: American Educational Research Association.

Schuman, H., & Kalton, G. (1985). Survey methods. In G. Lindzey & E. Aronson (Eds.), *Handbook of social psychology* (pp. 635–698). New York: Random House.

Scott, W. A. (1966). Brief report: Measures of cognitive structure. *Multivariate Behavior Research, 1*, 391–395.

Scott, W. A. (1968). Attitude measurement. In G. Lindzey & E. Aronson (Eds.), *Handbook of social psychology* (Vol. 2, pp. 204–273). Reading, MA: Addison-Wesley.

Skowronski, J. J., & Carlston, D. E. (1989). Negativity and extremity biases in impression formation: A review of explanations. *Psychological Bulletin, 105*, 131–142.

Snydersmith, M., & Cacioppo, J. T. (1992). Parsing complex social factors to determine component effects: I. Autonomic activity and reactivity as a function of human association. *Journal of Social and Clinical Psychology, 11*, 263–278.

Spector, P. E. (1992). *Summated rating scale construction: An introduction*. Newbury Park, CA: Sage.

Summers, G. A. (1970). *Attitude measurement*. Chicago: Rand McNally.

Taylor, S. E. (1991). Asymmetrical effects of positive and negative events: The mobilization-minimization hypothesis. *Psychological Bulletin, 110*, 67–85.

Tellegen, A. (1985). Structures of mood and personality and their relevance to assessing anxiety, with an emphasis on self-report. In A. H. Tuma & J. D. Maser (Eds.), *Anxiety and the anxiety disorder* (pp. 681–706). Hillsdale, NJ: Erlbaum.

Thompson, M. M., Zanna, M. P., & Griffin, D. W. (in press). Let's not be indifferent about (attitudinal) ambivalence. In R. E. Petty & J. A. Krosnick (Eds.), *Attitude strength: Antecedents and consequences*. Hillsdale, NJ: Erlbaum.

Thorndike, E. L., & Lorge, I. (1944). *The teacher's word-book of 30,000 words*. New York: Teachers College, Columbia University.

Thurstone, L. L. (1928). Attitudes can be measured. *American Journal of Sociology, 33*, 529–554.

Thurstone, L. L. (1931). The measurement of attitudes. *Journal of Abnormal Psychology, 26*, 249–269.

Thurstone, L. L., & Chave, E. L. (1929). *The measurement of attitude*. Chicago: University of Chicago Press.

Tomarken, A. J., Davidson, R. J., Wheeler, R. E., & Doss, R. C. (1992). Individual differences in anterior brain asymmetry and fundamental dimensions of emotion. *Journal of Personality and Social Psychology, 62*, 676–687.

Tourangeau, R., & Rasinski, K. A. (1988). Cognitive processes underlying context effects in attitude measurement. *Psychological Bulletin, 103*, 299–314.

Tulving, E. (1985). How many memory systems are there? *American Psychologist, 40*, 385–398.

Turner, C. F. (1984). Why do surveys disagree? Some preliminary hypotheses and some disagreeable examples. In C. F. Turner & E. Martin (Eds.), *Surveying subjective phenomena* (pp. 159–214). New York: Russell Sage Foundation.

Turner, C. F., & Krauss, E. (1978). Fallible indicators of the subjective state of the nation. *American Psychologist, 33*, 456–470.

Turner, C. F., & Martin, E. (1984). *Surveying subjective phenomena*. New York: Russell Sage Foundation.

Vrana, S. R., Spence, E. L., & Lang, P. J. (1988). The startle probe response: A new measure of emotion? *Journal of Abnormal Psychology, 97*, 487–491.

Watson, D. (1988). The vicissitudes of mood management: Effects of varying descriptors, time frames, and response formats on measures of positive and negative affect. *Journal of Personality and Social Psychology, 55*, 128–141.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology, 54*, 1063–1070.

Watson, D., & Tellegen, A. (1985). Toward a consensual structure of mood. *Psychological Bulletin, 98*, 219–235.

Wilson, T. D., & Schooler, J. W. (1991). Thinking too much: Introspection can reduce the quality of preferences and decisions. *Journal of Personality and Social Psychology, 60*, 181–192.

Zajonc, R. B. (1968). Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology Monograph Supplement, 9*(2, Pt. 2).

Zanna, M. P., Detweiler, R. A., & Olson, J. M. (1984). Physiological mediation of attitude maintenance, formation, and change. In W. M. Waid (Ed.), *Sociophysiology* (pp. 163–196). New York: Springer-Verlag.

Zanna, M. P., & Rempel, J. K. (1988). Attitudes: A new look at an old concept. In D. Bar-Tal & A. W. Kruglanski (Eds.), *The social psychology of knowledge* (pp. 315–334). Cambridge, England: Cambridge University Press.

Zanna, M. P., & Thompson, M. M. (1991, September). *Let's not be indifferent about (attitudinal) ambivalence.* Paper presented at the Ohio State University Symposium on Attitudes and Persuasion, Nags Head, NC.

Zevon, M. A., & Tellegen, A. (1982). The structure of mood change: An idiographic/nomothetic analysis. *Journal of Personality and Social Psychology, 43,* 111–122.

Received December 11, 1992
Revision received August 24, 1993
Accepted August 26, 1993 ■