Hedonic Tone, Perceived Arousal, and Item Desirability: Three Components of Self-reported Mood

Lisa Feldman Barrett

The Pennsylvania State University, USA

Self-reports of mood are the most frequently used measure of subjective emotional experience in studies of human emotion. The present study evaluated the degree to which self-reports of mood reflect the social desirability of an affective state, rather than the hedonic tone and the level of arousal associated with such states. The study produced three main findings. First, the desirability of a mood and the hedonic quality of a mood are related, but not identical entities. Secondly, the desirability of a mood is also related to the level of arousal the mood denotes. Thirdly, desirability components are related to the self-report ratings of mood, but the ratings also reflect the hedonic tone and level of arousal describing the internal state of the respondents. Social desirability does affect the self-report ratings that are often used in emotion research, but such ratings also reflect something about the internal state of the respondents.

INTRODUCTION

For better or worse, self-report represents the most reliable and possibly only window that researchers have on conscious, subjective emotional experience. Self-report measures of mood are one of the most commonly used tools in research on human emotions. Such measures are quick, easy, inexpensive to collect, and offer a more direct measure of conscious

emotional experience than behavioural, facial, and physiological measures. The latter types of measures all have uncertain and inconsistent relationships to consciously reported emotional states (see Feldman, 1993; Lang, 1994 for a discussion). In many circumstances, self-report measures are the only way to access aspects of emotional experience that occur outside the laboratory in the context of people's daily lives (e.g. Feldman, 1995a; Larsen, 1987; Oatley & Duncan, 1994; Penner, Shiffman, Paty, & Fritzsche, 1994).

Despite the advantages of using self-report measures of mood, there is legitimate concern over their validity (e.g. Carstensen & Cone, 1983; Gotlib, 1984; Gotlib & Cane, 1989; Linden, Paulhus, & Dobson, 1986; Tanaka-Matsumi & Kameoka, 1986). Self-report ratings of mood must validly assess some aspects of psychological experience, but the crucial question is: Which aspects? In particular, it is important to determine whether self-report ratings reflect a respondent's actual emotional state, or merely a defensive or self-presentational strategy. The present study assessed the degree to which self-reports of mood reflect conscious, affective experience rather than the desirability of emotional states.

A Theory of Conscious, Affective Experience

Investigations of conscious emotional experience suggest that the circumplex represents the most general structure of affective experience (Larsen & Diener, 1992; Russell, 1980; Schlosberg, 1941). The affective circumplex is defined as a circular arrangement of terms around two dimensions. Although different rotations and labellings of the circumplex dimensions exist (e.g. Diener, Larsen, Levine, & Emmons, 1985; Larsen & Diener, 1992; Thayer, 1989; Watson & Tellegen, 1985), all can be defined as combinations of the original valence/arousal dimensions (Russell, 1980; Schlosberg, 1954). The valence dimension typically refers to the hedonic quality of an affective experience (pleasant or unpleasant). The arousal dimension refers to the perception of arousal associated with an affective experience (Feldman, 1995b; Russell, 1989). The two dimensions have been identified in the semantic structure of affect terms (Feldman, 1995b; Russell, 1980), and together typically account for a substantial amount of variance in self-reports of affective experience (Russell, 1980; Watson & Tellegen, 1985). Previous research has suggested that the dimensions represent the semantic components that individuals use to interpret and communicate their conscious, affective experience (Feldman, 1995b; Russell, 1980).

Many investigators consider the valence of emotions to be the single most important dimension of affective experience (cf. Diener, 1993; Izard, 1977; Russell, 1991). Although both the valence and the arousal

dimensions are represented in the structure of affect ratings, valence plays a dominant role in self-report ratings of mood terms (Feldman, 1995b; Mayer & Gaschke, 1988; Meyer & Shack, 1989; Russell, 1978, 1980; Russell & Mehrabian, 1977; Smith & Ellsworth, 1985). Given the importance of valence in affective experience, it is necessary to demonstrate that the valence dimension of the circumplex refers to the hedonic tone of the mood (e.g. "I feel good" vs. "I feel bad"), rather than to an evaluation of the mood (e.g. "this is a good feeling to have" vs. "this is a bad feeling to have").

The Relationship between the Valence and the Social Desirability of Mood Terms

The general thesis to be explored in the present study is that the hedonic tone denoted by an affect term can be distinguished from its social desirability. In previous research, the main axis of the affective circumplex has been referred to as "evaluation" or "desirability", rather than hedonic tone or valence, and the second axis as "arousal" (Fisher, Heise, Bohrnstedt, & Lucke, 1985). Using desirability ratings from Norman (1967, as cited in Fisher et al., 1985) and likability ratings from Anderson (1968), Fisher et al. found that item desirability was highly correlated with the evaluation axis ($r = 0.94$), but weakly correlated with the arousal axis ($r = -0.21$). On the basis of this finding, Fisher et al. (1985) suggested that item desirability was the key component of the similarity among mood terms. They reasoned that the ratings of words that share similar levels of desirability would be more highly correlated than the ratings of words that are less similar.

The observed correspondence between hedonic quality and desirability may take one of three forms. A first possibility, as suggested by Fisher et al. (1985), is that the desirability of mood states, rather than hedonic tone, is the key component in the semantic structure of emotion words that influences how individuals label their mood. Social desirability or appropriateness is considered by many to be a meaningful component of emotional experience and expression. Emotions are in part social phenomena, and the meaning of mood words may primarily indicate social impact or appropriateness (Hochschild, 1979; Kemper, 1978).

Secondly, valence and desirability may be separate, but completely overlapping constructs. In Western cultures, positive moods are more desirable than negative moods (Morgan & Heise, 1988; Sommers, 1984). As pleasant emotions are evaluated positively and unpleasant emotions are evaluated negatively, participants might respond similarly to items indicating unpleasant emotions versus those reflecting pleasant emotions, simply based on the level of desirability connoted by the item. For example, an

individual may use words such as *happy* and *enthusiastic* to describe his or her experience because they are similar in desirability, rather than because they are similar in hedonic tone. In this scenario, emotion ratings might reflect the cultural norms that specify both desirable or appropriate emotions (Diener, 1993) as well as the hedonic quality of the emotion *per se*. Were this the case, a combination of desirability and hedonic quality would constitute the major dimension of affective experience.

A third possibility is that hedonic quality and social desirability may be related, but distinct conceptual phenomena that are associated with affective experience. Following this hypothesis, affective experience would be determined in part by the hedonics of internal experience and in part by the social appropriateness or communicative function of the affect being experienced.

This study directly compared the contribution of desirability and hedonic quality both to the conceptual structure of affect terms and to self-report ratings of conscious, affective experience. Affect terms were rated for their social desirability, their semantic meaning, and were used as self-report descriptors of actual affective experiences by several different but comparable samples of university students. The desirability ratings, semantic ratings, and self-report ratings were compared to test two main hypotheses.

The first hypothesis tested was that the hedonic quality denoted by the affect terms and the desirability connoted by them reflect distinct components in the conceptual structure of the words. Conceptual structure was operationalised as the similarity or proximity between the mood words (e.g. D'Andrade, 1974; Lamiell, Foss, & Cavenee, 1980; Shweder, 1975, 1977, 1980). Three types of desirability ratings of the mood terms ("likeability", "social desirability", and "personal desirability" ratings) were obtained from three samples of participants. These ratings were compared to one another and transformed to reflect desirability-based similarity in the mood words. Semantic similarity ratings of the terms were obtained from fourth group of participants. These ratings were analysed and transformed to reflect valence-based and arousal-based semantic similarity in the mood words. The first prediction was that the pattern of valence-based semantic similarity of the mood words would be distinct from the pattern of desirability-based similarity. The relationship between arousal-based semantic similarity and desirability was also explored in keeping with the analyses performed by Fisher et al. (1985), although no predictions were made.

The second hypothesis tested was that hedonic quality would be an influential conceptual component associated with the self-report ratings of mood independent of desirability. Six additional groups of participants provided self-reported ratings of their affective states. The mood ratings

were transformed to reflect the proximity of each word to every other word, and these proximity matrices were compared to the semantic- and desirability-based matrices. The second prediction was that the pattern of self-reported mood ratings would be uniquely related to both the valence-based semantic similarity and the desirability-based similarity of the mood words.

METHOD

Types of Affect Ratings

Sixteen mood terms were sampled from the affective circumplex: *aroused*, *surprised*, *peppy*, *enthusiastic*, *happy*, *satisfied*, *calm*, *relaxed*, *quiet*, *still*, *sleepy*, *sluggish*, *sad*, *disappointed*, *nervous*, *afraid*. Two mood terms represented each pole of each dimension in order to ensure that all octants of the circumplex were equally represented. Three types of data were obtained on these items to test the hypotheses of the present study. The first three samples of participants provided desirability ratings of the mood terms. The fourth sample of participants provided semantic similarity ratings of the mood terms. Six more samples of participants provided self-report ratings of the mood terms. A summary of the types of ratings and samples used is presented in Table 1.

Desirability Ratings   Participants in Sample 1 rated the *likeability* of the 16 circumplex terms by rating each term according to how much they would like a person described by the word (0 = not at all likeable; 3 = moderately likeable; 6 = extremely likeable). Participants in Sample 2 were asked to rate the *social desirability* of each term using a similar 7-point Likert scale (0 = not at all desirable; 3 = moderately desirable; 6 = extremely desirable). Participants in Sample 3 were asked to rate the extent to which they found the mood terms *personally desirable* by rating how much they themselves would like to be in this mood (0 = not at all; 3 = moderately; 6 = extremely). The three samples of desirability ratings were subjected to a weighted Euclidean multidimensional scaling (MDS) analysis.

Semantic Similarity Ratings   Participants rated the semantic similarity of all 120 possible pairs of the 16 circumplex terms. Participants were asked to rate the extent to which the words were similar in meaning using a 7-point scale (1 = extremely dissimilar; 4 = unrelated; 7 = extremely similar). Each mood term appeared as the first member in exactly half of

TABLE 1
Types of Ratings and Samples Reported in the Present Study

Type of Rating	Sample No.	Origin of Sample	Sample Size			
			Men	Women	Unknown	Total
Desirability						
Likeability	1	Waterloo	30	31	0	61
Social Desirability	2	Penn State	54	104	3	161
Personal Desirability	3	Penn State	33	78	16	127
Semantic Similarity	4	Penn State	2	22	0	24
Self-Report						
	5	Alberta	59	51	10	120
	6	Manitoba	177	135	0	312
	7	Penn State	127	134	14	275
	8	Penn State	29	104	4	137
	9	Penn State	–	–	83	83
	10	Penn State	14	19	0	35

Note: Waterloo, University of Waterloo; Penn State, The Pennsylvania State University; Alberta, University of Alberta; Manitoba, University of Manitoba.

the pairs in which it occurred, and the adjective pairs were presented in a single random order (Davison, 1983).[1]

The 120 semantic similarity ratings from Sample 4 were subjected to a weighted Euclidean multidimensional scaling (MDS) analysis. This analysis provided a representation of the semantic structure of affect (for a detailed description of these analyses see Feldman, 1995a).[2] The mood terms fell, as predicted, in a circular order around a valence and an arousal dimension. The semantic circumplex is presented in Fig. 1. A measure of valence-based semantic similarity was calculated by assessing the distance

---

[1] These semantic similarity data were obtained from subjects who participated in a study investigating the idiographic structure of affective experience (Feldman, 1995a).

[2] This scaling solution for the 16 circumplex terms was replicated using data from five male and five female undergraduate subjects sampled from the University of Waterloo (Feldman 1995b). The correlation between the valence coordinates of the two MDS solutions was 0.83; the correlation between the arousal coordinates was 0.90. Furthermore, three additional sets of circumplex items were subjected to MDS analyses and have yielded highly similar results to those presented in Feldman (1995 a,b). These latter data sets included some emotion terms that were different than the 16 items selected for the present study (e.g. cheerful, blue, angry, annoyed, excited, elated, bored, energetic, and pleased).
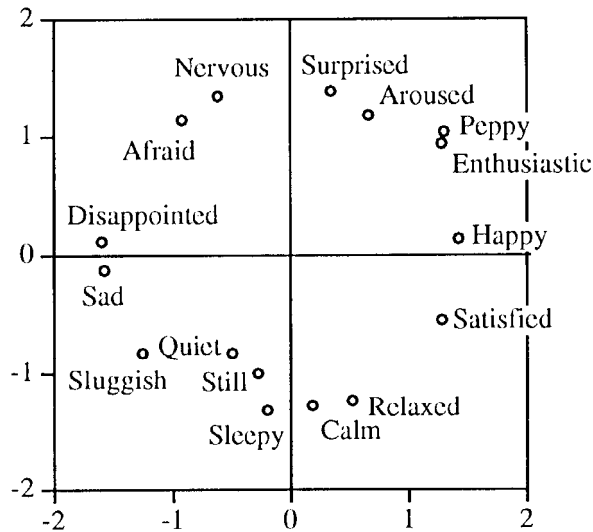
FIG. 1. The semantic-based circumplex structure of affect (Feldman, 1995a). Valence is the horizontal axis and arousal is the vertical axis.

(the absolute value of the difference) between MDS valence dimension coordinates for all 120 pairs of mood terms. Similarly, a measure of arousal-based semantic similarity was calculated by using the MDS arousal dimension coordinates. The smaller the distance between two terms on a dimension, the more similar those terms on the attribute represented by the dimension.

Self-reported Mood Ratings   Participants in Samples 5, 6, 7, 8, and 10 completed a mood questionnaire that included the 16 mood terms chosen from the circumplex, as well as many other terms found in commonly used self-report measures of mood (Feldman, 1995a).[3] Participants in Sample 9 completed a brief questionnaire that included only the 16 circumplex terms listed above. Participants indicated on a 7-point scale the extent to which each adjective described their mood (0 = not at all; 3 = a moderate amount; 6 = a great deal). Participants were asked to rate how they felt right at that moment. The samples were analysed separately for two

---

[3] The University of Alberta and University of Manitoba samples were previously analysed in Feldman (1995b) in order to demonstrate variations in the circumplex structure of affect. Specifically, people weigh the arousal dimension less than the valence dimension when making judgements of their mood, but both dimensions are weighed equally in the semantic structure of affect.

reasons. First, Samples 9 and 10 differed from the others in that these participants completed a training procedure designed to increase their perceptions of bodily arousal (Chang, Feldman, & Ravizza, 1994) whereas participants in the other samples did not. Secondly, analysing the samples separately demonstrates the replicability of the findings.

To compute proximity matrices, the self-report ratings of each circumplex adjective were correlated with the ratings of every other adjective across participants, producing 120 correlations. The similarity between different moods was indexed by these correlations. Large correlations indicated that the moods were being rated in a similar fashion, whereas smaller correlations indicated less similarity. One matrix was computed for each sample. Fisher r-to-z transformations were performed on these inter-item correlations so that they could be compared with the other similarity matrices.

RESULTS

Desirability of Mood Terms

Before proceeding to tests of the main hypotheses, the desirability ratings were examined to determine if the different types of desirability ratings were comparable. The mean desirability ratings for each mood term are presented in Table 2. Mood terms that denote positive hedonic quality (*enthusiastic*, *peppy*, *happy*, *satisfied*, *calm*, and *relaxed*) were rated as more desirable than those denoting negative hedonic tone (*nervous*, *afraid*, *sad*, *dissatisfied*, *sleepy*, and *sluggish*) in all three samples $\{t(59) = 26.84, P < 0.001; t(160) = 48.09, P < 0.001;$ and $t(126) = 45.24, P < 0.001\}$, respectively. Mood terms denoting high arousal (*nervous*, *afraid*, *aroused*, *surprised*, *enthusiastic*, and *peppy*) were rated as more desirable than those denoting low arousal states (*sleepy*, *sluggish*, *quiet*, *still*, *calm*, and *relaxed*) in all three samples $\{t(59) = 5.12, P < 0.001; t(160) = 9.21, P < 0.001;$ and $t(126) = 9.35, P < 0.001\}$, respectively.

Furthermore, the level of arousal denoted by a mood influenced the desirability level of moods with different hedonic qualities. As is evident in Table 3, positive moods denoting neutral arousal levels (*happy*, *satisfied*) were rated as more desirable than those denoting either high (*enthusiastic*, *peppy*) or low (*calm*, *relaxed*) levels of arousal. High and low arousal positive moods differed significantly in their degree of personal desirability, but not in their likeability or social desirability. Negative moods denoting high levels of arousal (*nervous*, *afraid*) were rated as more desirable than either those denoting neutral levels of arousal (*sad*, *disappointed*) or low levels of arousal (*sleepy*, *sluggish*). Negative moods

TABLE 2
Item Desirability of the 16 Mood Terms

Mood Adjective	Mean Item Desirability Score		
	Sample 1 (Likeability) (N = 60)	Sample 2 (Social Desirability) (N = 161)	Sample 3 (Personal Desirability) (N = 127)
Happy	5.37	5.55	5.48
Enthusiastic	5.00$_b$	5.47$_a$	5.04$_b$
Calm	4.73$_a$	4.52$_a$	3.93$_b$
Relaxed	4.67	4.95	4.77
Satisfied	4.62	4.86	5.04
Aroused	4.23	4.65	4.63
Peppy	4.10	4.26	4.57
Surprised	3.42	3.42	3.54
Quiet	2.85$_a$	3.14$_a$	2.30$_b$
Still	1.85$_b$	1.81$_b$	2.50$_a$
Nervous	1.27$_a$	1.46$_a$	0.92$_b$
Disappointed	1.23$_a$	1.13$_a$	0.61$_b$
Afraid	1.22	1.16	1.07
Sleepy	1.20	1.29	1.22
Sad	0.90$_a$	1.05$_a$	0.67$_b$
Sluggish	0.57$_b$	0.86$_a$	0.37$_b$

Note: Larger numbers indicate greater social desirability. The degrees of freedom for each analysis were (2,345). Mood terms appearing in **bold type** differed in their mean desirability rating across the three types of desirability ratings, $P < 0.01$, 2-tailed. Means with different subscripts within each row were statistically different, $P < 0.05$ using Newman-Keuls comparisons.

denoting neutral arousal were more likeable in a person than low arousal, negative moods.[4]

The mean ratings for 8 of the 16 circumplex terms were significantly different across the three samples (see Table 2). Two of the mood items denoted positive valence (*enthusiastic* and *calm*), two denoted neutral valence (*quiet* and *still*), and four denoted negative valence (*disappointed*, *sad*, *nervous*, and *sluggish*). Social desirability ratings were higher than personal desirability ratings in seven of the eight cases where differences occurred. There was no discernible pattern for the likeability ratings. Despite the differences noted, the correlations between mean

---

[4] One MANOVA comparing all sets of positive and negative mood items was computed for each sample of desirability ratings. All MANOVAs were statistically significant ($P < 0.01$), indicating that the significant results reported in Table 3 were not due to inflated Type I error rates.

TABLE 3
Desirability of Pleasant and Unpleasant Mood Items denoting Different
Levels of Arousal

	High Arousal	Neutral	Low Arousal
Pleasant Moods			
Likeability Ratings	4.55_a	4.99_b	4.70_a
Social Desirability Ratings	4.87_a	5.21_b	4.74_a
Personal Desirability Ratings	4.83_a	5.26_b	4.32_c
Unpleasant Moods			
Likeability Ratings	1.24_a	1.16_a	0.89_b
Social Desirability Ratings	1.31_a	1.09_b	1.08_b
Personal Desirability Ratings	1.01_a	0.63_b	$0.80_{a,b}$

Note: $N = 60$, Likeability ratings; $N = 161$, Social Desirability ratings; $N = 127$, Personal Desirability ratings. Means were compared within each row using repeated measures ANOVAS. The degrees of freedom were (1,58) for comparisons involving the Likeability ratings, (1,159) for the Social Desirability ratings, and (1,125) for the Personal Desirability ratings. Means with different subscripts within each row are statistically different, $P < 0.01$.

likeability ratings (Sample 1), mean social desirability ratings (Sample 2), and mean personal desirability ratings (Sample 3) were equal to or above $r = 0.98$. These correlations indicate that although the overall estimate of desirability may have varied for some mood terms, the mood terms displayed a very similar pattern across all three sets of ratings.

Due to the large correlations between different samples of desirability ratings, the proximity matrices of all three samples were analysed using a weighted Euclidean multidimensional scaling model.[5] A plot of the Stress values by the number of dimensions for the MDS solution, shown in Fig. 2, revealed a clear elbow at the two dimensional solution, suggesting the suitability of the two-dimensional MDS solution, Stress = 0.18.[6] Figure 2 also displays the squared correlations (RSQ) for solutions of each dimensionality; the RSQ represents the proportion of variance that the scaling solution accounts for in the distances between terms. The two-dimensional solution had an RSQ = 0.83. The Stress and RSQ statistics for each sample indicated that there was not much variation in the desirability-based structure of the mood terms across the three samples. (Stress ranged from 0.13 to 0.20; RSQ ranged from 0.78 to 0.91.)

---

[5] The primary approach to ties (allowing data to become untied) was used in the analysis because it typically results in a better fit to the data (Davison, 1983, p. 86).

[6] Although Kruskal and Wish (1978, p. 56) caution against accepting solutions with a Stress value above 0.10, the "elbow" in the plot clearly appeared for the two-dimensional solution (see Davison, 1983, p. 69). Furthermore, ease of interpretability is considered to be an acceptable criterion for selecting the dimensionality of a solution (Davison, 1983).
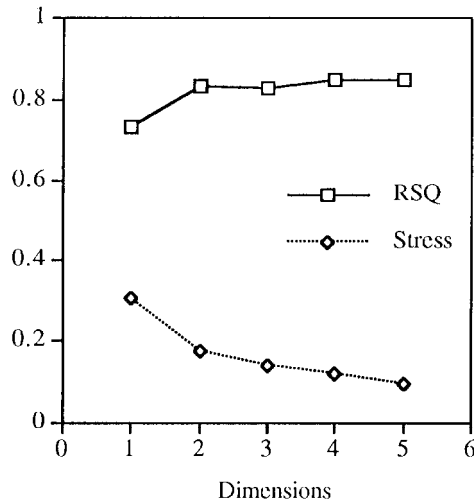
FIG. 2.   Changes in Stress and squared correlations with dimensionality for the MDS analysis of the desirability ratings.

The MDS solution, shown in Fig. 3, suggested that the desirability connoted by the affect terms was bidimensional rather than unidimensional. Furthermore, the desirability structure was somewhat circumplex-like. Inspection of the arrangement of terms along each dimension suggested that the horizontal axis represented the desirability of moods with different hedonic quality. The vertical axis corresponded to desirability of moods denoting different levels of arousal. Unlike the unidimensional desirability ratings presented in Table 2, the MDS solution demonstrates how much the two recovered dimensions contribute to the desirability of each affective state. All affect terms were dispersed along the horizontal dimension, suggesting that hedonic quality contributed to the desirability of all 16 affective states. The terms denoting pleasant moods (shown on the right side of Fig. 3) were more widely dispersed than the terms denoting unpleasant moods (shown on the left side of Fig. 3), suggesting the level of arousal denoted by positive mood states contributed the desirability of those states more so than for negative mood states.

Connections among Mood Terms: Semantic Circumplex vs. Desirability

The next set of analyses evaluated the overlap between the valence and arousal dimensions derived from the semantic ratings of the affect terms and those derived from the desirability ratings. Four similarity matrices were used to estimate the degree of relationship between the semantic-
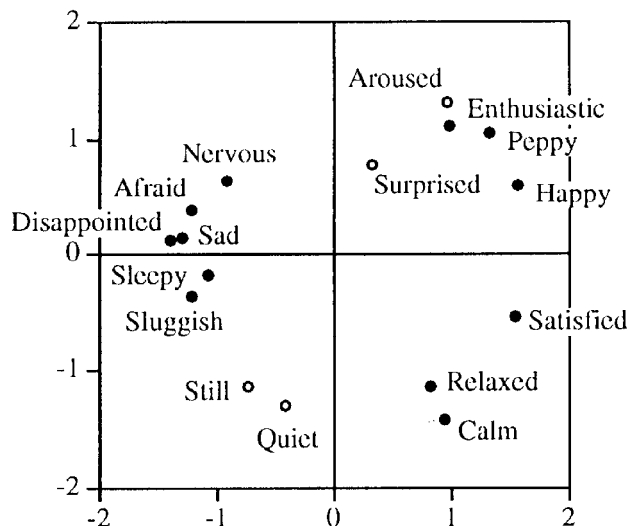
FIG. 3.   The desirability-based structure of affect. Positively and negatively valenced terms are denoted by filled circles.

based and desirability-based valence and arousal attributes of the mood words, one matrix per dimension. The valence-based semantic similarity matrix was calculated by measuring the distance between the MDS coordinates for all 120 pairs of mood words along the semantic valence dimension. The arousal-based semantic similarity matrix was calculated analogously using the semantic arousal dimension. Similarly, the valence-based and arousal-based desirability matrices were calculated from the valence-based and arousal-based desirability dimensions, respectively.

Each attribute similarity matrix was correlated with every other across the 120 pairs of mood terms to produce indices of the agreement between the semantic- and desirability-based valence and arousal attributes, presented in Table 4. There was substantial overlap between both valence-based matrices and both arousal-based matrices. The correlations in Table 4 are not low enough to permit the unequivocal assertion that the semantic and desirability components should be treated as different elements in the conceptual structure of mood adjectives.

The Relationship of Semantic Structure, Desirability, and Self-reports of Mood

The next set of analyses evaluated whether despite substantial overlap, the two semantic and the two desirability components constituted unique, influential elements in the self-report ratings of mood. To produce an

TABLE 4
Correlations between the Attribute Similarity Matrices

	1	*2*	*3*	*4*
1. Valence-based Desirable Matrix	–			
2. Arousal-based Desirability Matrix	0.03	–		
3. Valence-based Semantic Matrix	0.83*	0.05	–	
4. Arousal-based Semantic Matrix	0.12	0.70*	0.12	–

Note: The degrees of freedom for these analyses were based on the number of mood pairs ($N = 120$).

* $P < 0.001$, 2-tailed.

index of the degree to which conceptually similar words are rated in the same way, the attribute similarity data (both the semantic- and desirability-based matrices) were correlated with the inter-item correlations of the self-report mood ratings across the 120 pairs of mood terms. These analyses were computed for the self-reported mood data from Samples 5 through 10, and the correlations are presented in Table 5. The pattern of relationships between the four attribute matrices and the self-report matrices was similar for Samples 5 to 8. Both the semantic-based and the desirability-based valence similarity matrices were strongly related to the patterns of mood ratings, whereas the semantic-based and desirability-based arousal similarity matrices were related to the pattern of mood ratings to a smaller degree.

In Samples 9 and 10, the self-report matrices were more highly correlated with the arousal-based semantic matrix than in the other four samples {comparing correlations listed in second data column of Table 5: $c^2(5) = 34.15$, $P < 0.001$, 2-tailed; Meng, Rosenthal, & Rubin, 1992}. This finding likely occurred because the participants in these samples completed a training procedure designed to focus them on their physiological state. As a result, the perceived arousal component of the participants' mood ratings increased (Chang, Feldman, & Ravizza, 1994). There were also differences in the correlations between the self-report matrices and the valence-based semantic matrix reported in the first data column of Table 5 {$c^2(5) = 39.56$, $P < 0.001$, 2-tailed}. Although the size of the semantic components varied across samples, both components together accounted for a similar amount of variance in the pattern of self-reported mood in all six samples (see first data column of Table 7).

The contribution of the valence- and arousal-based desirability components varied across samples, and paralleled the variation in the valence- and arousal-based semantic components {for valence-based desirability: $c^2(5) = 80.46$, $P < 0.001$, 2-tailed; for arousal-based desirability: $c^2(5) =$

TABLE 5

Correlations between the Attribute Similarity Matrices and the Correlations
in Self-report Ratings of Mood

Self-report Ratings	Semantic Distances		Desirability-based distances	
	Valence	Arousal	Valence	Arousal
Sample 5	0.67**	0.32**	0.75**	0.31**
Sample 6	0.71**	0.21*	0.81**	0.23*
Sample 7	0.74**	0.27**	0.78**	0.36**
Sample 8	0.56**	0.36**	0.66**	0.42**
Sample 9	0.58**	0.53**	0.60**	0.55**
Sample 10	0.41**	0.64**	0.41**	0.58**

Note: $N = 120$, Sample 5; $N = 312$, Sample 6; $N = 275$, Sample 7; $N = 142$, Sample 8; $N = 83$, Sample 9; $N = 35$, Sample 10. The degrees of freedom for each analysis were based on either the number of participants involved in the calculation or the number of mood pairs, whichever was lower. This produced the most conservative estimate of degrees of freedom given that the pairs of mood terms were not independent from one another. The absolute values of the correlations are presented.
  * $P < 0.05$, 2-tailed; ** $P < 0.01$, 2-tailed.

23.42, $P < 0.005$, 2-tailed}.[7] Although the size of the desirability components varied across samples, both components together accounted for a similar amount of variance in the pattern of self-reported mood in all six samples (see first data column in Table 6).

Thus far, the analyses of zero-order correlations indicated that both semantic and desirability components were highly related to the mood ratings. The sizes of the desirability components were approximately equal to those of the semantic components in all six samples of mood ratings. The next set of analyses investigated whether or not the semantic and desirability components were significantly uniquely related to the variance in each sample of mood ratings.

The unique contribution of valence-based and arousal-based semantic similarity. Hierarchical regression analyses were conducted to test the hypothesis that semantic-based similarity was associated with unique variance in the pattern of self-report mood ratings over and above desirability-based similarity. One analysis was conducted on each sample of self-report data. In each analysis, both desirability-based similarity matrices were entered as the predictors in Step 1 of the regression, followed by the valence-based and arousal-based semantic similarity matrices in

---

[7] The amount of variance in the self-report ratings accounted for by valence-based semantic similarity is typically inversely related to that accounted for by arousal-based similarity (Feldman, 1995a, b). The two are not perfectly negatively correlated, however.

TABLE 6

Estimating the Unique Contribution of Semantic Similarity to Self-reported Affective States Controlling for Desirability-based Similarity

Mood Ratings	Step 1 (Variance due to Desirability)	Step 2 (Total Variance)	DR^2 at Step 2
	R^2	R^2	
Sample 5	0.67**	0.75**	0.08**
Sample 6	0.72**	0.76**	0.04**
Sample 7	0.75**	0.81**	0.06**
Sample 8	0.62**	0.66**	0.04*
Sample 9	0.67**	0.80**	0.13**
Sample 10	0.52**	0.70**	0.18**

Note: $N = 120$, Sample 5; $N = 312$, Sample 6; $N = 275$, Sample 7; $N = 142$, Sample 8; $N = 83$, Sample 9; $N = 35$, Sample 10. The degrees of freedom for each analysis were based on either the number of participants involved in the calculation or on the number of mood pairs, whichever was lower. This produced the most conservative estimate of degrees of freedom given that the pairs of mood terms were not independent from one another.

* $P < 0.01$; ** $P < 0.0001$, 2-tailed.

Step 2. The squared multiple correlation for Step 1 represents the predictive contribution of item desirability. The incremental change in R^2 from Step 1 to Step 2 represents an estimate of the unique contribution of the semantic similarity of the moods words in predicting the inter-item correlations of the self-reported mood states. The results are presented in Table 6. As indicated in the third data column of Table 6, valence-based and arousal-based semantic similarity contained predictive power beyond what was attributable to the desirability of the mood terms. This finding was replicated across all six samples of self-report ratings of mood (changes in R^2 ranged from 0.04 to 0.18). Thus, the pattern of mood ratings was similar to the pattern of semantic similarity after controlling for the influence of desirability, suggesting that self-reports of mood do describe something about the internal states of the respondents.

The unique contribution of valence-based and arousal-based desirability. Further analyses were conducted to investigate whether desirability-based similarity had a unique association with the pattern of self-report mood ratings. As in the previous analyses, a series of hierarchical multiple regressions were conducted, one for each sample of self-report data. In these analyses, the valence-based and arousal-based semantic similarity matrices were entered as the predictors in Step 1 of the regression, followed by both desirability-based similarity matrices in Step 2. The incremental change in R^2 from Step 1 to Step 2 represents an estimate of

TABLE 7

Estimating the Unique Contribution of Desirability-based Similarity
to Self-reported Affective States Controlling for Semantic Similarity

Mood Ratings	Step 1 (Variance due to Semantic Similarity)	Step 2 (Total Variance)	DR^2 at Step 2
	R^2	R^2	
Sample 5	0.62**	0.75**	0.13**
Sample 6	0.60**	0.76**	0.16**
Sample 7	0.69**	0.81**	0.12**
Sample 8	0.50**	0.66**	0.16**
Sample 9	0.71**	0.80**	0.09**
Sample 10	0.66**	0.70**	0.04*

Note: $N = 120$, Sample 5; $N = 312$, Sample 6; $N = 275$, Sample 7; $N = 142$, Sample 8; $N = 83$, Sample 9; $N = 35$, Sample 10. The degrees of freedom for each analysis were based on either the number of participants involved in the calculation or on the number of mood pairs, whichever was lower. This produced the most conservative estimate of degrees of freedom given that the pairs of mood terms were not independent from one another.

* $P < 0.001$; ** $P < 0.0001$, 2-tailed.

the unique contribution of desirability-based similarity of the mood words in predicting the inter-item correlations of the self-reported mood states (see Table 7). Valence- and arousal-based desirability contained predictive power beyond what was attributable to the semantic meaning of the affect terms. Again, this finding was replicated across all six samples of ratings (changes in R^2 ranged from 0.04 to 0.16).

DISCUSSION

Summary of Results

The results of the present study demonstrate that the desirability of an affective state is empirically related to both its hedonic quality and the level of arousal denoted by that state. Semantic and desirability judgements produced related patterns of associations among the mood words. Both semantic-based and desirability-based conceptual similarity were related to the way in which participants rated their subjective emotional states, but each set of dimensions or attributes made a unique contribution.

The results of this study suggest several things. First, the hedonic quality of an affective state and the desirability of that state are not identical entities. This finding challenges previous concerns that hedonic quality is not a veridical component of mood ratings (Fisher et al., 1985), and calls

into question the claim that self-reported emotion or related reports are primarily determined by social desirability concerns (Carstensen & Cone, 1983). The results of this study demonstrate that individuals' reports of their affective experience reflect the hedonic quality of the state, as well as their perception of arousal, in addition to the desirability associated with the valence and arousal denoted by each state.

In addition to demonstrating that desirability does not mask important dimensions of affective experience, the results of this study suggest that desirability is an important component in the conceptual connections among mood words. Social desirability has a unique, substantial relationship to the ratings of affective experience. Although the findings clarify the presence of desirability, they do not address the meaning of the desirability component in the self-report patterns of affective states. The degree to which desirability is a proper part of affective experience is a topic of considerable debate. Theorists who believe in a more biological basis for affective experience (e.g. Ekman, 1971; Ekman, Levenson, & Friesen, 1983; Plutchik, 1962) would consider the desirability components of mood ratings to be a source of response bias. Theorists who believe in a more social basis for affective experience (e.g. Fridlund, 1992; Goffman, 1974; Hochschild, 1979; Kemper, 1978; McCrae, 1986; Rorer, 1990; Sarbin, 1989) would consider the desirability components of mood ratings to be substantive properties of the experience. The data from the present study do not speak to the validity of either theoretical stance. They do, however, clearly indicate that the social desirability of affective states is communicated in self-reports of mood.

Potential Limitations of the Present Study

The interpretations of the findings are qualified somewhat by the potential limitations of the present study. First, it might be argued that a restricted range of affective experiences was sampled in this study. Several emotion terms that are central to some emotion theories (e.g. terms to represent anger, guilt, or shame) were not included. The goal of this study, however, was to investigate *patterns* of affective responding, rather than the experience of any given emotion *per se*. The inclusion of only 16 words is not a serious shortcoming of the present study unless there is evidence to suggest that desirability will differentially affect the experience or expression of a particular affective state when compared to others. No such evidence exists to date.

Secondly, the differences in the semantic-based and desirability-based matrices may reflect random fluctuation. One substantially different dimension coordinate in the two solutions will reduce the correlation between the resulting similarity matrices. As a consequence, these data may under-

estimate the degree of overlap in the semantic-based connections and desirability-based connections among the mood words. If differences in dimension coordinates are due to random error, then the differences found in the similarity matrices may be meaningless. Were this the case, the findings of this study would not rule out the possibility that the semantic similarity and desirability-based similarity are identical entities at a conceptual level. Steps were taken to avoid this possibility by obtaining several samples of data. The desirability structure presented in Fig. 3 is very similar to the semantic structure presented in Fig. 1, however. In addition, the correlations between the semantic-based and the desirability-based matrices were high. Both sets of ratings yielded corresponding, although not identical, dimensions. If the meaning of a mood word includes social impact or appropriateness (Hochschild, 1979; Kemper, 1978), then the correspondence may indicate that both semantic meaning and desirability are components of the basic descriptive dimension in the structure of affective experience. Experimental evidence is needed to confirm the hypothesis that the two are separate. For example, this study involved mood ratings that were obtained under relatively mundane classroom conditions. Perhaps the relative sizes of the semantic and desirability components in the samples of self-reported mood may differ when assessing emotional states that are obtained under evocative (perhaps directly interpersonal) conditions.

Thirdly, it might be argued that the global ratings obtained in this study provided a limited assessment of the desirability of an affective state. The desirability of an affective state was measured using three criteria: (1) how much a person would be liked if he or she possessed each emotion; (2) how socially desirable it would be to experience each emotion; and (3) how personally desirable it would be to experience each emotion. The three types of desirability ratings were highly similar, suggesting that, for example, self-imposed feeling rules (personally desirable) are closely related to socially imposed feeling rules (social desirability). This interpretation is consistent with the notion that display rules become related to emotional experience through the socialisation of emotion expression (e.g. Barrett & Campos, 1987; Izard & Malatesta, 1987). Some differences may exist, however, in the personal desirability and social desirability of affective states. Almost half of the mood terms were rated as more socially desirable than personally desirable.[8] This finding has some interesting implications. Perhaps individuals judge others less harshly for expressing a negative mood than they judge themselves. Some indirect evidence for this hypothesis comes from a study conducted by Funder and Sneed (1993).

---

[8] I thank an anonymous reviewer for this observation.

Although the expression of negative emotion was considered by partici-pants to be diagnostic of neuroticism, acquaintances did not use this as an indicator when judging the behaviour of their target friends. Thus, acquain-tances did not use the expression of negative emotion as a criterion for making negative judgements about their friends.

Moreover, participants in this study made desirability ratings without reference to the social context. The desirability of an emotional state is in part defined by when and where an emotion is experienced or expressed. For example, it is expected (some might even say desirable) to become sad at a funeral, angry in the face of a racist remark, or tired when one has stayed up all night studying for an exam. Asking participants to rate the desirability of an emotional state in a contextless format amounts to asking them to make the rating, all other things being equal. We experience emotions in a social context that changes from moment to moment, however; so all other things are rarely equal. Not only do the prescriptions change for different social contexts, but an individual's goals may also change from situation to situation. One affective state might be more desirable than another in a given situation because it may afford a host of cognitive and motivational aspects that are congruent with the goals of the individual (Parrott, 1993). Additional research is needed to determine if the desirability of an emotion is consistent across different social circum-stances, with certain interaction partners, or when individuals have differ-ing goals.

Finally, the findings of this study may be culture-bound. Additional research is needed to determine if the desirability of an affective state is consistent across cultures. Other cultures, or even subcultures, have social rules to govern the experience and expression of emotional experience that may differ from those of the dominant North American culture from which the participants of the present study were sampled. To the extent that there are cultural differences, the relationship between the descriptive and the prescriptive aspects of affective experience may vary.

CONCLUSIONS

The results of this study clarify several important issues about affective experience. First, self-reports of mood reflect the hedonic tone, the level of arousal, and the desirability associated with emotional states. Secondly, the social desirability of an affective state is based not just on hedonic quality, but also on the level of arousal experienced. Thirdly, desirability concerns do affect the self-report ratings that are often used in psychological research, but such ratings also reflect some aspects of the respondents' internal state. These findings do not rule out the possibility that some individuals may be responding to self-report mood measures in a biased

manner, but such responding does not seem to account consistently for all of the variance in reports of affective experience.

REFERENCES

Anderson, N.H. (1968). Likeableness ratings of 555 personality-trait words. *Journal of Personality and Social Psychology*, *9*, 272–279.

Barrett, K., & Campos, J. (1987). Perspectives on emotional development II: A functionalist approach to emotions. In J. Osofsky (Eds), *Handbook of infant development* (2nd ed., pp. 555–578). New York: Wiley.

Carstensen, L.L., & Cone, J.D. (1983). Social desirability and measurement of psychological well-being in elderly persons. *Journal of Gerontology*, *38*, 713–715.

Chang, S.K., Feldman, L.A., & Ravizza, R. (1994, August). *The effect of breathing on attention to arousal in self reports of emotion.* Paper presented at the annual meeting of the American Psychological Association, Los Angeles.

D'Andrade, R.G. (1974). Memory and the assessment of behavior. In T. Blalock (Ed.), *Measurement in the social sciences* (pp. 159–186). Chicago: Aldine-Atherton.

Davison, M.L. (1983). *Multidimensional scaling.* Toronto: Wiley.

Diener, E. (1993). Challenges in measuring subjective well-being and ill-being. *Social Indicators Research*, *28*, 1–20.

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.

Ekman, P. (1971). Universals and cultural differences in facial expressions of emotion. In J. Cole (Ed.), *Nebraska Symposium on Motivation* (Vol. 19, pp. 207–282). Lincoln: University of Nebraska Press.

Ekman, P., Levenson, R.W., and Friesen, W.V. (1983). Autonomic nervous system activity distinguishes among emotions. *Science*, *221*, 1208–1210.

Feldman, L.A. (1993). Distinguishing depression from anxiety in self-report: Evidence from confirmatory factor analysis on nonclinical and clinical samples. *Journal of Consulting and Clinical Psychology*, *61*, 631–638.

Feldman, L.A. (1995a). Valence-focus and arousal-focus: Individual differences in the structure of affective experience. *Journal of Personality and Social Psychology*, *69*, 153–166.

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

Fisher, G.A., Heise, D.R., Bohrnstedt, G.W., & Lucke, J.F. (1985). Evidence for extending the circumplex model of personality trait language to self-reported moods. *Journal of Personality and Social Psychology*, *49*, 233–242.

Fridlund, A.J. (1992). The behavioral ecology and sociality of human faces. In M.S. Clark (Ed.), *Review of Personality and Social Psychology*, *13*, 90–121.

Funder, D.C., & Sneed, C.D. (1993). Behavioral manifestations of personality: An ecological approach to judgmental accuracy. *Journal of Personality and Social Psychology*, *64*, 479–490.

Gotlib, I.H. (1984). Depression and general psychopathology in university students. *Journal of Abnormal Psychology*, *93*, 19–30.

Gotlib, I.H., & Cane, D.B. (1989). Self-report assessment of depression and anxiety. In P.C. Kendall & D. Watson (Eds), *Anxiety and depression: Distinctive and overlapping features* (pp. 131–169). Toronto: Academic.

Goffman, E. (1974). *Frame analysis.* New York: Harper & Row.

Hochschild, A.R. (1979). Emotion work, feeling rules, and social structure. *American Journal of Sociology*, *85*, 551–575.

Izard, C.E. (1977). *Human emotions.* New York: Plenum.

Izard, C.E., & Malatesta, C.Z. (1987). Perspectives on emotional development I: Differential emotions theory of early emotional development. In J. Osofsky (Ed.), *Handbook of infant development* (Vol. 2, pp. 494–554). New York: Wiley.

Kemper, T.D. (1978). Toward a sociological theory of emotions: Some problems and some solutions. *American Sociologist*, *13*, 30–41.

Kruskal, J.B., & Wish, M. (1978). *Multidimensional scaling.* Beverly Hills: Sage.

Lamiell, J.T., Foss, M.A., & Cavenee, P. (1980). On the relationship between conceptual schemes and behavior reports: A closer look. *Journal of Personality*, *48*, 54–73.

Lang, P.J. (1994). The varieties of emotional experience: A meditation of James-Lange theory. *Psychological Review*, *101*, 211–221.

Larsen, R.J. (1987). The stability of mood variability: A spectral analytic approach to daily mood assessment. *Journal of Personality and Social Psychology*, *52*, 1195–1204.

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

Linden, W., Paulhus, D.L., & Dobson, K.S. (1986). Effects of response styles on report of psychological and somatic distress. *Journal of Counseling and Clinical Psychology*, *54*, 309–313.

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

McCrae, R.R. (1986). Well-being scales do not measure social desirability. *Journal of Gerontology*, *41*, 390–392.

Meng, X., Rosenthal, R., & Rubin, D.B. (1992). Comparing correlated correlation coefficients. *Psychological Bulletin*, *111*, 172–175.

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.

Morgan, R.L., & Heise, D. (1988). Structure of emotions. *Social Psychology Quarterly*, *51*, 19–31.

Oatley, K., & Duncan, E. (1994). The experience of emotions in everyday life. *Cognition and Emotion*, *8*, 369–381.

Parrott, W.G. (1993). Beyond hedonism: Motives for inhibiting good moods and for maintaining bad moods. In D.M. Wegner & J.W. Pennebaker (Eds), *Handbook of mental control* (pp. 278–305). Englewood Cliffs, NJ: Prentice Hall.

Penner, L.P., Shiffman, S., Paty, J.A., & Fritzsche, B.A. (1994). Individual differences in intraperson variability in mood. *Journal of Personality and Social Psychology*, *66*, 712–721.

Plutchik, R. (1962). *The emotions: Facts, theories, and a new model.* New York: Random House.

Rorer, L.G. (1990). Personality assessment: A conceptual survey. In L.A. Pervin (Ed.), *Handbook of personality: Theory and research* (pp. 693–720). New York: Guilford.

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

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

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

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

Sarbin, T.R. (1989). Emotions as narrative emplotments. In M.J. Packer & R.B. Addison (Eds), *Entering the circle: Hermeneutic investigators in psychology* (pp. 185–201). Albany: SUNY Press.

Schlosberg, H. (1941). A scale for the judgment of facial expressions. *Journal of Experimental Psychology, 29,* 497–510.

Schlosberg, H. (1954). Three dimensions of emotion. *Psychological Review, 61,* 81–88.

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

Shweder, R.A. (1975). How relevant is an individual difference theory of personality? *Journal of Personality, 43,* 455–484.

Shweder, R.A. (1977). Likeness and likelihood in everyday thought: Magical thinking in judgments about personality. *Current Anthropology, 18,* 637–648.

Shweder, R.A. (1980). Factors and fictions in person perception: A reply to Lamiell, Foss, and Cavenee. *Journal of Personality, 48,* 74–81.

Sommers, S. (1984). Reported emotions and conventions of emotionality among college students. *Journal of Personality and Social Psychology, 46,* 207–215.

Tanaka-Matsumi, J., & Kameoka, V.A. (1986). Reliabilities and concurrent validities of popular self-report measures of depression, anxiety, and social desirability. *Journal of Consulting and Clinical Psychology, 54,* 328–333.

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

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