Psychological Bulletin
1999, Vol. 125, No. 1, 3–30

0033-2909/99/$3.00

On the Bipolarity of Positive and Negative Affect

James A. Russell and James M. Carroll
University of British Columbia

Is positive affect (PA) the bipolar opposite of, or is it independent of, negative affect (NA)? Previous analyses of this vexing question have generally labored under the false assumption that bipolarity predicts an invariant latent correlation between PA and NA. The predicted correlation varies with time frame, response format, and items selected to define PA and NA. The observed correlation also varies with errors inherent in measurement. When the actual predictions of a bipolar model are considered and error is taken into account, there is little evidence for independence of what were traditionally thought opposites. Bipolarity provides a parsimonious fit to existing data.

> Man! Thou pendulum betwixt a smile and tear.
> —Lord Byron, *Childe Harold's Pilgrimage*

Science has repeatedly shown that things are not necessarily the way they appear. The sun does not rise in the east, a solid rock consists mainly of empty space, the continents move about, and panda bears are not bears. Science seems about to destroy yet another long-held belief.

Zautra, Potter, and Reich (1997) observed that "most of us believe that positive feelings are the opposite of negative feelings, and that a person who is unhappy is also sad. These statements are truisms in the language of feelings, affects, and emotions, as fundamental as one plus one equals two" (p. 75). Psychologists have uncovered evidence that positive affective feelings are not, in fact, the bipolar opposite of negative affective feelings: It seems that a human being is not a pendulum, moving between opposite feelings. A pendulum can be in only one place at a time, but a human being can be both happy and unhappy. Zautra, Potter, and Reich went on to invite their readers "to consider the possibility that in many cases one and one does not equal two, at least when it comes to comparing positive and negative affective states. Instead, . . . most of the time, positive and negative feeling states are independent of one another" (p. 75).

The National Advisory Mental Health Council (1995) similarly advised "while one would ordinarily think that positive and negative emotions are opposites, apparently this is not the case. . . . This finding [is] one of the most interesting results of emotion research" (p. 839). In an article on the measurement "and mismeasurement" (p. 267) of mood, Watson and Clark (1997) declared as a "fundamental psychometric principle" (p. 282) that "oppositely valenced affects tend to be only weakly negatively correlated with one another" (p. 282). They therefore posited that "variations in positive and negative mood are largely independent of one another" (p. 270). Costa and McCrae (1980) called independence of positive and negative affect a "paradox that has never been fully explained" (p. 669). They remarked that "the repeated observation that the pleasantness and unpleasantness of one's life are uncorrelated is a puzzling phenomenon the explanation for which is of considerable theoretical importance" (p. 670).

Evidence has challenged the bipolar view so often that it now seems on its deathbed, and independence has taken its place as the prevailing assumption. At the same time, some researchers have insisted that the evidence for independence might be an artifact of method, and others have found reason to maintain the assumption of bipolar opposites (Eich, 1995; Parducci, 1995; Reisenzein, 1994; Smith & Ellsworth, 1985). Bipolarity is assumed in Helson's (1964) adaptation level theory and Osgood's (1969) semantic differential technique.

The resulting controversy is central to the psychology of affect. Its resolution touches on such basic issues as the processes involved in affect and its causes and consequences and what strategies to use against crippling negative emotions (Is negative affect counteracted by, or independent of, improvements in positive affect?). Some theories assume independent unipolar dimensions of positive and negative affect (Mayer & Gaschke, 1988; Meyer & Shack, 1989; Morris, 1989; Tellegen, 1985; Watson & Clark, 1997)—as do some measures of affect (Watson, Clark, & Tellegen, 1988). Other theories assume a bipolar positive–negative dimension (Feldman, 1995; Lang, Greenwald, Bradley, & Hamm, 1993; Ortony, Clore, & Collins, 1988; Reisenzein, 1994; Russell, 1989)—as do some measures of affect (Cuthbert, Bradley, & Lang, 1996; Russell, Weiss, & Mendelsohn, 1989). Different lines of research are accumulating on each of these fundamentally different assumptions.

In this article, we focus on the bipolarity of one quality of affect, variously called its hedonic quality, its valence, its pleasant–unpleasant quality, or its positive–negative quality—positive affect (PA) and negative affect (NA). We later consider how to define these concepts. By *affect,* we have in mind genuine subjective feelings and moods (as when someone says, "I'm feeling sad"), rather than thoughts about specific objects or events (as

---

James A. Russell and James M. Carroll, Department of Psychology, University of British Columbia, Vancouver, British Columbia, Canada.

This study was funded by the Social Sciences and Humanities Research Council of Canada through a predoctoral fellowship to James M. Carroll and through a grant. We thank James H. Steiger, Shawn Reynolds, Sherri Widen, Lisa Feldman Barrett, Kim Barchard, and Darrin Lehman for their advice in formulating the analysis presented here and for comments on a draft of this article.

Correspondence concerning this article should be addressed to James A. Russell, Department of Psychology, 2136 West Mall, University of British Columbia, Vancouver, British Columbia, Canada V6T 1Z4. Electronic mail may be sent to jrussell@neuron3.psych.ubc.ca.

when someone calmly says, "The crusades were a sad chapter in human history"). Affect can be assessed at any given moment, rather than only in times of great emotion. We limit our analysis to the psychometrics of affect as experienced, where almost all of the challenge to bipolarity has taken place, and we do not consider the underlying biological mechanisms.[1]

We argue that the controversy's persistence stems from various misconceptions held by writers (including ourselves) on both sides of the debate. So, before examining the accumulated evidence, we examine the logic behind the research. We argue that the bipolar view of affect has been misunderstood, that incorrect predictions have been derived from it, and that existing data have not been properly interpreted. Indeed, even though bipolarity has been repeatedly defended or rejected, no model of bipolarity with precise psychometric predictions has been specified. When we formulated a bipolar model, some of its predictions surprised us, and not all commonly used methods of testing bipolarity turned out to be appropriate. Reviewing the evidence, we found little support for the independence of what were once thought opposites. A bipolar view of affect provides a parsimonious fit to existing data. For affect, one plus one does equal (approximately) two.

Background: The Persistent Debate

Traditionally, pleasure and displeasure were assumed to be opposites. With the introduction of psychometrically sophisticated correlational techniques, pioneers such as Nowlis (Nowlis & Nowlis, 1956) and Bradburn (1969; Bradburn & Caplovitz, 1965) sought to establish the structure of affect empirically. They anticipated bipolar dimensions, but their results forced them to question this assumption. They and other early researchers (Borgatta, 1961; Clyde, 1963; McNair & Lorr, 1964; Thayer, 1967) found that factor analysis yielded two independent unipolar factors where one bipolar factor had been expected. The correlation between positive and negative affect was surprisingly low.

Defenders of bipolarity soon appeared (Bentler, 1969; Meddis, 1972; Russell, 1979), and the debate was on. The debate initially centered on whether bipolarity was being masked by errors of measurement. Random noise is well known to attenuate a correlation coefficient. The more random error there is, the closer an observed correlation will be to zero. The more unreliable two scales are, the more independent they appear. Thus, now that Bradburn's (1969) scales of PA and NA have been shown to have relatively low reliability (Watson, 1988), his research should be cited for its pioneering contribution but no longer as credible evidence of the independence of PA and NA.

In a classic article, Bentler (1969) showed that a nonrandom error, specifically, an acquiescent response style, can also mask bipolarity by shifting the correlation in a positive direction. *Acquiescence* refers to individual differences in the tendency to agree or disagree with an item regardless of its content. Consistent with Bentler's (1969) analysis, an acquiescent response style has been repeatedly found to influence affect ratings (Davison & Srichantra, 1988; Lorr, 1989; Lorr & McNair, 1982; Lorr, McNair, & Fisher, 1982; Lorr & Shea, 1979; Lorr, Shi, & Youniss, 1989; Lorr & Wunderlich, 1980; Russell, 1979; Russell & Mehrabian, 1977; Tellegen, Watson, & Clark, 1994, in press). Many other forms of systematic (nonrandom) error exist (Hunter & Schmidt, 1990), although their role in the appearance of independence or bipolarity of affect has yet to be systematically examined.

However, it would be false to portray the study of affect as focused on measurement error. Most researchers implicitly assumed that the correlation between *observed* scores provides a reasonable estimate of the true correlation between the *latent* variables. Indeed, in the 1980s, interest in the role of measurement error faded. In their review of the history of this debate, Green, Goldman, and Salovey (1993) pointed to three highly influential articles that appeared around the same time advocating independence in one form or another (Diener & Emmons, 1984; Warr, Barter, & Brownbridge, 1983; Zevon & Tellegen, 1982). After this trio, Green et al. (1993) observed a "virtual cottage industry developed with the goal of demonstrating that positive and negative affect were indeed independent across a variety of contexts" (p. 1031).

In 1993, a time when measurement errors were all but forgotten in the study of affect, Green et al. (1993) published a powerful reminder of their potential impact. They also proposed a new approach. Instead of examining each type of error separately, Green et al. (1993) used a variety of different response formats to measure the same concept and analyzed the resulting data with a structural equation model. To the extent that the formats vary in the kinds of systematic error to which they are susceptible, it should be possible to estimate the zero-order correlation between latent variables relatively free of both systematic and random error. Green et al. (1993) found that observed correlations were systematically biased away from bipolarity and toward independence. For example, the observed correlations between happy and sad scales ranged from $-.25$ to $-.69$; the true correlation between their latent scores was estimated to be $-.84$. Diener, Smith, and Fujita (1995) observed that "the work of Green et al. clearly demonstrated the absolute necessity of controlling measurement error when examining the structure of emotions" (p. 131).

Nevertheless, consideration of measurement error has not sufficed to resolve the controversy. Watson and Clark (1997) reviewed relevant evidence and conceded that random error, acquiescence, and other systematic errors do exist and do alter the observed correlation, thereby biasing results away from bipolarity and toward independence. Nonetheless, they concluded that PA and NA are still largely independent. When they examined each source of error separately, the amount of variance accounted for by each one was not large. Watson and Clark also argued that their own scales from the Positive Affect and Negative Affect Schedule (PANAS; Watson et al., 1988) are only weakly correlated. Their estimate of this correlation after controlling random and one systematic error (acquiescence) was $-.43$, a figure that fell in a region Tellegen et al. (1994) had defined as "largely independent." Even

[1] Perhaps some day behavioral and neurochemical measures of affect will be developed, but they will presumably have to be validated at least to some extent against self-reported experience, and so we are back to self-report. There is no way to bypass solving the problems of bipolarity and independence within the realm of self-reported experience. Cacioppo & Bernston (1994) discuss bipolarity and independence in this broader context. They point out that bipolarity should be empirically tested rather than presupposed and that testing requires that alleged opposites be treated within a bivariate framework. They also argue that positive and negative affect likely result from separate neurological processes.

with their full procedure, Green et al. (1993) estimated a similar correlation between the constructs underlying these scales. Later, we return to these scales.

The debate between Green et al. (1993) and Watson and Clark (1997; see also Tellegen et al., 1994) raised a number of unresolved questions. Foremost, once measurement error is controlled, what correlation indicates bipolarity? What are we to make of empirically obtained figures such as $-.84$ and $-.43$? And why is there a discrepancy? As our review will show, this broad range exemplifies the findings in this field. Although it is an essential consideration, measurement error alone cannot tell the whole story.

A second consideration that is essential in resolving this controversy is time. Because one can feel happy today and unhappy tomorrow (in fact, we suspect that one's mood can change over the course of completing a long mood questionnaire), any analysis of bipolarity and independence requires a very careful consideration of time. Some research is cross-sectional, concerning the mood of different individuals at one point in time, whereas other research is longitudinal, concerning the various moods of the same individual over some extended period of time; these two designs do not yield exactly the same estimate of the correlation between PA and NA (Watson & Clark, 1997).

Time was also central to two landmark articles that proposed conceptual resolutions to the puzzle. Warr et al. (1983) pointed out that Bradburn's (1969) findings are less paradoxical when time is considered. Bradburn had asked each respondent to rate retrospectively the affect of the previous few weeks. Over that length of time, the number of pleasant episodes might well be unrelated to the number of unpleasant episodes, even though the proportion of time spent in a pleasant state is inverse to the proportion of time spent in an unpleasant state. Diener and Emmons (1984) then proposed, and offered evidence to show, that measures of momentary affect suggest a more bipolar relation between PA and NA than do measures of affect extended over time.

On the other hand, Watson (1988) offered evidence of his own that PA and NA are independent across a range of time frames. Green et al. (1993) offered evidence of their own that PA and NA are bipolar both at one time and when extended over time. And so, time must be considered, but once again it is not sufficient to resolve the controversy.

A third essential consideration is the multidimensional nature of affect. Affect involves various dimensions or components (Russell, 1978; Smith & Ellsworth, 1985; Scherer, 1984), and therefore the issue of bipolarity must be distinguished from the issue of independence. The question of the bipolarity of any one dimension is not equivalent to the question of how many independent dimensions (or components) are required to describe affect (just as answering the question of whether extraversion is the bipolar opposite of introversion would not specify the total number of independent dimensions required to describe the domain of personality, or vice versa). Still, the questions are closely related and easily confused. Two variables that are bipolar opposites are the whole of or parts of one dimension; two variables that are independent or even separable are two dimensions. The multidimensional nature of affect thus opens the door to substantive confounds. If bipolarity is taken as predicting one dimension where independence predicts two, then evidence of two or more substantive dimensions in the domain of affect could conceivably be mistaken for evidence against bipolarity. If measures of positive and negative affect were both confounded with the same component, then their observed correlation would be shifted in a positive direction.

Once again, however, acknowledgement of the multidimensional nature of affect is not sufficient to resolve the controversy. Noting at least two dimensions of affect, Green et al. (1993) argued that Watson et al.'s (1988) scales of PA and NA both contained a component of arousal, and it is this shared substantive component that accounts for their correlation being shifted in a positive direction. Tellegen et al. (1994) countered that Green et al.'s bipolar model was unidimensional and was thus contradicted by evidence of the multidimensional nature of affect.

To summarize, as soon as affect was measured with modern correlational techniques in the 1950s, evidence began to emerge that challenged the traditional assumption of bipolarity. In turn, the possibility was raised that bipolarity was being masked through errors in those same modern techniques. The ensuing debate has been especially vexing for its persistence. The problem seems to defy empirical solution. Both sides in the dispute have presented a similar range of correlations, including values high, medium, and low in magnitude. More than additional data, resolution would seem to depend on making sense of the range of already available results. Much progress has been achieved, with three considerations now understood to be essential in resolving the controversy: (a) the role of random and systematic errors of measurement, (b) time, and (c) the multidimensional nature of affect.

Our analysis, to which we now turn, builds on these three factors. The remainder of this article is divided into two major sections based on time. In the first section, the research examined concerns affective feelings at a given moment. We call this *momentary affect.* In the second, the research examined concerns affective feelings extended over a good length of time, such as happiness over several months. We call this *extended affect.* The psychometric considerations in momentary and extended affect are different. We build on the multidimensional nature of affect by looking at its implication for item selection. We also point to an additional factor (response format) so far not sufficiently appreciated. Our theme is that the original three factors plus response format are interactive and must therefore be considered simultaneously. We emphasize the need to articulate an explicit representation of bipolarity that specifies its predictions as these four factors vary. We then compare those predictions with the available data.

Momentary Affect

Affective feelings ebb and flow over the course of a day. Like the weather, they sometimes change slowly, sometimes quickly. Because affect is transitory, the first case to consider is affect at a slice in time. When the voice of tradition says that one cannot be both happy and sad, or both tense and relaxed, or both elated and depressed, it means both at the same time.

The studies that have challenged bipolarity of momentary affect have been of three basic types: (a) a factor analysis that yields two unipolar factors, positive as one and negative as another; (b) external correlates of PA and NA that are not the mirror image of each other (for instance, whereas PA correlates with Extraversion but not Neuroticism, NA correlates with Neuroticism but not Extraverson); and (c) a weak correlation between separate unipolar scales of PA and NA. All three types of study ultimately rely on the correlation coefficient, which is thus the statistic to focus on.

We examine the three types of evidence, but before doing so, we take a step curiously overlooked in previous tests of the bipolar view. We develop an explicit statement of that view or what we call a *bipolar model.* To interpret empirically obtained correlations, we need to know the correlation predicted by a bipolar model. To derive this predicted value, we must carefully consider how the abstract notions of PA and NA are made operational. We argue that the predicted correlation varies with both item selection and response format.

Building a Model of Bipolarity

Definitions, Semantics, and Item Selection

To test bipolarity, one must specify just what is supposed to be bipolar to what. Our hypothesis comes from everyday language. As Zautra et al. (1997) noted, it is in the everyday "*language* [italics added] of feelings, affects, and emotions" (p. 75) that bipolarity is a truism. The claim that PA is independent of NA is counterintuitive, because, in large part, the words *positive* and *negative* are antonyms.

In language, opposites are antonyms: *happy* and *sad, tense* and *relaxed, elated* and *depressed.* Not every positive affect term is the antonym of every negative affect term. It is *sad* rather than *tense* that is the opposite of *happy,* but it is *tense* rather than *sad* that is the opposite of *relaxed.* Thus, there is not one but many bipolarity hypotheses. When the topic is the bipolarity of a specific affective state (such as feeling guilty), then its specific semantic opposite (feeling innocent) is the proper target of a test of its bipolarity.

Bipolarity has not been challenged at the level of such specific items as happy and sad or guilty and innocent but rather at a more abstract level: factors or scales named positive and negative. For example, Watson and Clark (1997) acknowledged that happiness and sadness form a bipolar pair, even as they proclaimed and defended the independence of "positive and negative mood" (p. 270) in general.

On this more abstract level, the key question becomes this: What counts as PA and NA in general? Some terms are already more general (*positive, happy, negative, unhappy*), others are more specific (*proud, guilty*). A test of bipolarity of affect in general could rely either on more general terms or on a broad representative sample of more specific terms. In either case, the items chosen must still be semantic opposites.

Sampling raises the question of how to specify the domain from which the word samples are to be drawn. Presumably, PA in general is simply all affective states that are pleasant, and NA in general is all affective states that are unpleasant. As it happens, the key issue has been the definition of affect. Unfortunately, the word *affect* (like *emotion, mood,* and *feeling*) is not consensually defined. Psychologists have not agreed which specific states (and hence which words in an affect questionnaire) are to be included and which are to be excluded. Lack of agreement is a serious difficulty, because one definition can pick out a set of states that manifest bipolarity and another definition can pick out a set of states that do not. Not all researchers have explicitly defined their terms, although most seem to share a common-sense notion that affect includes the self-reportable feelings of happiness, sadness, elation, depression, tension, relaxation, and countless others, including but not limited to those involved in mood and emotion. Because of disagreements and lack of explicitness, a review of this literature cannot impose a definition on the domain. We therefore sought to present an analysis of definitions as independent as possible of any one conceptualization of the domain.

Semantic studies of the affect domain show that affect words are defined by a number of components—affect is multidimensional. Valence is but one component. In words from *thrilled* to *tranquil,* our English lexicon recognizes many different types of PA. Similarly, in words from *distressed* to *depressed,* it recognizes many different types of NA. To account for this diversity, we require components beyond valence. Here we emphasize one, namely what has been variously called *arousal, activity,* or *activation.* It is an empirical finding that both pleasant and unpleasant words vary in the level of activation they imply (Averill, 1975; Bush, 1973; Neufeld, 1975, 1976; Russell, 1978; Thayer, 1989; Whissell, 1981). Some pleasant words imply activation (*elated, thrilled*), and others imply deactivation (*serene, calm*). Some unpleasant words imply activation (*upset, distressed*), and others imply deactivation (*lethargic, depressed*).

Figure 1 shows an array of affect-related words, each of which varies both in valence (the x axis) and activation (the y axis). Not all bipolar pairs lie directly on the x axis. Activation has fanned them out. Positive terms lie on the right half of Figure 1 and vary in level of activation. Negative terms lie on the left half and also vary in level of activation. Notice that a rigid rod placed through the center of Figure 1 would mark out specific bipolar opposites. Thus, exact antonyms fall 180° away. (Quasi antonyms fall near 180°, just as quasi synonyms are separated by small angles.) The opposite of a pleasant activated word (such as *elated*) is an unpleasant deactivated word (*depressed*). Thus, bipolar opposites are opposite on both valence and activation. In testing bipolarity, activation must be taken into account. A two-dimensional representation of the semantics of affect is admittedly oversimplified, but Figure 1 provides a first approximation of the bipolarity implicit in the everyday language of affect. The important point is that Figure 1 is thoroughly bipolar. In fact, we cannot imagine a model more bipolar.

Now, the question becomes, where in Figure 1 is affect? Different researchers have picked out different regions of Figure 1 as affect. Rather than impose our own definition of this domain, we offer Figure

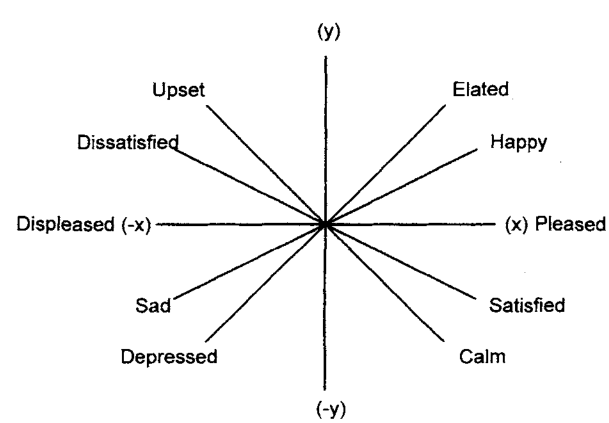

Figure 1. A semantic analysis of affect terms as composed of two components: x = pleasantness, y = activation.

2 as a broad perspective that encompasses various possible definitions. On the right side of Figure 2, we have defined three clusters of positive items. PA/HighAct refers to a cluster of positively valenced items that are also high in activation, with items such as enthused, excited, and energetic. PA/MediumAct refers to a cluster of positively valenced items that are medium (or noncommittal) in activation, with items such as happy, gratified, and content. PA/LowAct refers to a cluster of positively valenced items that are also low in activation, with items such as calm, serene, and relaxed. Three clusters are used for convenience; we believe that positive items vary more or less continuously in the amount of activation that they denote.

In parallel fashion, we have defined three clusters of negative items on the left side of Figure 2. NA/HighAct refers to a cluster of negatively valenced items that are also high in activation, such as jittery, tense, and nervous. NA/MediumAct refers to a cluster of items that are medium (or noncommittal) in activation, such as unhappy, miserable, and troubled. NA/LowAct refers to a cluster of negatively valenced items that are low in activation, such as depressed, lethargic, and down. Again, the number of clusters is merely a convenience.

Any or all of these clusters might be used to define affect. So, even at this abstract level, there is more than one bipolarity hypothesis. According to the bipolar model of Figures 1 and 2, the relationship between PA and NA depends on which item clusters are selected. It is essential to distinguish among types of PA and among types of NA. Only some pairs of clusters include semantic opposites. For instance, PA/HighAct is semantically opposite (180° away from) NA/LowAct but is not opposite either NA/MediumAct or NA/HighAct. Exclude the cluster 180° away and you have excluded bipolarity by definition.

Figures 1 and 2 are nothing more than a working semantic model meant to clarify one feature of the language of affect: bipolarity. In trying to be as clear as possible, we may seem to imply that emotion concepts can be defined as precisely as can types of triangles. A more thorough analysis of the semantics of affect would discuss such properties as fuzziness, hierarchical relations, and semantic components beyond valence and activation (Fehr & Russell, 1984; Bullock & Russell, 1984; Russell, 1978, 1991a, 1991b, 1991c, 1997; Russell & Bullock, 1986; Russell & Fehr, 1994). Although our treatment is brief, we rely on extensive evidence from unidimensional and multidimensional scaling and

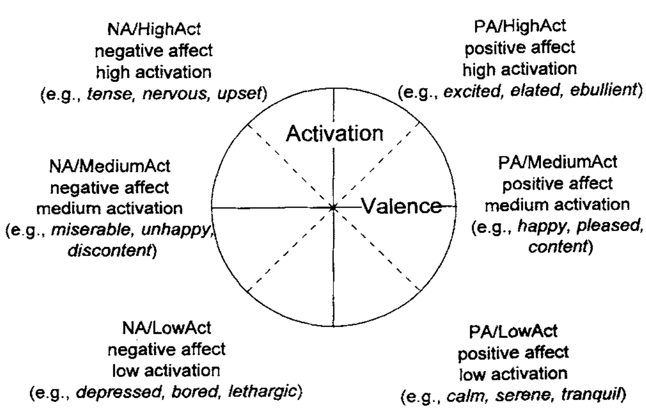

Figure 2. Six clusters of affect items defined by valence and activation. PA = Positive Affect; NA = Negative Affect.

other semantic studies of affect across diverse languages (Averill, 1975; Bush, 1973; Osgood, 1969; Russell, 1991c; Russell, Lewicka, & Niit, 1989; Whissell, 1981). Our aim in this section is not to use semantic studies as evidence for bipolarity, but as a means to develop a hypothesis. In later sections, we take up the entirely different question of whether affective experience actually conforms to the semantic hypothesis of Figures 1 and 2.

However simple, our model of bipolarity highlights certain requirements in tests of bipolarity. In testing the bipolarity of a specific affective state, it is necessary to select its specific semantic opposite. We know of no challenge to the bipolarity of specific affect items and thus focus on what has been challenged: the bipolarity of PA and NA in general. This challenge raises more difficult issues, but it is just as necessary to specify a bipolarity hypothesis and to select items that test that hypothesis. Items selected must actually be the hypothesized bipolar opposites and must represent PA and NA in general. In the multidimensional domain of affect, bipolarity is an angle of 180° between items in Figures 1 and 2. Bipolarity is thus understood as one relation among a range of relations. As we discuss shortly, the predicted correlation between affect items varies with the angle between them. But before developing that point, we must examine the response format used to assess each variable.

Response Format and Part Versus Whole Definitions

On our bipolar model, happy, sad, elated, depressed, or any other item can be conceptually and operationally defined in one of two ways. It can be defined either as its whole underlying bipolar continuum or as a region of that continuum. To test whether two items (or clusters) are bipolar opposites, one must measure the two items separately. To do so requires a unipolar response format. Although Meddis (1972) and Warr et al. (1983) emphasized the possible effects of response format, the consequences of this act of separation and this use of a unipolar format for potentially bipolar items have gone largely unrecognized in writings on bipolarity. As we explain in this section, separation and unipolar formats can force each item to be defined as a region of rather than the whole of the full continuum—thereby producing a relation (and hence correlation) between PA and NA that is not what researchers have generally presupposed. (An analysis in agreement with ours was made by Diener and Iran-Nejad, 1986.)

Defined as whole continuum. On the bipolar model, one way to define an item conceptually and operationally is as its full underlying continuum. In this case, the positive item is a dimension that extends all the way from the most extreme negative feeling through neutral to the most extreme positive feeling. To achieve this definition, the scoring procedure must assign the lowest score to the most negative feelings and its highest score to the most positive. A neutral feeling is assigned a score approximately midway between these extremes. Only some response formats can do the job. Here is one candidate for the item *happy*:

Circle the number that describes your present mood:

−7	−6	−5	−4	−3	−2	−1	0	1	2	3	4	5	6	7
(very sad)							(neutral)							(very happy)

Of course, *sad* and *happy* could be replaced with any pair of valenced antonyms.

A negative item can also be defined as the full underlying bipolar continuum, simply looked at and scored in the opposite direction. In this case, the negative item starts at (and the lowest score is assigned to) the most positive feeling, goes through neutral (which is assigned an intermediate score), and ends at (and the highest score is assigned to) the most negative. To ensure that a respondent defines a negative item as the whole continuum, the response format must assign scores along the full continuum. The following response format appears to do the job for *sad:*

Circle the number that describes your present mood:

−7 −6 −5 −4 −3 −2 −1 0 1 2 3 4 5 6 7

(very happy) (neutral) (very sad)

We define a *strictly bipolar format* as one that explicitly specifies the bipolar opposites and that succeeds in assigning scores along the full underlying bipolar continuum.

Defined as part of the continuum. Bipolarity also allows another conceptual and operational definition of any affective item. Because the bipolar formats just given presuppose bipolarity, researchers reasoned that PA and NA must be separated and unipolar formats must be used. In this case, according to the bipolar model, the positive item is being defined as a part of the full continuum, namely the part above neutral. Similarly, the negative item is being defined as a part of the full continuum, namely the part below neutral.

Here is a response format that seems to do the job for *happy:*

Do you feel happy? yes no Circle one. If you circled yes, please indicate how much:

1——2——3——4——5——6——7

(slightly) (moderately) (extremely)

The response of "no" is assigned a zero. Similarly, here is a response format that seems to do the job for *sad:*

Do you feel sad? yes no Circle one. If you circled yes, please indicate by how much:

1——2——3——4——5——6——7

(slightly) (moderately) (extremely)

We define a *strictly unipolar format* as one that succeeds in defining an item as the appropriate part of the full underlying continuum. With a strictly unipolar format, a neutral feeling is assigned the lowest possible score; in this example, it is the zero as indicated by circling "no."

Actual response formats. Whether a unipolar response format in actual use is strictly unipolar is an empirical question and depends on how respondents subjectively construe the options provided. For instance, how do they construe the option "not at all" as a response to the item "I am in a good mood"? They might take it to mean "no positive amount of good mood" or, then again, they might take it to mean "I am in a quite bad mood." Similarly, what does "strong disagreement" mean as a response to the item "I am in a cheerful mood"? Does it mean neutral or gloomy?

The bipolarity thesis—that people's implicit response dimension is really a bipolar continuum—suggests that respondents might readily reinterpret ostensibly unipolar formats as if they were bipolar. Consider the following item: "Please describe your mood right now" coupled with the following apparently unipolar response scale:

1 2 3 4 5 6 7

not happy happy

The respondents are allowed to construe each option as they see fit. How do they construe "not happy"? Do they take it to mean "neutral (a lack of happiness)" as required for a strictly unipolar format? Or, do they take it to be synonymous with "sad or miserable"? If the latter, the format is effectively bipolar.

To explore the specific item just discussed, Carroll and Russell (1998) showed it to 20 respondents drawn from the general public and asked them to supply a word describing each response option. No respondent construed the seven options as required for a strictly unipolar format. Different respondents defined the various options differently, but a typical respondent placed "neutral" not at Option 1 but around the middle of the scale, usually 4. (Two independent raters judged the neutral point implied by the full set of words each respondent supplied; neutral fell in the range of 3 to 4.5, $M = 4.0$, on the 1 to 7 scale.) Most telling were the labels supplied for Option 2; rather than meaning a bit of happiness as required for a strictly unipolar scale, Option 2 was thought by all respondents to be negative (e.g., sad, glum, bad). These respondents had apparently defined this ostensibly unipolar response scale as bipolar, extending from NA through a neutral point around 4 to PA.

What about other formats? Table 1 shows five hypothesized classes of response format (I–V). We know of no way to verify the exact nature of each type (this is a job for future research), but inspection suggests the following hypothesis. Let us begin with the last one shown. The last format type is likely strictly bipolar in that both bipolar opposites are specified (happy and sad) and neutral is explicitly placed in the middle of the scale. All the other formats are ostensibly unipolar in that no bipolar opposite is specified. Return to the top of the list. The strictly unipolar format explicitly divides responses into parts (happy and not happy with not happy assigned a zero). The remaining formats are more ambiguous. The format we call *ambiguous–likely unipolar* fails to specify where a neutral feeling should be placed but does imply that it would receive the lowest score. The format we label simply *ambiguous* leaves the respondent to define the anchors (Does "not at all" happy mean neutral or miserable?). The format we call *ambiguous–likely bipolar* again leaves to the respondent the definition of the anchors; but by providing degrees of not happy, it suggests that neutral should fall toward the middle of the format; that is, the bipolar symmetry of the response options violates the assumption of a strictly unipolar format and strongly suggests bipolarity.

Whether respondents interpret a specific item on a questionnaire as unipolar or bipolar depends not only on its response format but on the specific affective words involved (Carroll & Russell, 1998). Semantically unmarked words are typically used to refer to a whole dimension, whereas marked words refer to part of a dimension (Clark & Clark, 1977). For example, the unmarked word *happiness* can refer to the full happiness–unhappiness dimension, whereas the word *unhappiness,* which is marked by the prefix *un,* refers to only part of that dimension.

Which formats are legitimate? One cannot test bipolarity with a response format that imposes bipolarity on the respondent. Thus,

Table 1
Five Hypothesized Types of Response Format

Type	Code	Example
I. Strictly unipolar	1	If you feel happy, tick here ____. If you ticked, please indicate by how much: 1——2——3——4——5——6——7 (slightly) (moderately) (extremely)
II. Ambiguous–likely unipolar	2	Circle the number that describes the degree to which the statement "I am happy" describes your present mood: 1————2————3————4————5 (not at all or slightly) (moderately) (very much)
III. Ambiguous	3	Circle the number that describes the degree to which the statement "I am happy" describes your present mood: 0——1——2——3——4——5——6 (not at all) (very much)
	4	Circle the number that describes the degree to which you agree with the statement "I am happy": 1———2———3———4 (not at all) (very well)
	5	Circle the number that describes the degree to which you feel happy: 1————2————3————4 (definitely do not feel) (cannot decide) (slightly feel) (definitely feel)
	6	Circle the number that describes the degree to which you feel happy: 1————2————3————4 (not at all) (a little) (quite a lot) (extremely)
IV. Ambiguous–likely bipolar	7	Circle the number that describes the degree to which you feel happy: 1————2————3————4 (definitely do not feel) (do not feel) (slightly feel) (definitely feel)
	8	Circle the number that describes the degree to which you agree with the statement "I am happy": 1————2————3————4————5 (strong disagreement) (strong agreement)
	9	Circle the number that describes your mood: 1——2——3——4——5——6——7 (not happy) (happy)
	10	Circle the number that describes the degree to which the statement "I am happy" describes your present mood: 1————2————3————4 (not at all) (not very well) (somewhat) (very well)
V. Strictly bipolar	11	Circle the number that describes your present mood: −7 −6 −5 −4 −3 −2 −1 0 1 2 3 4 5 6 7 (very sad) (neutral) (very happy)

Note. No distinction was made between response formats that differed only in the number of response options or in seemingly irrelevant phrasing or configuration of the question.

strictly bipolar formats are out. That respondents treat ambiguous but ostensiby unipolar formats as bipolar might, at first blush, suggest that those formats too are illegitimate in tests of bipolarity. However, in this case, the researcher does not impose bipolarity on the respondent; no bipolar opposite is specified. Rather, the respondent apparently imposes bipolarity onto what the researcher believed was a unipolar format and one that, indeed, would allow a unipolar interpretation. We believe, therefore, that ambiguous formats allow a legitimate test of bipolarity. Their ambiguity does, however, introduce another problem, to which we now turn.

What Correlation Indicates Bipolarity?

Finally, we arrive at the key question in the debate surrounding bipolarity: What is the correlation between bipolar opposites?

At times, bipolarity is almost equated with a correlation of −1.00. Most writers (including Russell, 1979) implicitly assumed that the correlation coefficient is the proper statistic to describe a bipolar relation, that bipolarity requires a negative correlation of large magnitude between the alleged opposites, and that the required correlation is invariant with time frame and operational definition of the hypothesized opposites. Two exceptions are Diener and Iran-Nejad (1986) and Van Schuur and Kiers (1994). The specific value of the predicted correlation is rarely mentioned, but when a number is specified, it is −1.00, although, of course, no one expects an actual observed correlation to match the predicted value precisely. What is more telling is that a correlation of −.70 has been assumed to be a challenge to bipolarity and to be evidence of independence (Tellegen et al., 1994). None of these assumptions is correct.

What is needed is a point value for the theoretic correlation between PA and NA predicted by a bipolar model. By *theoretic correlation,* we mean the correlation predicted for the population of true scores (i.e., free from all error). To derive a prediction from a bipolar model (or any other), it is necessary to make assumptions that go beyond bipolarity per se. We assume, for example, that each variable is measured on a continuous dimension, that the frequency distribution of the underlying bipolar affect dimension is normal, and that the mean on that distribution is the point that divides PA from NA. Although the theoretical model we offer is highly simplified, some model is essential. Only with a theoretic value can a researcher then decide, based on the quality of the data and reasonableness of the assumptions for a given case, whether specific data are or are not consistent with bipolarity.

The theoretic correlation predicted by a bipolar model is not, as commonly assumed, invariant. Carroll, Russell, and Reynolds (1997) provided mathematical arguments for predicting the theoretic correlation, and their results are shown in Figure 3. The correlation varies both with item semantics (represented by the angle between the variables in Figures 1 and 2) and with response format. In a later section, we show how it also depends on time frame. Here, we illustrate those results by focusing on four idealized cases that are central in interpreting the available data on the bipolarity of affect.

Case 1. Suppose that items or scales are selected for PA and NA that, according to the semantic analysis summarized in Figures 1 and 2, are separated by 90°. In this case, the theoretic correlation between them is zero, whether they are each defined as the whole underlying continuum or as a unipolar part. The effect of response format disappears at 90°. Further, at 90°, random error does not bias the observed correlation. For example, 90° is approximately the angle separating PA/HighAct and NA/HighAct in Figure 2. If the angle is exactly 90°, and if we set aside systematic error, our bipolar model anticipates a robust empirical correlation between these two sets near .00, whatever the response format or the reliability of the two scales involved.

Case 2. In the remaining cases, suppose that items are selected for PA and NA that are separated by 180°. At 180°, response format is a powerful force. In this case, suppose that PA and NA are each conceptually and operationally defined as a full underlying bipolar continuum (i.e., strictly bipolar formats). The relation between the two variables is linear: specifically, scores on one are simply the inverse of scores on the other. The correlation coefficient is an appropriate statistic to describe that relation. The theoretic correlation between them equals −1.00.

Case 3. Again consider two items 180° apart, but now suppose that each item is conceptually and operationally defined as exactly half of the underlying bipolar continuum, the median of which is zero. In this case (i.e., strictly unipolar formats), scores on the two items are mutually exclusive, their relation is nonlinear, the correlation coefficient is not the proper statistic to describe their relation, and the theoretic correlation (in error-free measurement) is not −1.00. The theoretic correlation is −.467. This prediction can be seen in Figure 3, and the precise value is given by Carroll et al. (1997).[2]

Why is the correlation coefficient not appropriate in this case? When PA and NA are each defined as parts, their relation is not linear. On our bipolar model, a person's true score on a full bipolar continuum falls into either the PA region or the NA region but not into both. Suppose it falls in the NA region; wherever it falls within the NA region, it is simply not in the PA region—there is no PA score other than not-PA, which is assigned a zero in a strictly unipolar format. Therefore, a zero on PA is consistent with any score on NA. Respondents who score zero on PA have not specified whether they feel neutral or negative; and if they feel negative, they have not specified how negative.

Why is the correlation so modest in magnitude in this case? Strictly unipolar formats assess only part of the information of a bipolar format, and it is this loss of information that accounts for the reduction in the theoretic correlation. For two scales to correlate −1.00, the information provided by one scale must be completely redundant with the information provided by the other. Figure 4 shows what is happening graphically. In each case, the axes consist of 7 options assumed to form a strictly unipolar scale. Imagine that these unipolar scales, one for happy and the other for sad, are administered to a sample of respondents. Imagine that happy and sad are actually bipolar opposites and that our measurement procedure is completely free from any errors of measurement. Ask yourself this: What would the results have to look like if they are to produce a correlation of −1.00? All data would have to fall as shown in Figure 4A. Those who rate themselves as moderately sad would also have to rate themselves as moderately happy—they could not rate themselves as not at all happy. This result would violate rather than confirm the basic assumption of

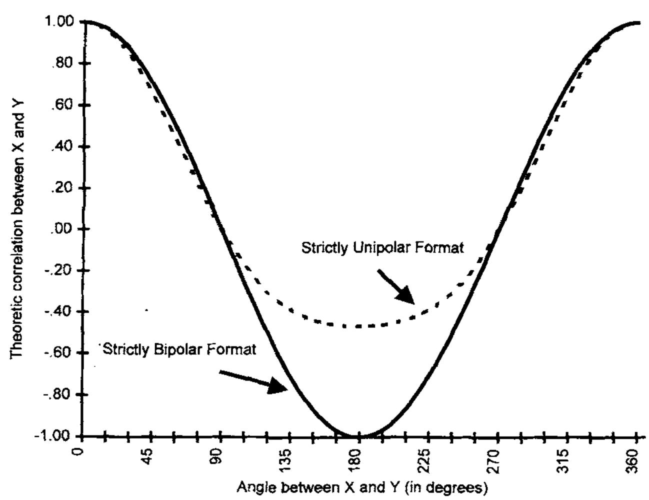

Figure 3. Theoretic correlation between two variables, X and Y, as a function of the angle between them and part (strictly unipolar format) versus whole (strictly bipolar format) operational definition.

[2] The prediction of $r = -.467$ can be derived independent of the formula developed by Carroll et al. (1997). Assume that X and Y are two mutually exclusive parts of a single continuum in standard normal form with zero as the point of division between X and Y, that both X and Y are assessed with strictly unipolar response formats, that both are scored with positive numbers, and that they have the L-shaped bivariate distribution shown in Figure 4B. Both X and Y are thus positively skewed. Let ρ be the correlation between X and Y: $\rho = [E(XY) - E(X)E(Y)]\ \{[E(X^2) - (E(X))^2]\ [E(Y^2) - (E(Y))^2]\}^{-1/2}$. In this case, $E(XY) = 0$, because either X or Y is always zero. $E(X) = E(Y) =$ the height of the normal curve at the mean of the normal distribution $= (2\pi)^{-1/2}$ (Glass & Hopkins, 1984, p. 72). $E(X^2) = E(Y^2) = 1/2$, because $E(X^2) = 1$ for a normal distribution and the distribution of X is normal except that half the values are zero. With substitutions, $\rho = (1 - \pi)^{-1} = -.467$. The ρ just calculated is the usual Pearson product–moment correlation. When the same assumptions about X and Y are made and data are sampled from the resulting populations, correlations between X and Y of approximately −.47 result. When the same data are converted to a polycotomous format and polychoric correlations are calculated, correlations of approximately −1.00 result. More generally, for two variables separated by an angle of Θ degrees, their theoretic polychoric correlation equals $\cos(\Theta)$. The possibility of using polychoric correlations was inspired by an article by Tellegen, Watson, and Clark (in press). We return to this possibility in footnote 6.

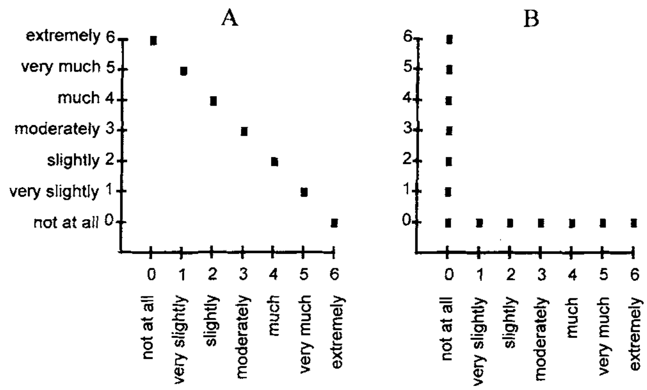

Figure 4. A: Data required to achieve a correlation of -1.00. B: Data required to show a bipolar relationship between two variables in which each is assessed with a strictly unipolar response format and therefore defined as mutually exclusive parts of a single bipolar continuum.

bipolarity that at a given moment a person falls at one and only one point along the full bipolar continuum. When strictly unipolar formats are used, ironically, an observed correlation of -1.00 between happy and sad would contradict bipolarity. With strictly unipolar scales, bipolar opposites yield the L-shaped distribution of responses shown in Figure 4B.

Case 4. Again consider items 180° apart, but now imagine the response format is not known to be either strictly bipolar or strictly unipolar. In this case, a predicted correlation is impossible to specify, except to say that it should fall between $-.467$ and -1.00. The more bipolar the format, the closer the correlation is to -1.00. As we indicated, we suspect that most ostensibly unipolar formats in actual use are ambiguous, and the theoretic value thus remains specified only to this range. The vagueness of prediction introduced here stems not from our model of bipolarity but from the ambiguity of the response format. The solution is research that establishes the properties of response format used.

Actual correlations. The preceding four cases are all based on highly simplified assumptions. We assumed that measurement occurred without error and that scores faithfully represented continuous variation. In Case 2, we further assumed a bivariate normal distribution. In Cases 3 and 4, we assumed a special distribution in which half of the scores are zero and the other half come from either the upper or lower half of a normal distribution. In actual data, these assumptions are unlikely to be met.

Nevertheless, these theoretic values serve as benchmarks against which observed values can be compared. Researchers have not known ahead of time the true relation between PA and NA. Instead, they gathered data and wanted to infer that relationship by calculating a correlation coefficient. To interpret the resulting evidence, it is therefore useful to know what correlation the bipolar model predicts even in cases when the correlation is inappropriate.

Our analysis has one further implication that should be underscored: In testing bipolarity, bipolarity cannot be equated with or inferred directly from a correlation coefficient, whatever its value. Under specified circumstances, bipolarity predicts a correlation of -1.00; under other circumstances, it is inconsistent with -1.00. Under specified circumstances, it predicts a correlation of $-.467$, but clearly a correlation of $-.467$ could arise from relationships other than bipolarity. To test bipolarity, obtained data must be compared to the actual predictions of a bipolar model for the specific method used to gather those data.

Watson, Clark, and Tellegen's PANAS

We can now turn to two actual scales that have played a large role in the debate concerning bipolarity and independence: the two scales of the PANAS (Watson et al., 1988). The positive cluster consists of active, alert, attentive, determined, enthusiastic, excited, inspired, interested, proud, and strong. The negative cluster consists of afraid, ashamed, distressed, guilty, hostile, irritable, jittery, nervous, scared, and upset. The response format provided is what we have called *ambiguous–likely unipolar* (see Code 2 in Table 1).

A number of writers have noted problems with the PANAS scales as measures of PA and NA in general (Carver, 1996; Green et al., 1993; Larsen & Diener, 1992; Morris, 1989; Nemanick & Munz, 1994). The PANAS items stand out as unlike those chosen by other researchers, and they violate the semantic requirements listed earlier for a test of bipolarity. The negative set includes none of the semantic opposites of the positive set. For *interested,* one might expect to see *bored*; for *strong,* one might expect *weak*; for *active,* one might expect *passive.* The clusters do not cover the full range of PA or NA. For PA, happy, positive, satisfied, serene, and pleased are missing and high arousal items predominate (active, alert, attentive, excited). For NA, unhappy, negative, depressed, and sad are missing, and high arousal items predominate (distressed, jittery, upset).

One could simply conclude that the PANAS scales are not relevant to a test of the bipolarity of PA and NA. A more useful analysis would be to state clearly the predictions of a bipolar model for these two sets of items. According to Watson et al. (1988), the positive items were chosen to tap a theoretical dimension (unfortunately given the very broad and unqualified name *Positive Affect* or *PA*) defined as activation plus pleasantness. The negative items were chosen to tap a theoretical dimension (given a name that is the antonym of the first scale: *Negative Affect* or *NA*) defined as activation plus unpleasantness. Because they are not opposite on activation, they are not opposites. In our terms, Watson et al. defined positive affect as PA/HighAct and negative affect as NA/HighAct, which are separated by 90° from one another in Figure 2. If the angle were indeed 90°, then the theoretic correlation between them would equal zero.[3] Feldman Barrett and Russell (1998) estimated the empirical angle to be 115°; if so, the theoretic correlation between them would be $-.423$ with a strictly bipolar format and $-.262$ with a strictly unipolar format.

[3] The prediction of a zero correlation here depends on the angle being exactly 90°. At 90°, response format does not alter the correlation, but if the actual angle is greater than 90°, the ambiguous but likely unipolar response format of the PANAS would limit the magnitude of any negative correlation. The prediction of a zero correlation does not depend on our specific model of bipolarity but follows more generally from the manner in which Watson et al. (1988) defined their constructs. Let y be activation, x be pleasantness, and $-x$ be unpleasantness. If we assume that pleasantness and activation are equal in their contribution, then, on this translation, their PANAS PA $= x + y$ and their PANAS NA $= -x + y$. If x and y are in standard score form, the correlation between $x + y$ and $-x + y$ is exactly zero. Because affect involves at least two dimensions, dimensions of affect can be defined that are independent of one another, and indeed, this formula shows how to do so.

Put differently, the extensive correlational research behind the PANAS shows that affect involves at least two substantive dimensions. Because affect is multidimensional, items can be selected, or clusters of items can be constructed, that bear various correlations to each other, including zero. Therefore, the clusters that are created—and what they are named[4]—are critical in any test of bipolarity. Our semantic hypothesis of Figures 1 and 2 does make a further prediction, however. Each dimension so created has a bipolar opposite. According to Figure 2, the opposite of the PANAS PA (in our terms PA/HighAct) is NA/LowAct. The opposite of the PANAS NA (in our terms NA/HighAct) is PA/LowAct. These predictions are borne out (Feldman Barrett & Russell, 1998; Yik, Russell, & Feldman Barrett, 1998; Watson & Tellegen, 1985; Watson et al., 1988). A bipolar interpretation of the dimensions underlying the PANAS is not a trivial matter. Although acknowledging bipolarity, Watson and Tellegen (1985) presented these two as the basic dimensions of mood, each conceptualized, labeled, and assessed as unipolar: high versus low PA and high versus low NA. In contrast, consider how low scores from each of the PANAS scales are understood on our bipolar interpretation. Whereas the phrase "low positive affect" might suggest mild comfort, we anticipate that low scores on their PA scale include states of depression, melancholy, and lethargy. Whereas the phrase "low negative affect" might suggest mild discomfort, we anticipate that low scores on their NA scale include states of relaxed tranquility and serenity.

Available Evidence

Zero-Order Correlation

We now review evidence on what has been considered the most definitive challenge to the bipolarity of PA and NA at a particular moment in time: the zero-order correlation. We were lenient in what we considered to be a moment. We include studies that have asked participants how they felt today, since this morning, right now, or about a brief incident.[5]

All studies to be reviewed in this section followed the same procedure of creating separate (ostensibly) unipolar scales of PA and NA. The correlation between the two scales was then calculated across participants. In what follows, our concern is the evaluation of bipolarity; but to evaluate properly, we must also consider measurement error, the semantics of the items, and the nature of the response format. We begin with studies that have not considered measurement error.

Uncorrected correlations and item semantics. Does the zero-order correlation between PA and NA vary with item content? Common sense, previous reviews, and evidence converge on an answer of yes (Lawton, Kleban, & Dean, 1993; Lawton, Kleban, Rajagopal, & Parmelee, 1992), Here, we quickly review one study (Watson, 1988), asking whether the variations are consistent with the semantic analysis presented earlier. This study did not control errors of measurement but did hold response format constant. The results are thus valuable less for the absolute magnitude of the observed correlations than for the relative magnitude when item pool changes.

Watson (1988) noted that Diener and Emmons (1984) had found correlations between PA and NA more negative than he had. Watson suggested item selection as one of the reasons. He then gave the same respondents three different item sets: (a) PANAS items, (b) items similar to those used by Diener and Emmons, and (c) items Watson termed *pleasant* versus *unpleasant.* All items were assessed with the same response format; in Table 1, it is called *ambiguous–likely unipolar.* (Hence, the absolute magnitude of any negative correlation can be predicted to be restrained; see Figure 3.) The specific items are given in Table 2. The PANAS items have already been discussed as sampling PA/HighAct and NA/HighAct. Diener and Emmons' set and Watson's pleasant–unpleasant set contained roughly antonyms, sampled widely from PA and NA. The correlations (also shown in Table 2) varied reliably with item pool. As expected, correlations from the PANAS were close to zero, whereas correlations from the other two item sets were more negative.

Uncorrected correlations and response format. Table 3 lists 31 data sets that yielded a zero-order correlation between PA and NA but that used a variety of response formats. We did not include in Table 3 studies that used the PANAS items, but we did include studies that used any other item pool.

The predicted correlation can be specified only roughly. Items varied across studies, and although the item pools were roughly semantic opposites, none were constrained to antonyms. The number of items was often small, and one might question how representative they were of PA or NA in general. In addition, these observed correlations failed to take into account errors of measurement. The results in Table 3 are thus valuable less for the absolute magnitude of the observed correlations than for the relative magnitude when response format changes. We used the tentative scheme of Table 1 to categorize the response formats; no

[4] With the benefit of hindsight, we believe that Watson & Tellegen's (1985) choice of labels for their concepts and scales was unfortunate. Evidence we review in this article shows that the PANAS scales do not measure what other positive and negative affect scales measure. *Positive affect* and *negative affect* were phrases used by earlier writers such as Bradburn (1969) and Costa & McCrae (1980) interchangeably with pleasant affect and unpleasant affect. For Bradburn and Costa and McCrae, the independence of PA from NA was a startling and important empirical finding. In contrast, Watson and Tellegen's system of definitions presupposes the bipolarity of pleasantness–unpleasantness, a distinction between pleasant and positive affect, and a distinction between unpleasant and negative affect. Their meaning of independence depends critically on the existence of another dimension of affect (engagement or activation). As Watson and Tellegen use the terms, to assert that PA and NA are independent is to assert that affect involves two independent substantive dimensions that can be additively combined.

[5] Ideally, we would focus on studies of affect at a particular instant, and this range of time frames is too liberal for momentary affect. Larson & Csikszentmihalyi (1980) found that mood at one moment bore no relation to mood just 3 hr earlier. Therefore, reports of daily mood surely mix together a range of events. As is described later in this article, the likely effect of too broad a time frame would be to obscure any negative correlation predicted by bipolarity of momentary affect. Later in this article, we suggest that bipolarity might be extended to *natural units* of time—blocks of time that the person thinks of as a single unit rather than as an aggregate of separate events. Likely, a day is a natural unit. We also doubt that questionnaires about mood should be mixed with questionnaires about affective reactions to specific events. Ignoring such distinctions, as we do, would operate against the model we propose; they are ignored here in the interest of having a large enough body of data to review, but future analyses might well want to separate these different topics.

Table 2
Correlation Between Different Sets of Positive and Negative Affect Items

Positive items	Negative items	Correlation: Moment (N = 574)	Correlation: Today (N = 657)
PANAS			
Active, alert, attentive, determined, enthusiastic, excited, inspired, interested, proud, strong	Afraid, ashamed, distressed, guilty, hostile, irritable, jittery, nervous, scared, upset	$-.14_a$	$-.12_a$
Watson's pleasant–unpleasant			
Happy, contented, at ease, calm, confident, friendly, sociable, warmhearted	Sad, blue, downhearted, alone, lonely	$-.40_b$	$-.42_b$
Diener & Emmons (1984) pleasant–unpleasant			
Happy, joyful, pleased, enjoying myself	Unhappy, frustrated, blue, angry, worried	$-.38_b$	$-.40_b$

Note. Results were reported by Watson (1988), who compared correlations within a column: Correlations between sets of items not sharing the same subscript were significantly different from one another ($p < .01$, two-tailed). PANAS = Positive Affect and Negative Affect Schedule.

response format was strict, but there was variation within the ambiguous category. We added a category for dichotomous formats (yes–no, true–false, and the like), which cannot be categorized as either unipolar or bipolar. In general, dichotomous formats are thought to be poor psychometrically.

All 31 correlations obtained were negative, ranging from −.25 to −.86, with a median of −.66. Because the dichotomous formats may introduce excessive measurement error, we focus on the polychotomous formats. Even here, the variability of correlations that is observed is much greater than the variability that is to be expected merely by sampling error, $\chi^2(26) = 412.0, p < .001$ (see Hunter & Schmidt, 1990). On the other hand, within each group with a similar type of response format, the variability is reduced, although it is still somewhat greater than what would be expected by sampling error alone in two of the three cases. Mean correlation obtained with an ambiguous–likely unipolar format was significantly closer to zero than the mean correlation obtained with either of the other two ambiguous formats. The more the format allowed a bipolar interpretation, the more negative was the resulting correlation.

Because of variations in items pool and because of the role of measurement error, the results of Table 3 cannot provide a rigorous test of our bipolar model. For both reasons, we must expect the obtained correlations to be attenuated relative to the theoretic values given earlier. Nevertheless, the pattern of results was highly consistent with our bipolar model. The mean correlation of −.41 that was obtained with the most unipolar format was only slightly closer to zero than the correlation predicted for Case 3 (−.467); and the mean of −.75 that was obtained with the ambiguous–likely bipolar format showed a substantial inverse relation.

Carroll and Russell (1998) provided a more direct test of the influence of response format on the correlation. They compared different response formats that were used to gather judgments of hot and cold (for verbally presented temperatures) as well as for happy and sad mood. The study was of single items, partly for simplicity, partly because single items can be valuable (Burisch, 1984), and partly because single items are the clearest candidates for bipolarity. Still, such data can only approximate the continuous distributions postulated in our theoretical model, and single items are especially subject to random and systematic measurement errors.

In one study, 120 participants were each given one verbally described temperature drawn from a normal distribution and then were asked to describe that temperature on 12 items, half for hot and half for cold. The items varied in format. In another study, 195 participants were asked to complete a small, anonymous mood survey. Participants were asked to describe their current mood with eight items, half for happy and half for sad. The items varied in format.

The results for both temperature and mood are shown in Table 4. The response format is keyed to Table 1. The observed correlations failed to match those derived theoretically in absolute magnitude, as would be expected from the nature of measurement error, from sampling, from the lack of a continuous scale, and so on. Still, the pattern of relative magnitudes was as predicted. The strictly bipolar format showed the strongest correlations (−.82 and −.79); the strictly unipolar format showed the weakest correlations (−.27 and −.46). The correlations yielded by ambiguous formats typically fell between these values (−.39 to −.57).

Figure 5 shows the most visually compelling summary of the mood data in the form of bivariate frequency distributions. The strictly bipolar format yielded the inverse relation that was anticipated, whereas the strictly unipolar format yielded the L-shaped, nonlinear relationship that was anticipated earlier in Figure 4B. Figure 5 also shows graphs for the two less strict formats. These latter results are more difficult to interpret, but they are consistent with the notion that some respondents interpreted the formats as unipolar but others interpreted them as bipolar. (That is, an L-shaped distribution plus a diagonal distribution coupled with

Table 3
Observed Correlation and Response Format, Momentary Affect

Study	Response format[a]	N	r_{pn}
Ambiguous format (likely unipolar)			
Watson (1988, Study 1)	2	574	−.38
Watson (1988, Study 1)	2	574	−.40
Watson (1988, Study 1)	2	657	−.43
Watson (1988, Study 1)	2	657	−.42
Mean correlation = −.41 (95% confidence interval: −.34 − −.47). $\chi^2(3) = 1.29$, *ns.*			
Ambiguous format			
Diener & Iran-Nejad (1986, Study 1)	3	72	−.39
Diener & Iran-Nejad (1986, Study 1)	3	72	−.38
Diener & Iran-Nejad (1986, Study 2, Group 1)	3	472	−.64
Diener & Iran-Nejad (1986, Study 2, Group 2)	3	472	−.71
Diener & Iran-Nejad (1986, Study 2, Group 3)	3	472	−.66
Russell (1979)	5	150	−.66
Russell (1979)	6	150	−.66
Green, Goldman, & Salovey (1993, Study 1, Time 1)	4	139	−.54
Green, Goldman, & Salovey (1993, Study 1, Time 1)	9	139	−.66
Green, Goldman, & Salovey (1993, Study 1, Time 2)	4	139	−.60
Green, Goldman, & Salovey (1993, Study 1, Time 2)	9	139	−.68
Green, Goldman, & Salovey (1993, Study 2)	4	250	−.69
Green, Goldman, & Salovey (1993, Study 2)	9	250	−.75
Feldman Barrett & Russell (1998, Study 2)	3	225	−.66
Feldman Barrett & Russell (1998, Study 3)	3	316	−.59
Mean correlation = −.65 (95% confidence interval: −.57 − −.72). $\chi^2(14) = 54.1, p < .01$.			
Ambiguous format (likely bipolar)			
Russell (1979)	7	150	−.71
Green, Goldman, & Salovey (1993, Study 1, Time 1)	8	139	−.64
Green, Goldman, & Salovey (1993, Study 1, Time 2)	8	139	−.67
Green, Goldman, & Salovey (1993, Study 2)	8	250	−.66
Feldman Barrett & Russell (1998, Study 2)	8	225	−.86
Feldman Barrett & Russell (1998, Study 2)	10	225	−.79
Feldman Barrett & Russell (1998, Study 3)	8	316	−.82
Feldman Barrett & Russell (1998, Study 3)	10	316	−.79
Mean correlation = −.75 (95% confidence interval: −.70 − −.81). $\chi^2(7) = 54.4, p < .01$.			
Dichotomous			
Russell (1979)		150	−.43
Green, Goldman, & Salovey (1993, Study 1, Time 1)		139	−.25
Green, Goldman, & Salovey (1993, Study 1, Time 2)		139	−.25
Green, Goldman, & Salovey (1993, Study 2)		250	−.40
Mean correlation = −.33.			

Note. An overall analysis for data gathered with any ambiguous format (i.e., excluding dichotomous formats) yielded a mean correlation of −.60 (95% confidence interval: −.52−−.67), $\chi^2(26) = 412.0, p < .001$. Although the samples are not always independent, we treated them as independent in the meta-analyses. N = number of observations; r_{pn} = correlation between positive and negative affect.
[a] Response format code is described in Table 1.

random and systematic error could yield the roughly triangular distributions seen.)

In all, variations in response format account for at least some, and possibly a fair amount, of the variation in the observed correlation between PA and NA. In evaluating the substantive hypothesis of bipolarity, it is necessary to remove this method-induced variance. Further, when interpreted in this light, the results so far are inconsistent with the claim that PA and NA are independent but are consistent with the predictions of our bipolar model.

Correlations corrected for measurement error. More telling data come from studies that considered measurement error. Although there are surprisingly few such studies, we can take very seriously the magnitude of the correlation obtained. These results are summarized in Table 5.

Begin with studies that used an ambiguous–likely bipolar response format. Early on, Russell (1979) found that both random and systematic measurement error had influenced the correlation. A disattenuated partial correlation estimated the latent correlation between PA/MediumAct and NA/MediumAct to be −.88. All the remaining results of Table 5 used structural equation modeling to control random and systematic error. Green et al. (1993) used their proposed multiformat procedure, which mixes together various re-

Table 4
Correlations Between Single Items With Different Response Formats

Response format[a]	Temperature (N = 120)	Mood (N = 195)
Strictly unipolar (1)	$-.27_{a}$	$-.46_{a}$
Ambiguous–likely unipolar (2)	$-.39_{abc}$	$-.51_{ab}$
Ambiguous (4)	$-.51_{c}$	$-.57_{b}$
Ambiguous–likely bipolar (8)	$-.34_{ab}$	
Ambiguous–likely bipolar (9)	$-.45_{c}$	
Strictly bipolar (11)	$-.82_{d}$	$-.79_{c}$

Note. Each respondent received a series of items each with a different response format. Within a column, correlations not sharing the same subscript were significantly different from one another ($p < .01$, two-tailed). Data were taken from Carroll and Russell (1998).
[a] Response format code given in parentheses is described in Table 1.

sponse formats but with a result tantamount to an ambiguous–likely bipolar format. As they expected, both random and systematic error had attenuated the correlation between observed scores. Their estimates of the correlation between the latent PA and NA scores ranged from −.84 to −.92. These figures were based on items that Green et al. called happy versus sad and that we would call a mix of PA/HighAct–PA/MediumAct versus NA/MediumAct–NA/LowAct. In their Study 3, they also included a set of items resembling (but not identical to) the PANAS. With these latter items, the estimated correlation between latent PA and NA was −.58.

In Green et al.'s (1993) third study, the time frame had been expanded to a month and thus does not fit within our criteria of momentary affect. Still, their data have been replicated in other studies that were restricted to current mood. In a study that followed Green et al.'s procedure closely, Feldman Barrett and Russell (1998) obtained correlations of −.92, −.93, −.93 (when the item pool was PA/MediumAct versus NA/MediumAct), and −.48 (when the items were similar to those of the PANAS).

Finally, Carroll, Yik, Russell, and Feldman Barrett (in press) reanalyzed two additional data sets gathered with Green et al.'s multiformat procedure, with results close to those already described. More important, Carroll et al. included all six clusters defined earlier in Figure 2. The six clusters were treated as scales, and two figures were calculated. First, the six scales were empirically placed within a two-dimensional space through a procedure described by Fabrigar, Visser, and Browne (1997). From this placement, the angle between each pair of scales was estimated. Second, the correlation between each pair of scales was estimated with a structural equation model that considered both random and systematic error.

The estimated correlations for all oppositely valenced pairs are shown in Figure 6 as a function of the angle between them. Also shown is a theoretical prediction based on the assumptions that all variance due to errors of measurement had been removed, that the scales were perfectly valid, and that response format was strictly bipolar. These data make three important points: First, the correlation varied dramatically with item content. This massive variation occurred even though every pair consisted of one scale called *positive affect* and another scale called *negative affect.* Second, the estimated empirical correlations were monotonically related to the angle between the scales, as predicted by our bipolar model. Third, empirical correlations were uniformly closer to zero than the theoretic correlations. Our guess on the explanation for the last conclusion is that response format was not entirely bipolar, that each scale contains some substantive components other than valence and activation, that systematic method variance is not entirely eliminated by Green et al.'s method, and that violations of other assumptions of the correlation coefficient (such as bivariate normal distributions on continuous dimensions) occurred.

To summarize so far, when the response format is ambiguous–likely bipolar, and when measurement error is taken into account, the correlation between scales called PA and NA varies with item selection. For PA/MediumAct versus NA/MediumAct (i.e., clusters that include hypothesized bipolar semantic opposites), the estimated latent correlation is between −.84 and −.92. For PA/HighAct versus NA/HighAct (i.e., the dimensions underlying the PANAS scales), the estimated latent correlation is between −.42 and −.58.

Finally, we arrive at one study that used an ambiguous–likely unipolar response format. Tellegen et al. (in press) created two sets of items for PA and NA and used the same response format for both. One set contained PA/MediumAct versus NA/MediumAct items such as happy and sad, respectively. The other set consisted of only those items used in the PANAS. Tellegen et al. calculated polychoric correlations and used a structural equation model to estimate the latent correlation between PA and NA separately for the two item sets. In a first analysis, only random error was considered. For the happy–sad item set, $r = -.73$; for the PANAS set, $r = -.28$. In a second analysis, both random error and acquiescent response style were considered, although other systematic errors were not. For the happy–sad set, $r = -.92$; for the PANAS set, $r = -.43$.

At first glance, Tellegen et al.'s (in press) results are highly consistent with the other results presented in Table 5. But recall that the response format was ambiguous–likely unipolar, and recall that we anticipated that correlations would be greatly attenuated with this format. So, at second glance, Tellegen et al.'s correlations appear much too substantial to be consistent with our analysis. But, finally, recall that these correlations are polychoric. A polychoric correlation estimates the correlation in a normally distributed population from a data set that is skewed. (Of course, *skew* is precisely one of the features we anticipate to result when a unipolar response format is applied to bipolar concepts). As we discuss in Footnote 2, for measurement with strictly unipolar scales on our model, the theoretic polychoric correlation for items 180° apart is −1.00 and for items 115° apart it is −.423.[6]

[6] A question for further study is whether a polychoric correlation can legitimately be used to test bipolarity. A polychoric correlation assumes a bivariate normal distribution, and therefore the use of a polychoric correlation requires a test of that assumption. Tellegen et al. (in press) provided no such test, and we do not know how such a test could be conducted. A polychoric correlation assumes part of the population not seen in the sample data and would thereby seem to assume that the response dimension itself extends beyond the neutral point into a region that is consistent with a bipolar dimension (which extends in both directions) but not with a unipolar dimension (which begins at the neutral point). Based on our Figure 3 and on the assumption that their response format was strictly unipolar, we guess that in Tellegen et al.'s (in press) study, the latent product–moment correlation is −.467 for their happy–sad item set (assuming them to be 180° apart) and −.262 for their PANAS set (assuming them to be 115° apart).

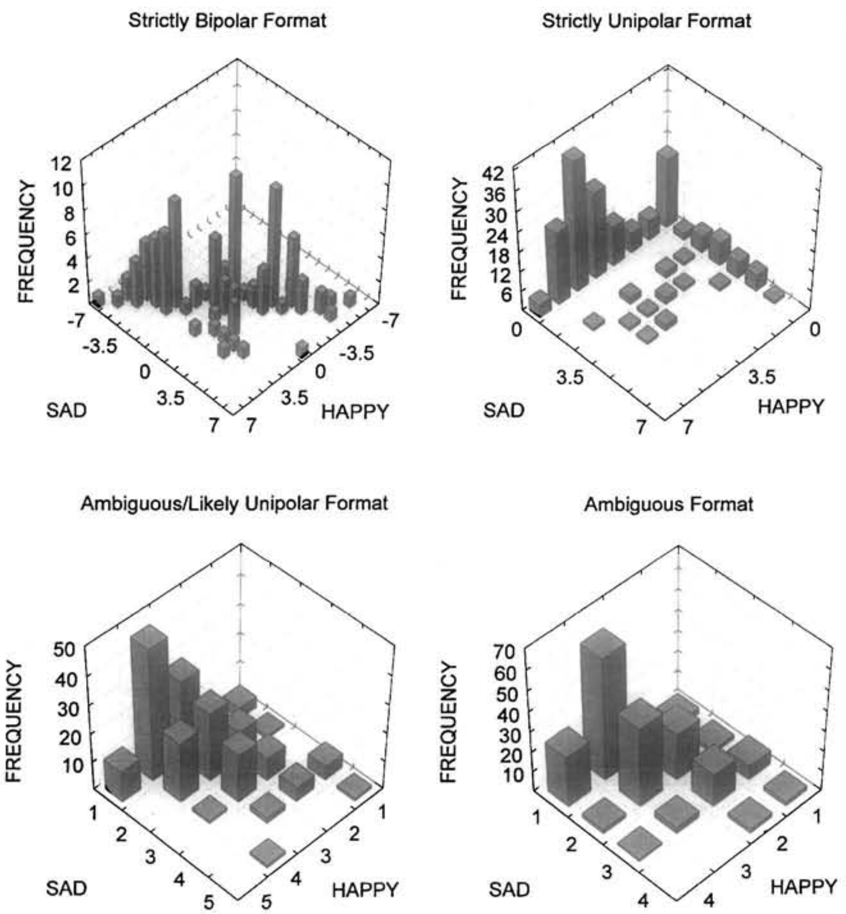

Figure 5. Bivariate frequency distributions for happy and sad assessed with four response formats (N = 195).

Summary. Tables 2, 3, 4, and 5 report correlations between PA and NA that range from −.12 to −.93. This huge range accounts for the vexing nature of the debate surrounding bipolarity and independence. We believe that a coherent account of these correlations is possible. Indeed, we find the results beautifully consistent. The correlation varies systematically with errors of measurement, item selection, and response format.

Correlations with the smallest magnitude result from the PANAS items. When measurement errors and response format are taken into account, the latent correlation between PANAS–PA and PANAS–NA is in the range of −.42 to −.58. Although closer to zero than the correlation obtained with other item pools, this correlation is still substantial in magnitude, which is surprising given the great effort to create the PANAS scales to make them independent and given our analysis that predicts a correlation of about zero. Our guess is that the displeasure component of Watson et al.'s NA scale is weighted more heavily than the activation component, and thus the scales do not capture exactly the vectors separated by 90° as they were intended.

Bipolarity is tested with item pools that include hypothesized opposites. When measurement error and response format are taken into account, and when the item pools contrast the semantically opposite pairs PA/MediumAct and NA/MediumAct, then the latent correlation is in the range of −.84 to −.93.

Factor Analysis

Many reports of the independence of PA and NA stem from factor analyses. Indeed, in an uncountable number of studies, affect items were incidentally administered to the participants, and those items were then submitted to exploratory factor analysis with varimax rotation. The result—one factor labeled *positive affect* and another labeled *negative affect*—is so commonplace nowadays as to warrant little more than passing mention of the mounting evidence for the independence of PA and NA. We do not summarize all this evidence here but instead argue that this method cannot be decisive.

Problems that plague studies of the zero-order correlation—item selection, response format, and measurement error—reappear. Item pools that lack semantic opposites do not yield bipolar factors. Because affect is multidimensional and because of method factors such as acquiescence, factor analysis can yield several independent factors. Independence of substantive factors can be mistaken for evidence against bipolarity. Such problems should be obvious by now.

Perhaps less obvious is the impact of a unipolar response format. The basic building blocks of factor analysis are the linear relationships among the variables, typically assessed in the form of a correlation matrix. Therefore, the nonlinearity introduced by a

Table 5
Between-Subject Analysis of Momentary Affect When Errors of Measurement Are Controlled

Study	N	PA/HighAct and NA/HighAct	PA/MediumAct and NA/MediumAct
Ambiguous–likely bipolar response format			
Russell (1979, Meddis Format)[a]	150		−.88
Green, Goldman, & Salovey (1993, Study 1, Time 2)[b]	139		−.85
Green, Goldman, & Salovey (1993, Study 2)[b]	250		−.91
Green, Goldman, & Salovey (1993, Study 3)[b, c]	304	−.58	−.86
Green, Goldman, & Salovey (1993, Study 1, Time 1)[b]	139		−.84
Feldman Barrett & Russell (1998, Study 1)[b]	129		−.92
Feldman Barrett & Russell (1998, Study 2)[b]	225		−.93
Feldman Barrett & Russell (1998, Study 3)[b]	316	−.48	−.93
Carroll, Yik, Russell, & Feldman Barrett (in press; Boston)[b]	198	−.39	−.91
Carroll, Yik, Russell, & Feldman Barrett (in press; Vancouver)[b]	217	−.20	−.90
M		−.39	−.90
95% confidence interval		−.49 – −.29	−.92 – −.87
Ambiguous–likely unipolar response format			
Tellegen, Watson, & Clark (in press)[a, d]	284	−.43	−.92

Note. Separate meta-analyses were done for each column of correlations. For correlations between PA/HighAct and NA/HighAct, $\chi^2(4) = 17.3$, $p < .01$. For correlations between PA/MediumAct and NA/MediumAct, $\chi^2(11) = 57.2$, $p < .001$. Although the samples are not always independent, in the meta-analyses we treated them as independent. All correlation coefficients shown have been corrected for random error. PA/HighAct = positively valenced items that are high in activation; NA/HighAct = negatively valenced items that are high in activation; PA/MediumAct = positively valenced items that are medium in activation; NA/MediumAct = negatively valenced items that are medium in activation.
[a] Identifies studies where acquiescence was controlled.
[b] Identifies studies where systematic method error was controlled with Green, Goldman, & Salovey's (1993) procedure.
[c] Instructions asked about the last month.
[d] The correlation coefficient calculated was polychoric.

unipolar response format can produce an artifact. Consider the case of one completely bipolar dimension. We assess four variables, two items for positive affect (P1, P2) and two items for negative affect (N1, N2), with a strictly unipolar response format. Suppose that measurement is free of error. Table 6 shows the intercorrelations and factor loadings. The result is an incorrect two-factor solution, which, following varimax rotation, looks like two relatively independent unipolar factors. The reason for this artifact is simple: The correlation coefficient presupposes a linear relationship and is inappropriate in this case.

External Correlates

Zevon and Tellegen (1982) argued that "further evidence in support of distinct dimensions of Positive and Negative Affect" (p. 121) can be seen in the patterns of correlations that PA and NA bear to other (external) variables, which might be causes, consequences, or correlates of affect. For example, whereas PA correlates with Extraversion, NA correlates instead with Neuroticism (Costa & McCrae, 1980; McFatter, 1994). Parkinson, Totterdell, Briner, and Reynolds (1996) considered such evidence to be the most persuasive case for the independence of PA and NA.

We do not review such evidence here because we believe that patterns of external correlations are not an appropriate test of bipolarity. When PA and NA are defined through item selection to be independent (as with the PANAS scales), they would be expected to have different external correlates.

When PA and NA are defined by unipolar scales of semantic opposites, we argue that patterns of external correlates can provide little or no support for or against a bipolar model. A practical problem is that the predicted magnitudes of correlations with external variables are almost inevitably smaller than the predicted correlations between (semantically opposite versions of) PA and NA themselves. Consequently, the examination of correlations with external variables provides a much less sensitive test of bipolarity than the examination of the correlation between PA and NA directly.

But the deeper problem is that this method presupposes a linear relationship between valence and the external variable. But perhaps the relationship is not linear. Suppose, for example, that PA and NA fail to show equal but opposite correlations with temperature. The explanation may be that people prefer intermediate temperatures over extreme hot or cold—a nonlinear relationship. Or, suppose that PA and NA fail to show equal but opposite correlations with Extraversion–Introversion. The explanation may be that people who score toward the middle of the personality scale (the ambiverts) are the happiest—again a nonlinear relationship. When unipolar formats force PA and NA to be defined as parts of the valence continuum, correlations with an external variable would be highly sensitive to any such nonlinearity. Our point here is not that relations of affect to temperature and Extraversion–Introversion are nonlinear but that when unipolar formats are used, such relationships would have to be known to be linear to test bipolarity by means of external correlations. They are not.

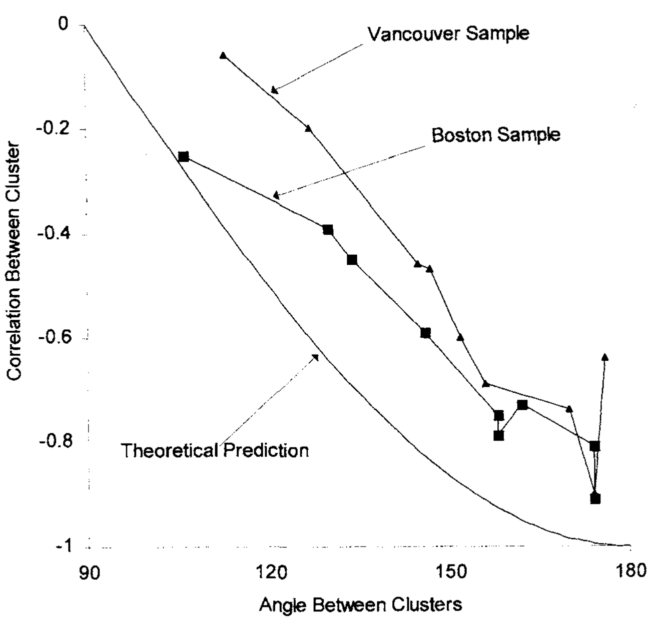

Figure 6. Estimated and predicted correlations between various positive affect and negative affect scales. Results were reported by Carroll, Yik, Russell, and Feldman Barrett (in press).

Conclusion About Momentary Affect

At a given moment, is PA independent of, or is it the bipolar opposite of, NA? This question presupposes that independence and bipolarity are mutually exclusive possibilities, but not all versions of independence are. Indeed, the widespread claims of independence require various clarifications. First, such claims are rarely phrased at a specific level (happy and sad, elated and depressed, tense and relaxed) but at a more abstract level (PA and NA). Advocates of independence need to clarify the logical connection between these two levels.

Second, independence of PA and NA has come in two different versions. Some writers (e.g., Bradburn, 1969; Costa & McCrae, 1980; Nowlis & Nowlis, 1956; Zautra et al., 1997) meant that what were traditionally thought of as bipolar opposites were, in fact, independent. On this version, genuine semantic opposites are claimed to be independent. This version is as surprising as one plus one is not two. Watson and Tellegen (1985) also claimed independence, but of a PA and NA redefined from what these terms had meant earlier. On Watson and Tellegen's definitions, the two independent dimensions are each bipolar. Their claim is equivalent to saying that affect involves two independent substantive dimensions and is consistent with what has always been believed. Confusion has resulted from treating these two versions of independence as if they were the same and from citing evidence gathered under one definition as supporting independence defined the other way.

The evidence that has been gathered to answer the question has been equally confusing. This evidence has relied on the correlation coefficient. Many different forces can shift that correlation toward zero. Unipolar response formats, errors inherent in measurement, inclusion of items that are not the hypothesized opposites, substantive confounds, and nonlinear relations can all produce correlations far from those once thought invariably predicted by bipolarity. Indeed, item selection alone can reduce the correlation to zero; thus, with error-free measurement and with strictly bipolar response formats, two scales that assess PA/HighAct and NA/HighAct can correlate near zero. In other contexts, it is a combination of forces that reduces the correlation to zero. Conversely, for a correlation to be near -1.00 (as many researchers had incorrectly defined bipolarity), all factors must simultaneously be right. Finding a substantial negative correlation is an uphill struggle, finding a weaker correlation is a downhill run. As a result, the literature on affect consists of a bewildering range of correlations.

In reviewing this evidence, we found no support for the independence of what were traditionally thought of as opposites. The counterintuitive quality, the surprise value, of the claim for independence comes from the version that is incompatible with bipolarity. The evidence for independence comes from the version that is compatible with bipolarity.

Does this same evidence establish bipolarity? No, but we are moving in that direction. The first step is to state what bipolarity predicts and how to test it. One cannot examine bipolarity per se, but only a specific and clear statement of bipolarity. We therefore defined a simplified and idealized but thoroughly bipolar model of momentary affect, and we laid out the major factors that influence the correlation between PA and NA: measurement error, item selection, and response format. When we compared the available evidence with the predictions of our model, we could find no basis whatsoever for a rejection of bipolarity.

Further, the available evidence begins to support bipolarity. No one piece of evidence is definitive, the evidence was not gathered to test our specific analysis, and the ambiguity of the response formats leaves all the evidence open to question. Still, the trend is clear: The more that biases against bipolarity are removed, the closer the data conform to the bipolar model. When time was restricted to a relatively brief period, when items were at least approximately the hypothesized opposites, when response format did not severely restrict negative correlations, and when measurement errors were controlled to the extent currently possible, then PA and NA emerged as substantially negatively correlated. In 11

Table 6
Theoretic Correlation Matrix and Factor Pattern Matrix for One Bipolar Dimension Assessed With a Strictly Unipolar Response Format

	True correlations				Factor loadings			
					Unrotated		Varimax	
Item	P1	P2	N1	N2	F1	F2	F1	F2
P1	—				−.86	.51	−.24	.97
P2	1	—			−.86	.51	−.24	.97
N1	−.47	−.47	—		.86	.51	.97	−.24
N2	−.47	−.47	1	—	.86	.51	.97	−.24

Note. The correlation matrix was specified a priori. P1, P2, N1, and N2 are items. F1 and F2 are factors from either common factor analysis or principal components analysis.

data sets that met these criteria, the mean correlation was −.90. Bipolarity has not been proven, but it is a good bet—for momentary affect.

Affect Extended Over Time

Diener and Emmons (1984) arrived at the conclusion that PA and NA are likely bipolar when examined at one time, but then they found evidence that PA and NA are independent when examined over time. The topic of affect extended over time is an intrinsically interesting one. For example, it necessarily arises in research on a person's subjective well being, the happiness of his or her life, temperament or characteristic mood, the cumulative effects of stress, and psychiatric problems of all sorts. When affect is viewed over time, the issues of bipolarity and independence arise (e.g., Bradburn, 1969) but in a more complex and potentially confusing manner.

We begin by distinguishing three methods that have sometimes been treated as equivalent but that are not. Each has very different implications. Indeed, the meaning of PA and NA changes from one to the other. The three are:

1. Within-S analysis. As in a study of momentary affect, the research participant provides momentary affect ratings but, in this case, continues to do so over an extended period of time (e.g., a diary or beeper study in which each participant provides momentary ratings once per day for 90 days). Each participant's data are analyzed separately.

2. Aggregation by researcher. Once again, the participant provides momentary ratings and continues to do so over time. Now, however, the researcher aggregates each participant's data in some manner.

3. Aggregation by participant. In this final case, the participants are never asked to rate their affect at one moment. Instead, the participants are asked for a global rating of the affect of a large chunk of time.

Within-S Analysis

With repeated momentary affect ratings, the questions of bipolarity and independence can be examined for each individual. Separate scales of PA and NA can be created. The raw data are the momentary ratings, and they are analyzed in an idiographic (within-S) fashion. For each participant, a correlation is calculated between PA ratings and NA ratings, with N equal to the number of days. (After the idiographic analysis, some summary statistic can be calculated, such as the mean correlation across participants. We think of this last step as a meta-analysis.)

The predictions of a bipolar model for this case parallel those for the case of momentary affect. For momentary affect, we discussed potential problems, including the effects of measurement error, expected correlation with unipolar scales, and item selection. Most such issues are identical and need not be repeated. Measurement error requires some comment.

Measurement Error

Random error produces the same effect in a within-S analysis as in a between-S analysis. One difference, of course, is that between-S data are independent. Ideally, every participant in a between-S study leads an independent life, and different participants do not influence one another. Independence cannot be guaranteed, and, indeed, it is highly unlikely in a within-S study. Imagine that 1 participant is consistently depressed for 1 month of a study of daily mood ratings. Those 30 ratings are in a sense 30 ratings of the same mood. The nominal N of, say, 90 days therefore is an inflated estimate of the number of independent ratings. In general, one may suppose that the nominal N in a within-S design exaggerates the number of independent observations. The result of this exaggeration is not a bias, but the correlation may be less replicable upon retesting than the nominal N would suggest.

Nonrandom error also appears in a within-S analysis. The major known source of a nonrandom error is acquiescence. As a difference between individuals, acquiescence contributes to the variance seen in a between-S analysis but is held constant in a within-S analysis. Acquiescence acts to shift a correlation in a positive direction; holding acquiescence constant should remove this one bias. All else being equal, a between-S correlation can therefore be expected to be somewhat more positive than the corresponding within-S correlation. Of course, the condition of "all else being equal" is hard to meet, but Watson and Clark (1997) came close by examining the same data set in both between- and within-S analyses. Their results were consistent with expectation. In seven of the eight cases, the mean between-S correlation was more positive than the mean within-S correlation. If acquiescence does indeed account for this difference, then the difference does not show that a bipolar model is more supported with within-S data and an independence model is more supported with between-S data. Acquiescence only makes it look that way.

Available Evidence

To our knowledge, no study has attempted to control random and nonrandom measurement error in a within-S analysis. Table 7 summarizes what results are available. The results for the PANAS item pool are separated from those from any other item pool. In

Table 7
Mean Correlation Between Observed Scores of Positive and Negative Affect for Within-Subject Data, Extended Affect

Study	Response format[a]	N	Obs	Mean r
PANAS items (ambiguous format likely unipolar)				
Watson (1988, Study 1)	2	123	44.5	−.20
Watson (1988, Study 1)	2	73	43.9	−.17
Watson (1988, Study 1)	2	80	44.4	−.31
Median correlation				−.20
Other items (ambiguous format)				
Diener & Emmons (1984, Study 3)	3	26	70.6	−.54
Diener & Emmons (1984, Study 4)	3	42	33.7	−.85
Diener & Emmons (1984, Study 4)	3	42	41.6	−.57
Diener & Emmons (1984, Study 4)	3	42	84	−.31
Diener & Emmons (1984, Study 5)	3	34	30	−.45
Glicksohn et al. (1996)	3	2	68	−.67
Median correlation				−.55

Note. N = number of participants; Obs = mean number of observations made by each participant; mean r = mean within-subject correlation across participants; PANAS = Positive Affect and Negative Affect Schedule.
[a] Response format code is described in Table 1.

studies that used the PANAS, correlations ranged from −.17 to −.31. In studies that did not use the PANAS, correlations ranged from −.31 to −.85. These results are entirely consistent with the uncorrected correlations seen earlier for between-S analyses (see Tables 2 and 3) and with our expectations. The differences that occurred could be due to the effect of controlling acquiescence noted previously and, for the non-PANAS item pools, to differences in item semantics and response format. In short, available results from within-S analyses reinforce our conclusions drawn from the study of momentary affect.

Aggregation by Researcher

We now consider another case in which each participant provides momentary ratings repeatedly over a period of time. In this case, however, the researcher aggregates each participant's data before the key correlation is calculated. For example, the researcher might average all the PA ratings to create an average positive score and average all NA ratings to create an average negative score. Or the researcher might count the frequency of happy days and the frequency of sad days for each participant. These derived scores (averages or frequencies), rather than the original momentary affect ratings, are then analyzed in some way, typically by calculating a single between-S correlation.

This particular case has been the most puzzling of a puzzling lot. Diener, Larsen, Levine, and Emmons (1985) discussed the dilemma of finding evidence that bipolarity seems to disappear in studies of the sort reviewed here. We make what may be the most surprising claim yet: When strictly unipolar scales are used to measure semantically opposite versions of PA and NA and when the ratings are then aggregated, any between-S correlation is consistent with bipolarity.

To understand research of this sort, a question must first be posed: What exactly are the predictions of a bipolar model? According to our bipolar model, when measured with antonyms, PA and NA constitute a single bipolar dimension (valence). Think of valence as the abscissa of a simple frequency diagram. The bipolar model requires that a participant's valence for a given moment can be properly represented as a single score somewhere along this continuum—that score constitutes one datum on the diagram. Suppose that ratings are taken daily. Each day, a new datum is provided; at the end of the study, all of the data from 1 participant are plotted. We call the resulting frequency distribution the individual's *Affect Distribution.*

All research of the type now being reviewed is really about properties of the Affect Distribution and comparisons of the Affect Distributions of different individuals or of the same individual at different times (odd numbered days vs. even numbered days or one month vs. another). In deriving predictions from a bipolar model, it is essential to distinguish the bipolar dimension on the abscissa from the frequency distribution above it. Only under some measurement procedures does the bipolar abscissa put constraints on the nature of the frequency distribution above it. Bipolarity per se does not specify the mean, variance, shape, or any other property of the Affect Distribution. For example, bipolarity does not specify that Sally's Affect Distribution for one set of observations bears any relation to her Affect Distribution for any other set. To be sure, common sense says that the distribution will have variance and be more or less bell shaped and that there is consistency in Sally's life. Individuals are likely to differ from one another in mean, variance, skew, and other properties of their Affect Distributions. A psychological theory might also specify that there will be a balance of PA and NA in one's lifetime (Parducci, 1995; Solomon, 1980). However, such possibilities are additional hypotheses to bipolarity per se, which demands none of these properties.

If the abscissa (valence) is assumed to be bipolar, what predictions does this specific assumption make about the Affect Distribution? The answer depends crucially on response format. We now consider how response format influences two of the measurement procedures typically used: proportions and means.

Proportions

On a bipolar continuum, there exists a neutral point, a threshold between PA and NA (or hot and cold or whatever). Suppose that the researcher calculates the proportion of cases (p) to the right of the neutral point and the proportion to the left (q). If no cases fall on the neutral point, then $p = 1 - q$. If the same procedure is repeated for each individual's Affect Distribution, then a p score and q score result for each individual. Across individuals, p and q correlate −1.00.

Suppose, however, that neutral is a region rather than a point; in this case, some cases are allowed to fall in this region. Instead of dichotomizing at the neutral point, two thresholds are used, resulting in a lower group of q scores, a middle group of m scores in the neutral region, and an upper group of p scores. In this case, knowing the value of p does not specify the value of q. If this procedure is repeated for each individual's Affect Distribution, then q, m, and p scores result for each individual. For any individual, $q + m + p = 1$. Even if all individuals use identical thresholds, p and q could have any correlation whatsoever because m is left to vary freely.

All this is clear in the abstract. In actual research, labels are given to these variables. Suppose that q is called *proportion of negative affect* and p is called *proportion of positive affect.* When $p + q = 1$, as in Diener et al.'s (1985) study, their correlation is −1.00. Of course, Diener et al. did not report this correlation, because it was a mathematical necessity given the way in which p and q were defined. When a set of m neutral ratings is allowed, however, as in Larson's (1987) study, then any correlation between p and q could result. Larson's reported correlations ranged from −.26 to +.26. In neither case is the correlation coefficient an interesting number to calculate.

Mean PA and Mean NA

More typically, the researcher calculates a mean PA and a mean NA score for each participant. A correlation between these two scores is then calculated across participants. This correlation is then used as a test of bipolarity. Again, the question is this: What, according to the bipolar model, are these two scores in terms of the individual's Affect Distribution and what is their theoretic correlation across individuals?

The answer depends on response format. Suppose that the abscissa (valence) is measured in its entirety. That is, PA is defined as the whole valence continuum and is assessed on a strictly bipolar rating scale going from *sad* through *neutral* to *happy*; neutral is defined as zero. The resulting scores yield the entire

Affect Distribution for one individual. The mean on this distribution is taken as mean PA. NA is defined as the inverse and is assessed on a second rating scale that is a mirror image of the first, going from *happy* through *neutral* to *sad.* This second rating scale also produces the entire Affect Distribution. The mean of this distribution is taken as mean NA. If we set aside the effects of random and nonrandom measurement error, the second Affect Distribution should be the mirror image of the first. Every score from the second equals the corresponding score from the first, multiplied by -1. Mean PA (the mean of the first distribution) is equal to the mean NA (the mean of the second distribution) multiplied by -1. The theoretic correlation between mean PA and mean NA across individuals would be -1.00 (in error-free data).

Suppose, however, that strictly unipolar response scales are used. PA is measured with a rating scale that concerns only cases above the zero (neutral) point of the valence dimension. NA is measured with a rating scale that concerns only cases that fall below the zero point of the same valence dimension. (In this case, zero corresponds to the neutral point, and the custom is to assign positive numbers to increasing degrees of NA.) Now, what are mean PA and mean NA, and what is their theoretic correlation across individuals? Because this is an especially important case, we dwell on it at length; but if your intuition is that the upper portion of the Affect Distribution puts no constraints on the lower portion, you have our bottom line. A more formal treatment of this topic is given in the Appendix (where we also demonstrate the parallel between our analysis and that of Diener et al., 1985). The importance of this question led us to illustrate our analysis with examples based on the bipolarity of temperature. Doing so helps bring out the nature of mean PA and mean NA when assessed with strictly unipolar scales; despite their names, we show that these are largely measures of the variability of the Affect Distribution.

Consider the case of temperature and, for the sake of argument, grant that hot and cold constitute a bipolar pair. Suppose that hot and cold are each measured on strictly unipolar response scales, such as the following:

1. *How hot is it right now?* The response is zero for temperatures at or below the worldwide annual median and is the actual temperature minus the annual median for temperatures above the annual median.

2. *How cold is it right now?* The response is zero for temperatures at or above the worldwide annual median and is the annual median minus the actual temperature for temperatures below the annual median.

Both scales result in positive numbers. Ratings are taken in various cities every day for a year. For each city, mean hot is the mean across days on Scale 1, and mean cold is the mean across days on Scale 2. For simplicity, suppose that all raters use the same zero point for the division into hot and cold (the worldwide median temperature). Suppose further that temperature is normally distributed within each city. Each city then has a Temperature Distribution. All scores above the zero point come from one rating scale; those below the zero point come from the other.

The mean of ratings on the hot scale is a measure of the distance of that subset of scores from the zero point. As such, it is a monotonically increasing function of the standard deviation of the overall Temperature Distribution. Indeed, under certain highly restrictive assumptions (such that the mean of the Temperature Distribution is zero and the distribution is normal with a standard deviation of σ; see Appendix, Equation 14), the following is the case:

$$\text{mean hot} = (2\pi)^{-1/2}\,\sigma = 0.4\,\sigma\,;$$

$$\text{mean cold} = (2\pi)^{-1/2}\,\sigma = 0.4\,\sigma.$$

However, more generally, mean hot and mean cold also vary with the mean on the overall distribution (see Appendix, Equations 5 and 6).

According to the bipolar model, what is the correlation between mean hot and mean cold, calculated across cities? When strictly unipolar response scales are used, the bipolar model does not yield a single expected correlation between mean hot and mean cold. It is difficult to prove a lack of prediction, but we can show that bipolarity is consistent with correlations ranging from -1 to $+1$.

First, consider two cities that differ in mean on the overall Temperature Distribution but have the same variance. The result might look like what is shown in Figure 7. Oslo is cold. Few days are above the zero point (worldwide median), and many days are below. So mean hot is small, and mean cold is large. Nairobi is hot. Most days are hot, and few are below the world median. Mean cold is far below mean hot. For these two cities, mean cold and mean hot are correlated -1.00.

More generally, imagine a fictional world in which this study is conducted. All cities have the same variance but differ in their means. A city that is hot in the summer is warm in the winter. A city that is cool in the summer is very cold in the winter. Something like this fictional world could exist if we examine cities that vary in latitude: A city near the equator has a high mean; a city near one of the poles has a low mean. In this world, the observed correlation between mean hot and mean cold would be negative. (To translate: According to the bipolar model, if all human beings had the same variance on their Affect Distribution and differed only in their mean level, then the larger the mean PA, the smaller the mean NA. Across individuals, the two would correlate negatively.)

Now consider two cities that differ in variance of temperatures but have the same mean. Boston and Vancouver yield the same overall mean; but in Boston, when it is hot, it is very hot, and when it is cold, it is very cold. In Vancouver, in contrast, it is never very cold or very hot. Summers are warm; winters are cool. The results might look like what is shown in Figure 8. For these two cities, mean cold and mean hot are correlated $+1.00$.

Imagine a fictional world in which this study is conducted. All cities have the same mean but differ in their variance. In this fictional world, the heat of the summer just balances the cold of the winter so that all cities have the same mean. This world would yield an observed positive correlation between mean hot and mean cold. (To translate: Imagine a fictional world in which all people have the same mean on their Affect Distribution but differ in their variance. Suppose further that the mean is zero. A person who experiences moments of great elation also experiences moments of severe misery. A person who is restricted to mild contentment even at the best of times experiences only mild unhappiness even at the worst of times. Brickman & Campbell, 1971, contemplated Helson's, 1964, adaptation-level theory and suggested that our world is such a world. In such a world, mean PA correlates positively with mean NA.)

Of course, in reality, cities (and people) vary simultaneously in

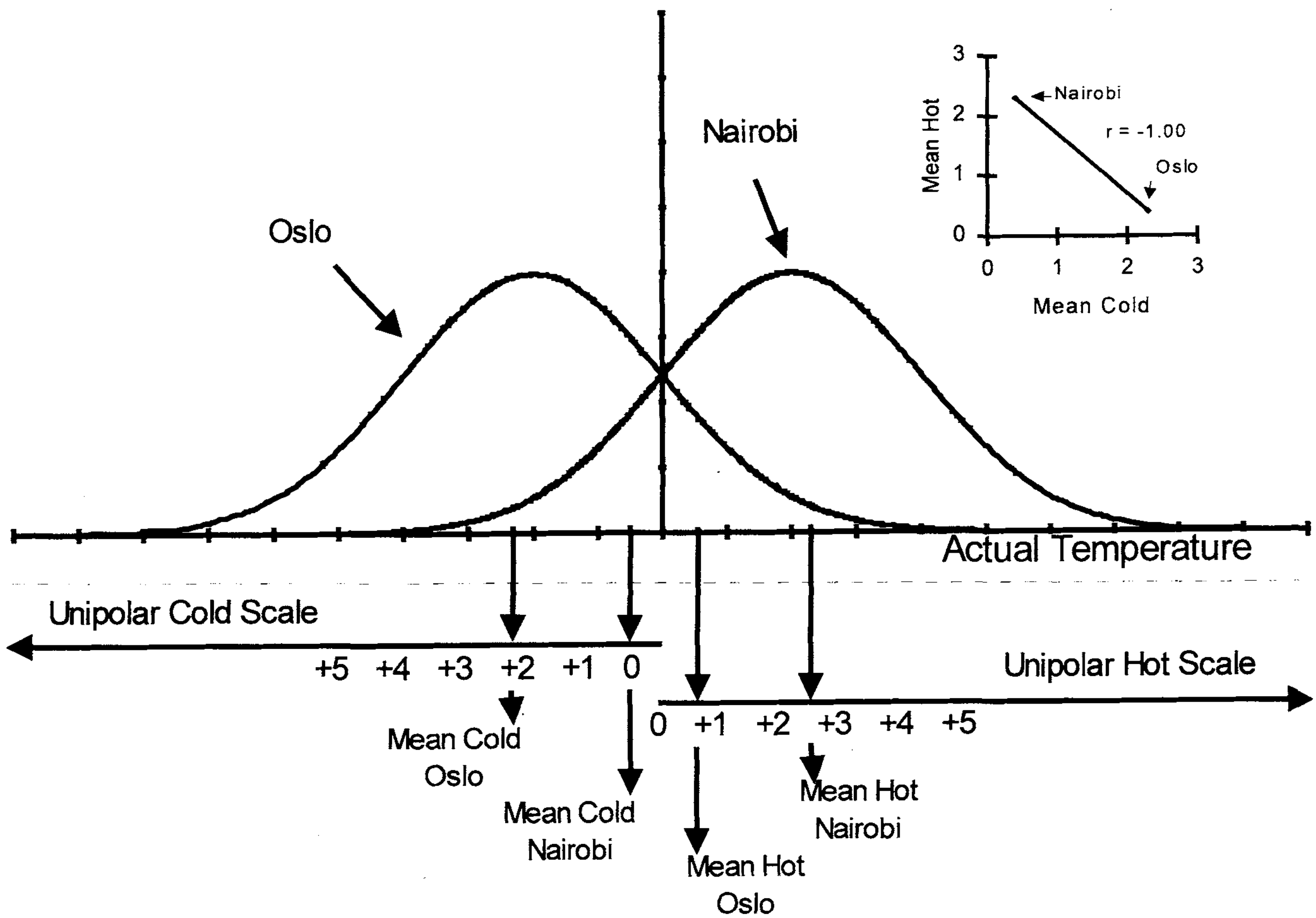

Figure 7. Fictional frequency distributions of temperature for Oslo and Nairobi. Both rated hot and rated cold are assigned positive numbers. In this case, the correlation between mean hot and mean cold is −1.00.

mean and variance. Thus, if we include all four cities of Figures 7 and 8 in one study, the correlation between mean cold and mean hot would fall between −1 and +1. These fictional cases thus establish that the bipolarity of hot and cold is consistent with a correlation between mean cold and mean hot of −1.00, +1.00, or something in between. Although the numbers are fictional for these cities, there is no reason to think that cities could not exist on this (or some) planet with these climates.

The bipolarity of hot and cold helps show that these statistics capture real phenomena: If heat were affect and cities were humans, Oslo would be frequently unhappy, suffering long winters of discontent. Nairobi would be frequently happy. Boston would be manic–depressive, with scorching summers of elation followed by severe winters of depression. Vancouver would be even-tempered—her PA would be a mild contentment, her NA would be a mild discomfort. Differences in variance are as important as differences in mean in determining an overall correlation. Of course, the analysis becomes more complicated in the real world, because the fictional world of temperature ratings had truly unipolar scales, error-free measurement, a normal distribution of the underlying temperatures, and the same temperature rating to divide hot from cold.

To summarize, we present a case in which bipolarity is assumed to be true: As demanded by bipolarity, in any one place at any one time, when it is hot, it is not cold, and when it is cold, it is not hot. Indeed, any specific temperature rating precludes any other. With this assumption, we showed that mean hot and mean cold (hence, mean PA and mean NA), assessed with strictly unipolar scales, could obtain any correlation whatsoever.

Available Evidence

The available evidence from studies of this design is summarized in Table 8. The results from the PANAS item pool are given separately. Results with other item pools suggest that mean PA and mean NA are negligibly correlated; observed correlations range from −.23 to +.26. Because bipolarity is consistent with any correlation obtained in a study of this sort, these figures say nothing about bipolarity.

What do these results mean? We might speculate that the near-zero correlations seen in Table 8 result from a balancing of forces. Differences among individuals in the mean on their Affect Distribution (a force for a negative correlation) are roughly equally balanced by differences among individuals in the variance of their

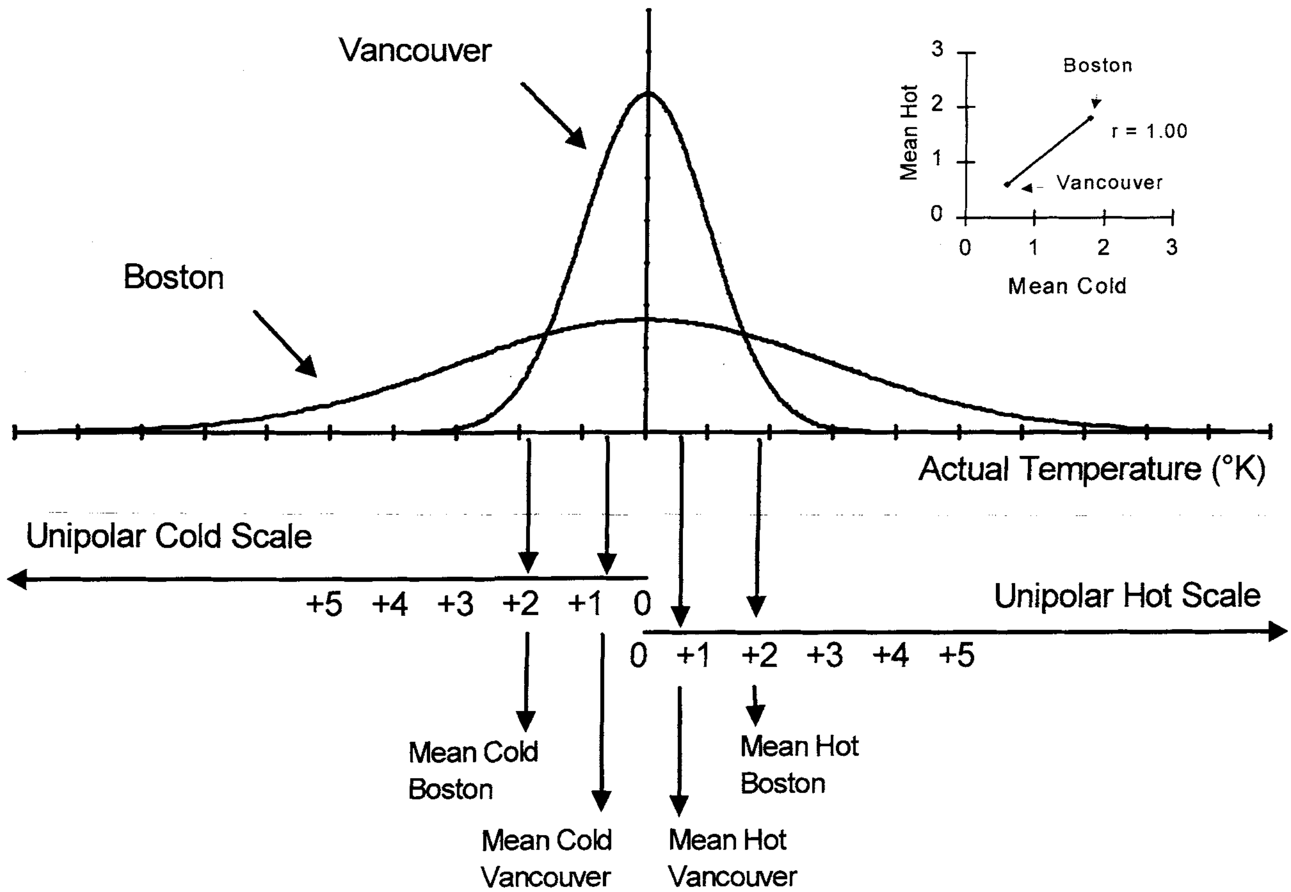

Figure 8. Fictional frequency distributions of temperature for Boston and Vancouver. Both rated hot and rated cold are assigned positive numbers. In this case, the correlation between mean hot and mean cold is 1.00.

Affect Distribution (a force for a positive correlation). The difficulty with this line of reasoning is that there are other factors influencing the correlation coefficient. Some individuals may project bipolarity onto an ostensibly unipolar rating scale, thereby pushing the correlation in a negative direction. On the other hand, individuals may differ in acquiescence, thereby pushing the correlation in a positive direction. (For certain response formats, a mean of repeated measures from the same individual can become a highly reliable index of acquiescence.) Of course, random noise pulls the correlation toward zero. Thus, there are too many unknowns to make much sense of the observed correlation.

Aggregation by Participant

In this final case, the participants are never asked about the moment. Instead, they are asked about large chunks of time. They are asked about their feelings over the last few weeks, months, or years, or over their entire life, or about their feelings in general. They might be asked the following: What was your average level of happiness? How frequently did you experience happiness? Have you ever been happy? What proportion of time were you happy? A standard between-S correlation is then calculated between the global rating for PA and the global rating for NA. In a study of this sort, the participants are being asked to re-create from memory their own Affect Distribution for a specified period of time and to report various facts about that distribution.

To test bipolarity with global ratings of this sort, the researcher faces two requirements. The researcher must state the prediction of an explicit bipolar model for the specific question asked and must demonstrate that humans are capable of answering that question in a sufficiently valid manner to test the prediction.

As to the first requirement, our bipolar model yields predictions only in certain circumstances. Thus, one must assume that human judges are exquisitely sensitive to the specific question asked. One must know how participants conceive of the particular chunk of time asked about. Consider four of the questions listed previously:

1. What was your average level of happiness over the last month? Of sadness? If the response format is strictly unipolar, then, as we have seen in the section on *Aggregation by Researcher*, any correlation between happiness and sadness is consistent with bipolarity. If the response format is strictly bipolar, then bipolarity predicts a theoretic correlation of -1.00, but even this prediction is not entirely clear. If a month is not a natural unit for participants, then they might not be able to conceive of it as a whole. Instead, a question about happiness might elicit memories of occasions different from those elicited by the question about sadness. For example, when asked "What was your average happiness?", the

Table 8
Between-Subject Correlations of Mean Positive Affect and Mean Negative Affect

Study	Response format[a]	*N*	Obs	Mean *r*
	PANAS items (ambiguous format likely unipolar)			
Watson (1988, Study 1)	2	123	44.5	.05
Watson (1988, Study 1)	2	73	43.9	−.01
Watson (1988, Study 1)	2	80	44.4	.10
Median correlation				.05
	Other items (ambiguous format likely unipolar)			
Diener & Emmons (1984, Study 3)	3	26	70.6	.26
Diener & Emmons (1984, Study 4, Measure 1)	3	42	33.7	−.15
Diener & Emmons (1984, Study 4, Measure 2)	3	42	41.6	.19
Diener & Emmons (1984, Study 4, Measure 3)	3	42	84	.09
Diener & Emmons (1984, Study 5)	3	34	30	−.23
Diener, Smith, & Fujita (1995)	[b]	≈212	≈52	.04
Median correlation				.07

Note. N = number of participants; Obs = mean number of observations made by each participant; mean r = mean within-subject correlation across participants; PANAS = Positive Affect and Negative Affect Schedule.
[a] Response format code is described in Table 1.
[b] Diener, Smith, and Fujita (1995) asked participant to rate the frequency of positive and negative affect experienced each day using a response format with the following anchors: 1 = *never,* 4 = *about half the time,* 7 = *always.*

participant might not conjure up the month as a whole but might remember the cluster of happy times at the beginning of the month and therefore report *their* average intensity. When next asked "What was your average sadness?", the participant might remember a different set of occasions, some sad times toward the end of the month, and then rate *their* average intensity. If so, bipolarity becomes mute about the possible correlation between happy and sad ratings.

2. How frequently did you experience happiness? Sadness? One key question is what each respondent takes to be the thresholds for happiness, for sadness, and for the neutral region in between. On our bipolar model, semantically opposite affect items form one continuous dimension with a neutral point or region in the center of that continuum. If all participants divide this continuum into two mutually exclusive categories and take the neutral point to be just a point that never actually occurs, then, roughly speaking, the frequency of happiness should be inversely correlated with the frequency of sadness. We say *roughly* because the division of affect into separate events is not a natural division. Imagine a man who feels moderately lousy continuously, except for the odd pick-me-up. How is he to answer a question about frequency? Strictly speaking, he might say that happiness is moderately frequent. Sadness is not frequent, but continuous. Consider another person. She takes the bipolar affective valence dimension to be implicitly divided into three roughly equal chunks: definitely unpleasant (q), relatively neutral (m), and definitely pleasant (p). The frequency of all three should be related, but any two (p and q) need not be.

3. Have you ever been happy? Sad? This question forces the rater to break the month into separate parts—to search for a happy event and to search separately for a sad event. One bears no relation to the other, and the bipolar model makes no predictions here and cannot be tested here. The analogy with temperature holds. Over the last month, has there been a hot day? A cold one? Nothing about the bipolarity of the hot–cold dimension precludes 2 hot days, 2 cold days, or 1 of each.

4. What proportion of time were you happy? Sad? This is probably a clearer version of the frequency question. If the participant divides the bipolar continuum into two mutually exclusive parts, then the bipolar model applies. But if the participant divides the continuum into three parts, with a neutral section dividing the two extremes, then no mathematical relation holds between any two.

The second requirement was that global ratings be sufficiently valid. Of course, validity is not an either–or matter. Humans can answer the questions listed previously, and their answers undoubtedly possess more validity than would random guesses. By *sufficient validity* we mean that the answers are accurate enough that any discrepancy from the predictions of a bipolar model could not plausibly be attributed to the reconstructive nature of memory or to biases or errors of judgment. For example, Schwarz and Clore (1983) found that questions about affect extended over time were influenced by current mood. Such a bias challenges any study about average levels of happiness over an extended period. Fredrickson and Kahneman (1993) described various studies in which participants provided both global affect ratings and moment-by-moment ratings. In general, the global ratings demonstrated what Fredrickson and Kahneman called *duration neglect:* The global ratings were insensitive to the proportion of time that was pleasant or unpleasant. This bias challenges any study that asks about proportions of time. Until sufficient validity of global ratings is demonstrated, any evidence gathered with this method remains open to question.

Of the three methods distinguished here for the study of extended affect, this third method is, to us, the least convincing. A review of published evidence using global ratings would require too much guesswork about just how the respondents interpreted the specific questions and conceived of the time period stated and would require too much faith in their answers. Still, we would not discourage studies about global ratings. People's answers to questions about extended affect and the processes used to arrive at those answers are fascinating and important topics in their own right. If used to study bipolarity, such data would best be complemented by the method we called *aggregation by researcher.* The most convincing evidence that an individual's global ratings are sufficiently valid would be a good equivalence of global ratings to the corresponding parameters of the individual's Affect Distribution.

Conclusion About Extended Affect

The study of affect extended over time encounters issues of bipolarity and independence. The findings in this area have been seen as a challenge to traditional beliefs, indeed to common sense. We disagree.

When measurements are taken in certain ways (a within-S correlation or, for aggregated data, proportions or means from fully bipolar scales), then our bipolar model does provide predictions. The within-S correlations obtained have been consistent with those predictions, but the other cases either seemed obvious or are mathematical necessities, and researchers have not calculated the correlations in these cases.

In many other cases, the data do not bear on bipolarity. For instance, when strictly unipolar formats are used, a correlation between mean PA and mean NA does not test bipolarity. Indeed, because of the many influences on this correlation, its meaning is difficult to determine and, under the circumstances we are familiar with, the correlation probably should not be calculated. Rather, we propose what seems to be a much simpler way of looking at this area. For each individual, a frequency distribution of affect scores is created. Then, the parameters of that frequency distribution can be examined. Each individual is thus characterized by statistics familiar to all: mean, variance, skew, and the like. (Larsen, 1992, discusses how such statistics and others might be put to good use.)

It is said in this literature that common sense teaches that mean PA and mean NA are highly negatively correlated. The finding of a low correlation between them was thus heralded as a counterintuitive finding. No evidence has been offered on just what common sense really says on this matter, and we have no such evidence either. But with an N of 2, we have a guess. Most ordinary people do not talk about means and correlations, but here are two ideas that they might endorse: (A) Some people are frequently happy; others are frequently unhappy. (B) Some people are intense, others are even-tempered. The intense people experience many moments of intense happiness in life but also experience many moments of intense unhappiness. The even-tempered people experience milder happiness but also milder unhappiness.

Our guess is that common sense endorses both of these propositions. Proposition A notes differences in means among individuals (in Diener et al.'s terms, differences in frequency of happiness). Proposition B notes differences in variance (differences in intensity). If A and B are both true, as seems likely, then common sense is in harmony with the published evidence and with our analysis. This is not to say that common sense is correct, but nothing so far challenges common sense (as we conceive it) on this matter.

Put differently, bipolarity says that when you are happy, you are not sad and that when you are sad, you are not happy. (Just as when you are hot, you are not cold.) Indeed, strictly speaking, the claim is very strong: Being at one point on the abscissa of an Affect Distribution precludes being at any other point—at a given moment. When ratings are aggregated across time (as in calculating a mean), this temporal linkage can be lost. When the link is lost, bipolarity cannot be assessed. Bipolarity does not specify where on the abscissa you will be at some other point in time. It does not specify how many happy days or sad days you will have and it does not specify just how happy are the happy days or how sad are the sad days—any more than the bipolarity of hot and cold predicts the climate in your city. After all, how could it?

Conclusion

Is a human being a pendulum betwixt a smile and a tear? Apparently so. Our simple but thoroughly bipolar model of affect provides a good fit to available data. Our review of the evidence turned up little or no substance to the psychometric challenge to bipolarity. For theories about affective feelings, bipolarity is a reasonable assumption. For the routine assessment of affective feelings, bipolar response formats are justified. (Of course, for tests of bipolarity, bipolar formats cannot be used.)

Like all scientific conclusions, ours is not final but is part of a dialectic. Nothing could seem more obvious than that people do not feel good and bad at the same time. But then it once seemed just as obvious that rocks are solid, that continents stay in one place, and that pandas are bears. The notion that PA and NA are bipolar opposites must be subjected to careful and continuing empirical scrutiny. Bipolarity is a fundamental assumption in our everyday thinking about affect and in many scientific accounts of affect. In science, all assumptions must be subjected to empirical scrutiny.

For bipolarity of affect, this scrutiny began when an unexpectedly weak correlation between PA and NA was found in a variety of research contexts. For over 40 years now, the scrutiny has continued, with much progress and much controversy. We now think we know why previous conceptual and empirical analyses produced controversy. Further progress (with perhaps less controversy) requires work on two fronts. Basically, the study of affect needs better models and better data.

By better models, we mean greater conceptual clarity, including explicit and precise models of independence, bipolarity, or any other alternative. Past controversy stemmed from the presentation of different versions of independence as if they were compatible or even identical. Independence in one context meant independence of what were traditionally thought of as opposites, but in other contexts, it meant independent components of affect. In yet another context, it had to do with the variability in an individual's affect distribution.

In the past, the bipolar view of affect seemed so obvious that it was not analyzed explicitly. And yet, our explicit bipolar model yielded surprising predictions about the correlation to be found in various circumstances. As an initial working model, ours required a number of simplifying assumptions and cannot be more than a first approximation. Further conceptual development can be achieved by stating more realistic bipolar models based, for example, on a more complex analysis of the semantics of affect. Our model explicitly concerned affective feelings, and part of the controversy stemmed from confusion over the definition of affect. We distinguished feelings from thoughts, and the study of thoughts might well yield different results. For example, nothing we have said in this article contradicts the notion that one can recognize both good and bad aspects of the same object or event.

By better data, we mean that the field needs better methods to test the models developed. Empirical testing of bipolarity turned out to be much more difficult than was imagined. This is not to say that bipolarity is nonfalsifiable. Bipolarity is a strong empirical hypothesis. But good research is difficult, as was research on the atoms in a rock, on plate tectonics, and on the genetic similarity between species. For many topics, good research requires precise measurement carried out within a framework that outlines the major influences on those measurements.

It is essential that those aspects of the data that arise from the nature of affect be separated from those aspects that are introduced by the process of measurement. In the controversy surrounding bipolarity and independence, this principle seems to have been forgotten until Green et al.'s (1993) powerful reminder. Their technique for the control of random and systematic error must be further developed and tested, and alternative approaches must be pursued; but the study of affect can ill afford another such lapse of memory. Tellegen et al.'s (1994, in press) alternative suggestion of direct assessment of systematic error should be examined, as should their recommendation to use polychoric correlations.

We were unable to predict the precise correlation between PA and NA when assessed with an ambiguous response format; that inability stems not from problems in the notion of bipolarity but from the widespread use of measuring instruments with properties that are not understood. We are uncomfortable that the principal data supporting bipolarity rely on an ambiguous response format. Provided that correct predictions are derived from an explicit bipolar model, strictly unipolar response formats provide a neglected but necessary test of bipolarity. An alternative research agenda could focus on how respondents interpret ambiguous formats. A complementary agenda could focus on the process whereby a person introspects a current affective feeling or remembers feelings over an extended period of time and translates the results into a rating on any questionnaire.

Available research relies too heavily on the correlation coefficient. Our model provides predictions about cases when the relation between NA and PA is not expected to be linear. For example, when strictly unipolar response formats are used (and other assumptions are made), our model predicts that the bivariate frequency distribution will have a specific L shape. In addition, the marginal frequencies of both separate unipolar scales will have a more positive skew than the corresponding bipolar distribution (Carroll et al., 1997). Bipolarity should be examined through univariate and bivariate frequency distributions.

The major challenge to bipolarity has been psychometric, and our analysis was therefore psychometric. Perhaps other approaches will raise other challenges. For example, Cacioppo and Berntson (1994) suggested that separate neurological processes underlie PA and NA. If so, and even if PA and NA are typically bipolar, they might be separable in specific circumstances. That possibility should be a high priority for empirical tests. For example, the evidence we have reviewed in this article came largely from asking respondents how they felt while sitting in a laboratory filling out a questionnaire. Perhaps different results would occur if different occasions were selected, such as moments of great emotion (but see Diener & Emmons, 1984) or times of conflict or decision. Perhaps different results would occur if affect were made operational through nonverbal means, such smiles and tears.

References

Averill, J. R. (1975). A semantic atlas of emotional concepts. *JSAS: Catalog of Selected Documents in Psychology, 5,* 330. (Ms. No. 421)

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

Borgatta, E. I. (1961). Mood, personality, and interaction. *Journal of General Psychology, 64,* 105–137.

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

Bradburn, N. M., & Caplovitz, D. (1965). *Reports on happiness: A pilot study of behavior related to mental health.* Chicago: Aldine.

Brickman, P., & Campbell, D. T. (1971). Hedonic relativism and planning the good society. In M. H. Appley (Ed.), *Adaptation-level theory* (pp. 287–301). New York: Academic Press.

Bullock, M., & Russell, J. A. (1984). Preschool children's interpretation of facial expressions of emotion. *International Journal of Behavioral Development, 7,* 193–214.

Burisch, M. (1984). You don't always get what you pay for: Measuring depression with short simple versus long and sophisticated scales. *Journal of Research in Personality, 18,* 81–98.

Bush, L. E., II. (1973). Individual differences multidimensional scaling of adjectives denoting feelings. *Journal of Personality and Social Psychology, 25,* 50–57.

Cacioppo, J. T., & Berntson, G. G. (1994). Relationship between attitudes and evaluative space: A critical review, with emphasis on the separability of positive and negative substrates. *Psychological Bulletin, 115,* 401–423.

Carroll, J. M., & Russell, J. A. (1998). *It is time to get strict with response format.* Unpublished manuscript.

Carroll, J. M., Russell, J. A., & Reynolds, S. (1997). *The psychometrics of bipolarity.* Manuscript submitted for publication.

Carroll, J. M., Yik, M. S. M., Russell, J. A., & Feldman Barrett, L. (in press). On the psychometric principles of affect. *Review of General Psychology.*

Carver, C. S. (1996). Some ways in which goals differ and some implications of those differences. In P. M. Gollwitzer & J. A. Bargh (Eds.), *The psychology of action* (pp. 645–672). New York: Guilford.

Clark, H. H., & Clark, E. V. (1977). *Psychology and language.* New York: Harcourt Brace Jovanovich.

Clyde, D. J. (1963). *Manual for the Clyde Mood Scale.* Coral Gables, FL: Biometric Laboratory, University of Miami.

Costa, P. T., Jr., & McCrae, R. R. (1980). Influence of extraversion and neuroticism on subjective well-being: Happy and unhappy people. *Journal of Personality and Social Psychology, 38,* 668–678.

Cuthbert, B. N., Bradley, M. M., & Lang, P. J. (1996). Probing picture perception: Activation and emotion. *Psychophysiology, 33,* 103–111.

Davison, M. L., & Srichantra, N. (1988). Acquiescence in components analysis and multidimensional scaling of self-rating items. *Applied Psychological Measurement, 12,* 339–351.

Diener, E., & Emmons, R. A. (1984). 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.

Diener, E., Larsen, R. J., Levine, S., & Emmons, R. A. (1985). Intensity and frequency: Dimensions underlying positive and negative affect. *Journal of Personality and Social Psychology, 48,* 1253–1265.

Diener, E., Smith, H., & Fujita, F. (1995). The personality structure of affect. *Journal of Personality and Social Psychology, 69,* 130–141.

Eich, E. (1995). Searching for mood dependent memory. *Psychological Science, 6,* 67–75.

Fabrigar, L. R., Visser, P. S., & Browne, M. W. (1997). Conceptual and methodological issues in testing the circumplex structure of data in personality and social psychology. *Personality and Social Psychology Review, 1,* 184–203.

Fehr, B., & Russell, J. A. (1984). Concept of emotion viewed from a prototype perspective. *Journal of Experimental Psychology: General, 113,* 464–486.

Feldman, L. A. (1995). Variations in the circumplex structure of mood. *Personality and Social Psychology Bulletin, 21,* 806–817.

Feldman Barrett, L. A., & Russell, J. A. (1998). Independence and bipolarity in the structure of current affect. *Journal of Personality and Social Psychology, 74,* 967–984.

Fredrickson, B. L., & Kahneman, D. (1993). Duration neglect in retrospective evaluations of affective episodes. *Journal of Personality and Social Psychology, 65,* 45–55.

Glass, G. V., & Hopkins, K. D. (1984). *Statistical methods in education and psychology* (2nd ed.). Boston: Allyn & Bacon.

Glicksohn, J., Gvirtsman, D., & Offer, S. (1996). The compensatory nature of mood: A single-subject time-series. *Imagination, Cognition and Personality, 15,* 385–396.

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.

Helson, H. (1964). *Adaptation-level theory.* New York: Harper & Row.

Hunter, J. E., & Schmidt, F. L. (1990). *Methods of meta-analysis: Correcting error and bias in research findings.* Newbury Park, CA: Sage.

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

Larsen, R. J. (1992). A process approach to personality psychology: Utilizing time as a facet of data. In D. Buss & N. Cantor (Eds.), *Personality psychology: Recent trends and emerging issues* (pp. 177–193). New York: Springer-Verlag.

Larsen, R. J., & Diener, E. (1992). Promises and problems with the circumplex model of emotion. In M. S. Clark (Ed.), *Review of personality and social psychology: Emotion* (Vol. 13, pp. 25–59). Newbury Park, CA: Sage.

Larson, R. W. (1987). On the independence of positive and negative affect within hour-to-hour experience. *Motivation and Emotion, 11,* 145–156.

Larson, R. W., & Csikszentmihalyi, M. (1980). Mood variability and the psychosocial adjustment of adolescents. *Journal of Youth and Adolescence, 9,* 469–490.

Lawton, M. P., Kleban, M. H., & Dean, J. (1993). Affect and age: Cross-sectional comparisons of structure and prevalence. *Psychology and Aging, 8,* 165–175.

Lawton, M. P., Kleban, M. H., Rajagopal, D., & Parmelee, P. A. (1992). The factorial generality of brief positive and negative affect measures. *Journal of Gerontology: Psychological Sciences, 47,* 228–237.

Lorr, M. (1989). Models and methods for measurement of mood. In R. Plutchik & H. Kellerman (Eds.), *Emotion: Theory, research, and experience: Vol. 4. The measurement of emotions* (pp. 37–54). New York: Academic Press.

Lorr, M., & McNair, D. M. (1982). *Manual: Profile of Mood States (bipolar form).* San Diego, CA: Educational and Industrial Testing Services.

Lorr, M., McNair, D. M., & Fisher, S. (1982). Evidence for bipolar mood states. *Journal of Personality Assessment, 42,* 432–436.

Lorr, M., & Shea, T. M. (1979). Are mood states bipolar? *Journal of Personality Assessment, 43,* 468–472.

Lorr, M., Shi, A. Q., & Youniss, R. P. (1989). A bipolar multifactor conception of mood states. *Personality and Individual Differences, 10,* 155–159.

Lorr, M., & Wunderlich, R. A. (1980). Mood states and acquiescence. *Psychological Reports, 46,* 191–195.

Mayer, J. D., & Gaschke, Y. N. (1988). The experience and meta-experience of mood. *Journal of Personality and Social Psychology, 55,* 102–111.

McFatter, R. M. (1994). Interactions in predicting mood from extraversion and neuroticism. *Journal of Personality and Social Psychology, 66,* 570–578.

McNair, D. M., & Lorr, M. (1964). An analysis of mood in neurotics. *Journal of Abnormal and Social Psychology, 69,* 620–627.

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

Meyer, G. J., & Shack, J. R. (1989). Structural convergence of mood and personality: Evidence for old and new directions. *Journal of Personality and Social Psychology, 57,* 691–706.

Morris, W. N. (1989). *Mood: The frame of mind.* New York: Springer-Verlag.

National Advisory Mental Health Council. (1995). Basic behavioral science research for mental health, a national investment: Emotion and motivation. *American Psychologist, 50,* 838–845.

Nemanick, R. C., Jr., & Munz, D. C. (1994). Measuring the poles of negative and positive mood using the Positive Affect Negative Affect Schedule and Activation Deactivation Adjective Check List. *Psychological Reports, 74,* 195–199.

Neufeld, R. W. J. (1975). A multidimensional scaling analysis of schizophrenics' and normals' perceptions of verbal similarity. *Journal of Abnormal Psychology, 84,* 498–507.

Neufeld, R. W. J. (1976). Simultaneous processing of multiple stimulus dimensions among paranoid and non-paranoid schizophrenics. *Multivariate Behavioral Research, 11,* 425–441.

Nowlis, V., & Nowlis, H. H. (1956). The description and analysis of mood. In *Annals of the New York Academy of Sciences* (Vol. 65, pp. 345–355). New York: New York Academy of Sciences.

Ortony, A., Clore, A. C., & Collins, A. (1988). *The cognitive structure of emotions.* New York: Cambridge University Press.

Osgood, C. E. (1969). On the whys and wherefores of E, P, and A. *Journal of Personality and Social Psychology, 12,* 194–199.

Parducci, A. (1995). *Happiness, pleasure, and judgment: The contextual theory and its applications.* New Jersey: Erlbaum.

Parkinson, B., Totterdell, P., Briner, R. B., & Reynolds, S. (1996). *Changing moods: The psychology of mood and mood regulation.* London: Longman.

Reisenzein, R. (1994). Pleasure-arousal theory and the intensity of emotions. *Journal of Personality and Social Psychology, 67,* 525–539.

Russell, J. A. (1978). Evidence of convergent validity on the dimensions of affect. *Journal of Personality and Social Psychology, 36,* 1152–1168.

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

Russell, J. A. (1989). Measures of emotion. In R. Plutchik & H. Kellerman (Eds.), *Emotion: Theory, research, and experience* (Vol. 4, pp. 83–111). San Diego, CA: Academic Press.

Russell, J. A. (1991a). Culture and the categorization of emotions. *Psychological Bulletin, 110,* 426–450.

Russell, J. A. (1991b). In defense of a prototype approach to emotion concepts. *Journal of Personality and Social Psychology, 60,* 37–47.

Russell, J. A. (1991c). Natural language concepts of emotion. In D. J. Ozer, J. M. Healy, Jr., & A. J. Stewart (Eds.), *Perspectives in personality: Self and emotion* (pp. 119–137). London: Jessica Kingsley.

Russell, J. A. (1997). How shall an emotion be called? In R. Plutchik & H. Conte (Eds.), *Circumplex models of personality and emotion* (pp. 205–220). Washington, DC: APA Press.

Russell, J. A., & Bullock, M. (1986). Fuzzy concepts and the perception of emotion in facial expressions. *Social Cognition, 4,* 309–341.

Russell, J. A., & Fehr, B. (1994). Fuzzy concepts in a fuzzy hierarchy: Varieties of anger. *Journal of Personality and Social Psychology, 67,* 186–205.

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.

Russell, J. A., & Mehrabian, A. (1977). Evidence for a three-factor theory of emotions. *Journal of Research in Personality, 11,* 273–294.

Russell, J. A., Weiss, A., & Mendelsohn, G. A. (1989). Affect grid: A single item scale of pleasure and arousal. *Journal of Personality and Social Psychology, 37,* 493–502.

Scherer, K. R. (1984). Emotion as a multicomponent process: A model and some cross-cultural data. In P. Shaver (Ed.), *Review of personality and social psychology* (Vol. 5, pp. 37–63). Beverly Hills, CA: Sage.

Schwarz, N., & Clore, G. L. (1983). Mood, misattribution, and judgments of well-being: Informative and directive functions of affective states. *Journal of Personality and Social Psychology, 45,* 513–523.

Smith, C. A., & Ellsworth, P. C. (1985). Patterns of cognitive appraisal in emotion. *Journal of Personality and Social Psychology, 48,* 813–838.

Solomon, R. L. (1980). The opponent process theory of acquired motivation: The costs of pleasure and the benefits of pain. *American Psychologist, 35,* 691–712.

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 disorders* (pp. 681–706). Hillsdale, NJ: Erlbaum.

Tellegen, A., Watson, D., & Clark, L. A. (1994, August). Modeling dimensions of mood. In L. A. Feldman (Chair), *Mood: Consensus and controversy.* Symposium presented at the 102nd Annual Conference of the American Psychological Association, Los Angeles, CA.

Tellegen, A., Watson, D., & Clark, L. A. (in press). On the dimensional and hierarchical structure of affect. *Psychological Science.*

Thayer, R. E. (1967). Measurement of activation through self-report. *Psychological Reports, 20,* 663–678.

Thayer, R. E. (1989). *The biopsychology of mood and activation.* New York: Oxford.

Van Schuur, W. H., & Kiers, H. A. L. (1994). Why factor analysis often is the incorrect model for analyzing bipolar concepts, and what model to use instead. *Applied Psychological Measurement, 18,* 97–110.

Warr, P., Barter, J., & Brownbridge, G. (1983). On the independence of positive and negative affect. *Journal of Personality and Social Psychology, 55,* 128–141.

Watson, D. (1988). The vicissitudes of mood measurement: 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. (1997). The measurement and mismeasurement of mood: Recurrent and emergent issues. *Journal of Personality Assessment, 86,* 267–296.

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.

Whissell, C. M. (1981). Pleasure and activation revisited: Dimensions underlying semantic responses to fifty randomly selected "emotional" words. *Perceptual and Motor Skills, 53,* 871–874.

Yik, M. S. M., Russell, J. A., & Feldman Barrett, L. (1998). *Structure of self-reported current affect: Integration and beyond.* Manuscript submitted for publication.

Zautra, A. J., Potter, P. T., & Reich, J. W. (1997). The independence of affects is context-dependent: An integrative model of the relationship between positive and negative affect. In M. P. Lawton & K. W. Schaie (Eds.), *Annual Review of Gerontology and Geriatrics, 17,* 75–102.

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

Appendix

Statistics of Aggregated Scores From Strictly Unipolar Scales Formed From a Bipolar Continuum

According to our bipolar model, positive affect (PA) and negative affect (NA) constitute a single continuum, referred to here as *valence.* Zero is the dividing point between positive and negative valence, with scores to the left assigned negative numbers and scores to the right assigned positive numbers. A person's affect at any moment (Xb) falls at a single point on the valence dimension. If an individual provides a set of momentary affect ratings on this valence continuum (e.g., once a day for a year), then a frequency distribution of scores can be created. The valence dimension is the abscissa; frequency is the ordinate. We refer to the resulting frequency distribution as the individual's affect distribution.

Suppose, however, that positive affect (Xp) is measured with a strictly unipolar rating scale that concerns only cases above the zero (neutral) point of the valence dimension (i.e., $Xp = 0$ for $Xb \leq 0$). Negative affect (Xn) is measured with a strictly unipolar rating scale that concerns only cases that fall below the zero point of the same valence dimension (i.e., $Xn = 0$ for $Xb \geq 0$). In assessing Xn, zero corresponds to the neutral point on valence, and it is the custom to assign positive numbers to increasing degrees of NA. Now what is mean PA (μp) and mean NA (μn) when unipolar scales are used, and what is their expected correlation across individuals?

According to our bipolar model, scores on the strictly unipolar scale of PA (Xp) are identical to those of the valence dimension (Xb) for values greater than zero and are equal to zero for values less than or equal to zero. Scores on the strictly unipolar scale of NA (Xn) are identical to the absolute value of scores on the valence dimension (Xb) for values less than zero and are equal to zero for values greater than or equal to zero. Therefore, for any single momentary rating,

$$Xb = Xp - Xn. \tag{A1}$$

Similarly, for the means of a population of momentary ratings,

$$\mu b = \mu p - \mu n. \tag{A2}$$

In this Appendix, we first show how mean PA (μp) and mean NA (μn) can be expressed as a function of the mean (μb) and standard deviation (σb) of the affect distribution. Then, we show why the bipolar model of affect is consistent with any reported correlation between mean PA and mean NA over time. Finally, we explore the relationship between the treatment developed here and the model developed by Diener, Larsen, Levine, and Emmons (1985). In our treatment, we assume each individual affect distribution is normally distributed with mean μb and standard deviation σb. (Note: μb is not necessarily the point dividing positive from negative on the valence dimension.)

Means on the Strictly Unipolar Scales

Mean PA (μp) and mean NA (μn) are each weighted means of the appropriate sections of the affect distribution. Mean PA is the mean of scores on the affect distribution that fall on the positive side of the valence dimension, weighted by the proportion of cases falling therein. Mean NA is the absolute value of the mean of scores on the affect distribution that fall on the negative side of the valence dimension, weighted by the proportion of cases falling therein.

The computational procedure requires three steps. First, we calculate the mean z score of the part of the affect distribution that falls on the positive side of the valence dimension (μpz^*) and the mean z score of the part of the affect distribution that falls on the negative side of the valence dimension (μnz^*). Following the general formula for obtaining the mean z score for a given section of the standard normal curve,

$$\mu pz^* = \frac{Ub - Ub\infty}{(p)} \tag{A3}$$

and

$$\mu nz^* = \frac{Ub - Ub\infty}{(1 - p)}, \tag{A4}$$

where Ub is the height of the normal curve corresponding to the z score of the affect distribution for the point that divides positive from negative on the valence dimension, $Ub\infty$ is the height of the normal curve corresponding to the z score of the affect distribution for the point that equals the largest possible score on the positive side of the valence dimension (this z score will always be equal to ∞ and the corresponding height on the normal curve will always be equal to zero), and p is the proportion of cases from the affect distribution that fall on the positive side of the valence dimension ($1 - p$ equals the proportion of cases that fall on the negative side of the valence dimension).

Next, we convert the mean z scores (μpz^* and μnz^*) to mean raw scores (μp^* and μn^*) based on the mean (μb) and standard deviation (σb) of the affect distribution, thus

$$\mu p^* = (\mu pz^*)\,(\sigma b) + \mu b \tag{A5}$$

and

$$\mu n^* = (\mu nz^*)\,(\sigma b) - \mu b. \tag{A6}$$

Finally, to obtain the mean score on the unipolar PA scale (μp), we weight (μp^*) by (p). Similarly, to obtain the mean score on the unipolar NA scale (μn), we weight (μn^*) by ($1 - p$). If we replace μpz^* and μnz^* with the formulas shown in Equations A3 and A4, then

$$\mu p = (Ub)(\sigma b) + (\mu b)\,(p) \tag{A7}$$

and

$$\mu n = (Ub)(\sigma b) + (\mu b)\,(p) - \mu b. \tag{A8}$$

We introduced Ub and p to compute μp and μn. Fortunately, both Ub and p can be computed from the mean (μb) and standard deviation (σb) of the affect distribution. To compute Ub in terms of μb and σb, we note that Ub is a simplified version of the function of the normal distribution,

$$U(x) = (2\pi\sigma^2)^{-1/2}\, e^{-((X-\mu)(X-\mu))/(2(\sigma)(\sigma))}, \tag{A9}$$

where $U(x)$ is the height of the normal curve above any given value for the variable X, e is the base of the system of natural logarithms (i.e., e = 2.71828 . . .), and μ and σ are the mean and standard deviation of the variable X.

In our treatment we use z scores; $X = -\mu b/\sigma b = Zv$ (the z score of the affect distribution for the point that divides positive from negative on the valence dimension), $\mu = 0$, and $\sigma = 1$. As a result of these simplifications,

$$Ub = (2\pi)^{-1/2}\, e^{-((Zv)(Zv))/(2)}. \tag{A10}$$

The proportion of cases that fall on the positive side of the valence dimension, p, can be computed from the cumulative distribution function of the normal curve, which can be expressed

$$p = \frac{1}{\sqrt{2\pi}} \int_{-Zv}^{\infty} e^{(-x^2/2)} dx. \tag{A11}$$

Mean PA and mean NA can now be written as functions of the mean and standard deviation of the affect distribution such that

$$\mu p = (2\pi)^{-1/2}\, e^{-((-\mu b/\sigma b)(-\mu b/\sigma b))/(2)}\,(\sigma b) + (\mu b)\left(\frac{1}{\sqrt{2\pi}} \int_{(-\mu b/\sigma b)}^{\infty} e^{-x^2/2} dx\right) \tag{A12}$$

and

$$\mu n = (2\pi)^{-1/2}\, e^{-((-\mu b/\sigma b)(-\mu b/\sigma b))/(2)}\,(\sigma b) + (\mu b)\left(\frac{1}{\sqrt{2\pi}} \int_{(-\mu b/\sigma b)}^{\infty} e^{(-x^2/2)} dx\right) - (\mu b). \tag{A13}$$

The formula for μn may appear different from that of Equation A6. This is simply because we multiplied out the expression $(\mu b)(1 - p)$.

Mean PA and Mean NA When μb Equals Zero

Equations A12 and A13 represent the general case, where the mean (μb) and standard deviation (σb) of the affect distribution can equal any value. For the special case where μb equals zero, the equations are very much simplified. If μb equals zero, $e^{-((-\mu b/\sigma b)(-\mu b/\sigma b))/(2)}$ equals 1 and drops from the equation; μb (p) and μb ($1 - p$) both equal zero and drop from the equation. In this case, μp and μn are equivalent, and the equations for μp and μn are simplified to the linear function of σb such that

$$\mu p = \mu n = (2\pi)^{-1/2}(\sigma b). \tag{A14}$$

Mean PA and Mean NA When σb Equals Zero

If σb equals zero, then all scores fall on a single point of the valence dimension. In this case, either μp or μn will equal μb, and the other will equal zero. Because all scores fall on one point of the valence dimension, (p) will equal 1 if μb is on the positive side of the valence dimension and will equal 0 if μb is on the negative side of the valence dimension. When the scores fall on the positive side of the valence dimension, μp equals μb and μn equals zero. When the scores fall on the negative side of the valence dimension, μn equals $-\mu b$ and μp equals zero.

Correlation Between Mean PA and Mean NA

Each individual is characterized by an affect distribution with its own statistical properties including μb and σb. The question addressed here concerns the correlation across individuals between μn and μp. We show that the bipolar model is consistent with a correlation as low as -1 and as high as $+1$. We turn to special cases to demonstrate our position. In the first case, μb is held constant across participants (i.e., μb is the same for all participants), and the predicted correlation between mean PA and mean NA is shown to be $+1$. In the second case, σb is held constant across participants (i.e., σb is the same for all participants), and the predicted correlation between mean PA and mean NA is shown to range from $-.467$ to -1, depending on the magnitude of σb.

Holding μb Constant

Rearranging Equation A2 shows that $\mu p = \mu n + \mu b$. If μb is a constant, then adding or subtracting a constant does not effect a correlation. Thus, μp and μn will show a perfect positive relationship.

(*Appendix continues*)

Holding σb Constant

If we hold σb constant across participants, then the correlation between mean PA and mean NA will depend on the magnitude of σb. As σb approaches zero, the correlation between mean positive and mean negative approaches $-.467$. As σb approaches infinity, the correlation between mean positive and mean negative approaches -1.00. First, look at what happens when σb approaches zero. We have defined Zv as $-\mu b/\sigma b$, where μb and σb are the mean and standard deviation of the affect distribution. As σb approaches zero, Zv approaches plus or minus infinity depending on the valence of μb. Now return to Equations A10 and A11. As Zv approaches plus infinity, both Ub and (p) approach zero. As Zv approaches minus infinity, Ub approaches zero and (p) approaches 1. If μb is positive, substituting these values in Equations A7 and A8 gives

$$\mu p = (0)(0) + \mu b\ (0) \tag{A15}$$

and

$$\mu n = (0)(0) + \mu b\ (0) - \mu b. \tag{A16}$$

If μb is negative, substituting these values in Equations A5 and A6 gives

$$\mu p = (0)(0) + \mu b\ (1) \tag{A17}$$

and

$$\mu n = (0)(0) + \mu b\ (1) - \mu b. \tag{A18}$$

It follows from Equations A15–A18 that whenever mean PA is not equal to zero, mean NA equals zero and that whenever mean NA is not equal to zero, mean PA equals zero. It also follows that mean PA and mean NA are always nonnegative values. Now, if we assume μb is normally distributed with mean $\mu_{\mu b}$ and standard deviation $\sigma_{\mu b}$, then the scatter plot of mean PA and mean NA would be L-shaped like the one shown in Figure 4B. Half the values for mean PA form half a normal distribution, the other half of the values equal zero. The same is the case for mean NA. In this case, the predicted correlation between mean PA and mean NA follows the derivation outlined in Footnote 2 (see Carroll, Russell, & Reynolds, 1997) and is equal to $-.467$.

Now, look at what happens when σb approaches infinity. We have defined Zv as $-\mu b/\sigma b$, where μb and σb are the mean and standard deviation of the affect distribution. So as σb approaches infinity, Zv approaches zero. Now return to Equations A10 and A11. As Zv approaches zero, Ub approaches $(2\pi)^{-1/2}$ and (p) approaches .5. If we place these values into Equations A7 and A8, mean PA and mean NA could be written like this:

$$\mu p = (2\pi)^{-1/2}(\sigma b) + \mu b\ (.5) \tag{A19}$$

and

$$\mu n = (2\pi)^{-1/2}(\sigma b) - \mu b\ (.5). \tag{A20}$$

Because $(2\pi)^{-1/2}(\sigma b)$ is a constant and adding or subtracting a constant does not effect a correlation, μp and μn in this case show a perfect negative relationship.

To summarize, we have shown that mean PA and mean NA can be derived from the mean and standard deviation of the affect distribution. Furthermore, we have demonstrated that a bipolar model of affect is consistent with correlations between mean PA and mean NA ranging from -1 to $+1$.

Diener, Larsen, Levine, and Emmons (1985)

Diener et al. (1985) proposed an analysis of affect extended over time consistent with ours, although phrased more psychologically. Frequency (of PA) and intensity were proposed as two separate, indeed uncorrelated, dimensions. In Diener et al.'s study of aggregated ratings, mean PA was found to be weakly correlated with mean NA, but the same correlation became substantially negative when intensity was controlled through partial correlation. This last result was presented as a restoration of consistency with bipolarity found for momentary affect. In this section, we analyze Diener et al.'s account in terms we have introduced already: the proportion p and the mean and variance of each individual's affect distribution. We assume that affect is measured with perfectly valid strictly unipolar scales consisting of exact antonyms.

In Diener et al.'s (1985) analysis, an individual's mean PA is the mean of the positive scale across days. Mean NA is the mean of the negative scale across days. As we have seen, with truly unipolar scales, bipolarity is consistent with any correlation between these two scores. Therefore Diener et al.'s empirical result of a correlation near zero is already consistent with bipolarity and not in need of reconciliation with results from momentary ratings; we return to this result shortly.

Diener et al. (1985) defined *frequency* (of positive affect) as the proportion of days on which the PA rating is greater than the NA rating. Thus, in our terms, frequency is the proportion p. On the assumption that each individual's distribution is normal, frequency can be defined in terms of the mean and standard deviation of the individual's affect distribution (see Equation A11). Furthermore, when the standard deviation is held constant, frequency is a monotonically increasing (although not a linear) function of the mean of that individual's affect distribution.

Diener et al. (1985) defined *positive intensity* as mean PA on those days when the positive score was greater than the negative score. *Negative intensity* is mean NA on those days when the positive score is less than the negative score. On the assumption that each individual's distribution is normal, positive intensity and negative intensity can be defined in terms of the mean and standard deviation of the individual's affect distribution (see Equations A5 and A6). So far in this discussion, we have assumed that the affect distribution is normally distributed and is therefore symmetric. In fact, a person's affect distribution need not be symmetric, and the variance to the left of the zero point need bear no relation to the variance to the right of the zero point. Therefore, it is an empirical finding, not a mathematical necessity, when Diener et al. (1985) found that positive intensity correlated .70 ($p < .01$) with negative intensity.

Diener et al. (1985) defined *overall intensity* as the mean across all days for the most intense score of each day, whether it was on the PA scale or the NA scale. They showed that when overall intensity is statistically controlled through partial correlation, the correlation between mean PA and mean NA became substantial and negative. This result can be demonstrated mathematically from the formula for the partial correlation because overall intensity is highly correlated with both mean PA and mean NA. (If we assume affect is bipolar and that truly unipolar scales are used, overall intensity is the sum of mean PA and mean NA.) The resulting shift is a mathematical necessity and therefore neither supports nor refutes an empirical thesis of bipolarity.

In short, Diener et al.'s (1985) account is phrased in more substantive terms, ours in more mathematical terms, but if we make certain assumptions, the two converge. Because every psychologist is familiar with the concepts of mean and standard deviation, we suggest that frequency and intensity can be clarified through their translation into our terms and vice versa. This translation allows a straightforward way of investigating the statistical and substantive properties of both accounts.

Received December 16, 1996
Revision received June 19, 1998
Accepted June 23, 1998 ■