A Circumplex Model of Affect

James A. Russell
University of British Columbia, Vancouver, Canada

Factor-analytic evidence has led most psychologists to describe affect as a set of dimensions, such as displeasure, distress, depression, excitement, and so on, with each dimension varying independently of the others. However, there is other evidence that rather than being independent, these affective dimensions are interrelated in a highly systematic fashion. The evidence suggests that these interrelationships can be represented by a spatial model in which affective concepts fall in a circle in the following order: pleasure (0°), excitement (45°), arousal (90°), distress (135°), displeasure (180°), depression (225°), sleepiness (270°), and relaxation (315°). This model was offered both as a way psychologists can represent the structure of affective experience, as assessed through self-report, and as a representation of the cognitive structure that laymen utilize in conceptuallyizing affect. Supportive evidence was obtained by scaling 28 emotion-denoting adjectives in four different ways: Ross' technique for a circular ordering of variables, a multidimensional scaling procedure based on perceived similarity among the terms, a unidimensional scaling on hypothesized pleasure–displeasure and degree-of-arousal dimensions, and a principal-components analysis of 343 subjects' self-reports of their current affective states.

Beginning with Nowlis (e.g., Nowlis & Nowlis, 1956), investigators who have factor analyzed self-reported affective states have typically concluded that there are between six and twelve independent monopolar factors of affect, such as degree of sadness, anxiety, anger, elation, tension, and the like (e.g., Borgatta, 1961; Clyde, 1963; Curran & Cattell, 1975; Hendrick & Lilly, 1970; Izard, 1972; Lorr, Daston, & Smith, 1967; McLaughlin, 1976; McNair & Lorr, 1964; Nowlis, 1965; Rymen, Biersner, & La Rocco, 1974). Although this result was somewhat unexpected—one might have thought that sadness would be inversely related to elation, for example—most psychologists today have accepted the conclusion that each such affective concept can be treated as a separate dimension. This assumption is included in Tomkins' (1962–1963) and Izard's (1972) theory of discrete emotions, it is included in Ekman's (e.g., 1972) cross-cultural work on facial expressions of emotion, and it is the basis for the self-report instruments most commonly used today in clinical, personality, and social psychology to assess affect (i.e., affect scales developed by Izard, 1972; McNair, Lorr, & Droppleman, 1971; Nowlis, 1965; and Thayer, 1967).

The notion has also persisted, however, that affective states are not independent of one another, but are related to each other in a highly systematic fashion. This view is illustrated by Schlosberg's (1952) proposal that emotions are organized in a circular arrangement, which means that they are adequately represented as only two bipolar dimensions, rather than as six to twelve monopolar ones. My thesis is that affective

I am very grateful to Karen Chapnick, Nicole Clement, Jasmin Jiwani, Doreen Ridgeway, and Klaus and Marsha Schroeder for their help in carrying out this study, and to Lisa Gaelick, Daniel Kahneman, Auke Tellegen, Larry Ward, and Jerry Wiggins for their thoughtful comments on an earlier version of this article. This study was supported by a grant from the Social Sciences and Humanities Research Council of Canada.

Requests for reprints should be sent to James A. Russell, Department of Psychology, University of British Columbia, Vancouver, British Columbia, Canada V6T 1W5.

Copyright 1980 by the American Psychological Association, Inc. 0022-3514/80/06-1161$00.75

1161
states are, in fact, best represented as a circle in a two-dimensional bipolar space. Two types of evidence will be considered: (a) evidence on how laymen conceptualize affective states and (b) evidence from multivariate analyses of self-reported affective states.

Laymen's Conceptualization of Affect

Psychologists are not the only ones who categorize affective states. In their daily interactions with others, most people interpret the moods of others, anticipate each other's emotional response, and seek to modify those emotional responses. In doing so, laymen must rely on their own knowledge about emotion. Laymen have learned about emotions, organizing and summarizing their knowledge into a cognitive structure. In turn, the cognitive structure helps to shape the perception and interpretation of specific events. It would be used in interpreting verbal descriptions of emotion, including anything from a subtle hint to an explicit declaration. It would be used in interpreting nonverbal evidence of emotional states, including facial expressions, tone of voice, slips of the tongue, overt actions, blushing, or any of a host of other possible cues. And to anticipate my thesis to be presented later, it would be used in conceptualizing and reporting one's own emotional state.

People might be said to have an implicit theory of emotions, suggesting an analogy to implicit personality theory or to a scientific theory. Indeed, the laymen's cognitive representation of emotion is presumably implicit in the sense that few if any could explicitly state their complete conceptual framework; it must be inferred from judgments made about emotion. It may also be useful to think of the cognitive representation as like a theory in that it includes a set of categories (each labeled by a term such as happy, excited, or sad) interrelated in a systematic fashion (an implicit taxonomy of emotions). But whether the cognitive representation for affect resembles a scientific theory in other ways remains to be seen. To take a more neutral stance on the topic, I shall simply say that persons possess a cognitive structure capable of representing affect.

Categorization of Nonverbal Emotional Expressions

The idea that people can be viewed as naïve psychologists has played an important role in personality and social psychology through the influence of Kelly (1955) and Heider (1958). Nevertheless, there have been few studies concerned with the laymen's theories of affect. These would be studies in which subjects must rely on their knowledge of affect rather than on introspection of their current affective state. In one such study, Schlosberg (1952) examined the errors that subjects made when categorizing facial expressions of emotions into experimenter-determined categories (such as surprise, disgust, and contempt). From these errors, he derived a circular representation of emotions involving dimensions of pleasantness-unpleasantness and attention-rejection. Schlosberg (1954) later suggested a third dimension, sleep-tension.

Abelson and Sermat (1962) had subjects rate the similarity-dissimilarity of pairs of facial expressions, thereby eliminating the need for any emotion categories imposed by the experimenter. The similarity data were analyzed by a multidimensional scaling procedure in which greater similarity between two expressions was represented by their closeness in a geometric space. The result was a two-dimensional space, the axes of which were pleasantness-unpleasantness and a combination of the sleep-tension and attention-rejection axes, which were essentially undifferentiable. Royal and Hays (1959), Shepard (1962), and Cliff and Young (1968) also multidimensionally scaled the judged similarity among facial expressions, and all found very similar structures, easily interpretable in terms of the same two dimensions. Green and Cliff (1975) investigated the way subjects categorized vocally expressed emotions. Again, two independent dimensions emerged, the first being pleasantness-unpleasantness and the second being some combination of Schlosberg's attention-rejection
and sleep—tension (which were again so highly correlated, \( r = .85 \), as to be undifferentiable).

Categorization of Verbal Emotional Expressions

Studies of the affective structure implicit in the English language have also supported Schlosberg’s (1952) hypothesis. Semantic differential studies indicate that dimensions of evaluation, activity, and potency are major components of the meaning of natural languages (Osgood, May, & Miron, 1975), and these dimensions have been interpreted as affective in nature (Osgood, 1969). Averill (1975) specifically studied semantic differential ratings of emotion terms and supported the evaluation (pleasantness—unpleasantness) and activity (arousal or activation) dimensions, but found two dimensions in place of potency: control and depth or importance. Russell and Mehrabian (1977) similarly found reliable and meaningful ratings of emotion terms on pleasure—displeasure and degree-of-arousal dimensions, plus a third dimension, dominance—submissiveness, related to potency.

More direct evidence on the structure of the language of affect comes from multidimensional scaling studies of affect terms. This evidence suggests three specific properties of the cognitive representation for affect. First, the pleasantness—unpleasantness and arousal—sleep dimensions account for the major proportion of variance in the judged similarities among emotion words (Bush, 1973; Dittmann, 1972; Neufeld, 1975, 1976; Russell, 1978). Additional dimensions are often obtained, but each accounts for a quite small proportion of variance, and there is little consensus on their interpretation. Russell (1978) specifically explored dimensions beyond pleasure and arousal and obtained dimensions interpreted as referring to (beliefs about) the antecedents or consequences of the emotion described rather than as referring to the emotion per se.

Second, the dimensions descriptive of affect are bipolar. Block (1957), Bush (1973), and Neufeld (1975, 1976) obtained semantic structures for affect that were interpreted as bipolar. Lundberg and Devine (1975) and Stone and Coles (1970) specifically examined the polarity of affect terms and obtained strong evidence of bipolarity. Complementary, although less direct, evidence can be seen in Bentler’s (1969) demonstration that the semantic differential factors are also bipolar.

If there are only two dimensions in the cognitive representation of emotions, one might wonder about the ability of such a representation to define the myriad of affect terms such as anger, anxiety, depression, elation, and the rest that are not synonymous with either pleasure—displeasure or degree of arousal. Russell and Pratt (1980) examined this question in the context of the affective qualities attributed to places (relaxing places, gloomy places, etc.) and found that, indeed, many affect terms were not synonymous with (did not cluster about) the pleasantness or arousal axes. Instead, terms fell meaningfully around the perimeter of the space defined by the axes. In other words, the third property of the language of affect is that any affect word could be defined as some combination of the pleasure and arousal components. This of course is Schlosberg’s (1952) idea of a circular order within a two-dimensional space. Fillenbaum and Rapoport (1971) reanalyzed data gathered by Block (1957) and obtained a similar result: Emotion names were arranged on the circumference of a circle in a two-dimensional space.

Cognitive Structure of Affect

These three properties of the cognitive representation of affect are summarized in Figure 1, where eight variables fall on a circle in a two-dimensional space in a manner analogous to points on compass. The horizontal (east—west) dimension in this spatial metaphor is the pleasure—displeasure dimension, and the vertical (north—south) dimension is arousal—sleep. The remaining four variables do not form independent dimensions, but help to define the quadrants of the space. Excitement need not be defined as an approximate north or as an approximate east; it can be defined precisely as falling at a point
in the northeast, the combination of high pleasure and high arousal. Excitement's bipolar opposite, depression, is then defined as a point in the southwest. Distress and its opposite, contentment, similarly form a bipolar dimension (northwest vs. southeast). All other affect terms would be similarly defined as vectors originating from the origin of the space.

In a first series of three studies on the layman's mental map of affective space, the hypotheses embodied in Figure 1 were tested. A sample of 28 words was chosen to represent the domain of affect. These 28 words were then scaled by techniques in which subjects would rely, not on their current affective states, but on their internal representations of affect. The first was based on Ross' (1938) technique specifically developed for variables falling in a circular ordering. (Kavanagh & McCormick, in press, have recently revived and discussed Ross' technique.) The same 28 words were also multidimensionally scaled to obtain scaling coordinates with a measurement model that does not assume a circular ordering. Finally, the 28 words were unidimensionally scaled along the proposed pleasure–displeasure and degree-of-arousal dimensions. These three scaling solutions were then quantitatively compared to assess the degree of convergence of the results across methods.

Circular Scaling of Terms

Method

Subjects. Subjects were 36 University of British Columbia undergraduates of both sexes who volunteered to participate in research.

Category-sort task. Each subject first performed a category-sort task. The 28 stimulus words (given in Figure 2) were described as "words or phrases that people use to describe their moods, feelings, temporary states, affect, or emotions." The subject was instructed to place each word, presented in alphabetical order, into one of eight categories (representing the concepts shown in Figure 1) labeled arousal, contentment, depression, distress, excitement, misery, pleasure, and sleepiness. There were no restrictions concerning how many, if any, of the 28 words were to be placed into any one category.

Circular ordering task. A second task was devised to complement the category-sort task utilized by Ross (1938). When the subject had completed the category-sort task, he or she was asked to place the eight categories—labeled aroused, contented, depressed, distressed, excited, miserable, pleased, and sleepy—into a circular order. The instructions stated: "Your task is to place the words around the edge of a circle in such a way that (1) words opposite each other on the circle describe opposite feelings, and (2) words closer together on the circle describe feelings that are more similar." Subjects were allowed to test various alternatives before giving their final solution to the experimenter.

Results

Circular ordering of eight categories. Ross' (1938) scaling technique requires that the order of the eight categories first be determined. Thus the data from the second task, involving the placement of the eight categories in a circular order, were analyzed first. Table 1 gives the frequency with which subjects placed the emotion categories in the eight slots around the circle. Since subjects' rotations of the circle were of no interest, the slot in which a subject placed aroused was assigned Number 1, with other slots assigned Numbers 2 through 8 in the order the subject placed them. If the data perfectly matched that predicted in Figure 1, all entries would fall along the diagonal of Table 1. As can be seen, the modal responses did fall along the diagonal, with the frequency of other responses rapidly falling off the further from the diagonal one goes. Indeed, 10 of the 36 subjects produced exactly the circular ordering
predicted. (Note that it was the ordering that was as predicted; reflection or rotation of the circle was ignored.)

For each subject, a distance matrix was created by assigning a score of 1 to the distance between terms placed adjacent on the circle, 2 to the distance between terms once removed, and so on until 4, which was assigned to the distance between terms placed opposite on the circle. Each subject's resulting matrix, consisting of 28 entries, was then correlated with the theoretical matrix predicted from Figure 1. The resulting 36 correlations ranged from .19 to 1.00, with a median of .80. (Ten of the 36 correlations were 1.00, which resulted from the 10 subjects whose circular orderings perfectly matched Figure 1.)

A final test of the resemblance of the obtained data to the theoretical structure was obtained by creating an average similarity matrix with means across the 36 subjects. This matrix was then analyzed by the Guttman-Lingoes (Lingoes, 1965, 1973) nonmetric multidimensional scaling procedure SSA-1. A two-dimensional solution resulted in a remarkable stress value of .001, with the eight categories ordered exactly as predicted in Figure 1. Indeed, the points plotted in Figure 1 are the empirical scale coordinates from the SSA-1 solution for the eight categories.

**Category sort task.** Table 2 gives the frequency with which each of the 28 stimulus words was placed in the eight emotion categories. The eight categories are ordered in the table as predicted (and supported) in Figure 1. Nevertheless, the data given in Table 2 could also be used to estimate the order of the eight categories, since Ross' procedure for discovering their ordering is to rearrange them by trial and error until the frequency data maximally resemble the pattern expected for a circular order. The pattern expected for a circular order is just the one approximated in Table 2; each row should show a symmetric distribution, with a modal frequency in one column, the next highest frequency in the two columns on either side, and so on, with the understanding that the end columns are also adjacent. As can be seen in Table 2, no trial and error procedure was needed since this pattern was already evident. There were deviations from this pattern in only 8 of the 28 cases, and these were for the most part trivial. (The most deviant case can be seen in the word sad.)

**Fuzziness of affect terms.** Table 2 also suggests another interesting property of the 28 emotion-denoting words. Each word appears to lack sharp boundaries in the sense that it was possible to place it in a number of categories, with the probability associated with placement in each category varying systematically with the position of the category in the circular order. Each emotion word can thus be considered a label for a fuzzy set, defined as a class without sharp boundaries, in which there is a gradual but specifiable transition from membership to nonmembership. Such a result for emotion-denoting words is consistent with the view that fuzziness is a characteristic of natural language categories in general (Hersh & Caramazza, 1976; Labov, 1973; Lakoff, 1973). It may therefore be fruitful to view the emotion-labeling process as a mapping function. Each internal emotional state would have associated with it a value (between 0 and 1) specifying its grade of membership in each fuzzy set that corresponds to

---

**Table 1**

<table>
<thead>
<tr>
<th>Term</th>
<th>Position on circle</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Aroused</td>
<td>36</td>
</tr>
<tr>
<td>Excited</td>
<td>24</td>
</tr>
<tr>
<td>Pleased</td>
<td>9</td>
</tr>
<tr>
<td>Contented</td>
<td>2</td>
</tr>
<tr>
<td>Sleepy</td>
<td>9</td>
</tr>
<tr>
<td>Depressed</td>
<td>1</td>
</tr>
<tr>
<td>Miserable</td>
<td>1</td>
</tr>
<tr>
<td>Distressed</td>
<td>2</td>
</tr>
</tbody>
</table>

*Note. For each subject (N = 36), position No. 1 was defined as wherever he or she placed the term *aroused*, hence the 36 entries for *aroused* at No. 1. Positions were then numbered consecutively for each subject. Diagonal elements are printed in italics.*
Table 2
Frequency of Placement of 28 Words Into Eight Categories

<table>
<thead>
<tr>
<th>Term</th>
<th>Pleasure</th>
<th>Excitement</th>
<th>Arousal</th>
<th>Distress</th>
<th>Misery</th>
<th>Depression</th>
<th>Sleepiness</th>
<th>Contentment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Happy</td>
<td>21</td>
<td>8</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5</td>
</tr>
<tr>
<td>Delighted</td>
<td>15</td>
<td>16</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>Excited</td>
<td>2</td>
<td>29</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Astonished</td>
<td>17</td>
<td>18</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aroused</td>
<td>14</td>
<td>21</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tense</td>
<td>8</td>
<td>18</td>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alarmed</td>
<td>6</td>
<td>19</td>
<td>11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Angry</td>
<td>5</td>
<td>21</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Afraid</td>
<td>2</td>
<td>11</td>
<td>22</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Annoyed</td>
<td>1</td>
<td>12</td>
<td>14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Distressed</td>
<td>4</td>
<td>25</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frustrated</td>
<td>2</td>
<td>5</td>
<td>19</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Miserable</td>
<td>3</td>
<td>23</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sad</td>
<td>10</td>
<td>6</td>
<td>19</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gloomy</td>
<td>2</td>
<td>11</td>
<td>22</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Depressed</td>
<td>4</td>
<td>7</td>
<td>24</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bored</td>
<td>3</td>
<td>2</td>
<td>14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Droopy</td>
<td>1</td>
<td>1</td>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tired</td>
<td>1</td>
<td>1</td>
<td>34</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sleepy</td>
<td>1</td>
<td>32</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Calm</td>
<td>4</td>
<td></td>
<td>32</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Relaxed</td>
<td>6</td>
<td></td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Satisfied</td>
<td>3</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>At ease</td>
<td>7</td>
<td></td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Content</td>
<td>6</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Serene</td>
<td>8</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glad</td>
<td>20</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pleased</td>
<td>22</td>
<td>2</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

an emotion label, such as happy or sad. As far as I know, empirical investigations of the implications of this approach to affect have not been carried out.

The fuzziness of the boundaries of affect words may provide an explanation for their circular ordering. Pleasure and excitement, for example, are close in the circular ordering because their fuzzy boundaries overlap considerably. A series of overlapping regions with fuzzy borders would result in the continuous placement of the affect terms around the perimeter of the affect space. Indeed, it was this fuzziness between boundaries of affect categories that allowed Schlosberg (1952) to construct a circular arrangement.

Polar coordinates for the 28 words. Ross' (1938) procedure was used to compute polar coordinates for each of the 28 words. To do this simply, the eight emotion categories were assigned scale coordinates based on their theoretical (rather than empirical) circular ordering shown in Figure 1: pleasure = 0°, excitement = 45°, arousal = 90°, distress = 135°, misery = 180°, depression = 225°, sleepiness = 270°, and contentment = 315°. Polar coordinates for the 28 words are available from the author and plotted in Figure 2. Ross' procedure also provides a value P for each word. P is the length of the vector from the origin and is a measure of the precision of the angle given. P can vary from zero (meaning that subjects sorted the word into categories randomly) to 1.0 (meaning that all subjects agreed on one category into which they sorted the word). The obtained values of P ranged from .71 to .97 and showed a high degree of precision, or intersubject agreement, in the placement of each word on the circle.
Figure 2 shows that all 28 terms fell meaningfully along the circle. Beginning with happy at 7.8°, we can see that increases in angle at this point in the circle correspond to the increases in arousal and slight decreases in pleasure. The next word, delighted (24.9°), involves pleasure with some arousal. Excited fell at 48.6° and involves more arousal. Further on, even more arousal but less pleasure is seen in astonished (69.8°) and finally aroused (73.8°). Beyond 90° further increases in the angle begin to involve displeasure and less arousal: tense (92.8°), alarmed (96.5°), and so on. The angle 180° should correspond to maximum displeasure, and further increases in angle correspond to falling arousal with less displeasure: from miserable (188.7°), sad (207.5°), through to droopy (256.6°) and tired (267.7°). The angle 270° should correspond to minimum arousal and neutrality on pleasure-displeasure, and the term sleepy fell at 271.9°. Further increases in angle involve increases both in pleasure and arousal, moving from sleepy through calm (316.2°), serene (328.6°), to pleased (353.2°), at which point we are again close to happy (7.8° = 367.8°).

Figure 2 thus supports the three major hypotheses stated earlier about the cognitive structure for affect. Affective space is bipolar, with antonyms falling approximately 180° apart. The horizontal and vertical axes are easily interpretable as the proposed pleasure-displeasure and degree of arousal dimensions. And affective space lacks “simple structure.” Rather than clusters of synonyms falling near the axes, terms spread out more or less continuously around the perimeter of the space.

**Multidimensional Scaling of Terms**

**Method**

**Subjects.** Subjects were 34 University of British Columbia undergraduates of both sexes paid to participate in the study. There was some worry that the results of the previous scaling were influenced by material learned in psychology courses, and therefore subjects who had never enrolled in a psychology course were recruited for this study.

**Procedure.** Each subject was given the set of 28 emotion terms, with each word printed on a separate card, and asked to sort the cards into 4, 7, 10, and 13 groups in successive trials. Instructions were to group together emotional states that were more similar.
The similarity of each pair of words for a subject was assessed by the number of trials in which the pair was placed in the same group, with the score for each trial (each sort) weighted by the number of alternatives available in that sort. For example, a score of 13 was given to a pair of words placed in the same group during the trial in which the subject sorted into 13 groups. In addition, a score of 1 was added to each pair, since they presumably would have been placed in the same group in a degenerate sort into one group. Thus, minimum similarity was 1; maximum possible similarity was $1 + 4 + 7 + 11 + 13 = 36$, which would have occurred had the subject placed the pair in the same group on all trials. This sorting procedure was employed because it is an easy and quick task for subjects and has been shown to yield a similarity measure as adequate as that yielded by a paired-comparison procedure (Ward, 1977). A final similarity matrix was formed by taking the mean entry across subjects for each cell of the matrix.

**Results**

The resulting similarity matrix was analyzed by the Guttman-Lingoes (Lingoes, 1965, 1973) SSA-1 procedure. This multidimensional scaling procedure provides a geometric representation of the relations among the 28 words by placing them in a space (Euclidean space was used here) of specified dimensionality. Greater similarity between two words is represented by their closeness in the space. Solutions for one to six dimensions were obtained; their stress (badness of fit) values were: .288, .073, .053, .039, .029, and .022. Since stress values must decline or remain constant with increasing dimensionality simply by fitting error variance, the point of diminishing returns is often taken as one sign of the appropriate number of dimensions in the final solution. This “elbow” clearly occurred at two dimensions, since further dimensions produced only trivial decrements in stress. Moreover, in no case was a higher dimension readily interpretable. Thus, all indications supported the two-dimensional solution, and coordinates from this solution (available from the author) are plotted in Figure 3. Figure 3 shows a remarkable resemblance to Figure 2, despite differences in the measurement model and procedures employed. The similarity between these two solutions will be discussed later.

**Unidimensional Scaling**

From the results so far, it appears that the 28 terms are definable in terms of a two-dimensional bipolar space. It was also hypothesized that this space could be defined in terms of two orthogonal dimensions,
pleasure-displeasure and degree of arousal. Results of a unidimensional scaling of the 28 words along these two dimensions are therefore shown in Figure 4.

Some data for this scaling were taken from previous scaling work reported by Russell and Mehrabian (1977) and Russell (1978). Two words for which no previous data were available were scaled with the same procedure as described in the earlier articles, and the procedure will therefore only be described briefly here. Each word was rated on Mehrabian and Russell's (1974) scales of pleasure-displeasure and degree of arousal. These are six-item scales in a semantic differential format. There had been a minimum of 27 subjects who rated each term in the previous studies; 51 subjects rated the two additional words. In Figure 4, the pleasure and arousal axes are assumed to be orthogonal. In the sample of 28 words, the actual correlation was .03.

Quantitative Comparison Among Scalings

From Figures 2, 3, and 4, it appears that the three scaling techniques yielded equivalent results. A quantitative assessment of this equivalence was made by calculating the “average redundancy” yielded by a canonical correlational analysis between the scalings, taken two at a time. Redundancy is the amount of variance accounted for in one set of dimensions (e.g., the two-dimensional multidimensional-scaling solution) by another (e.g., the pleasure and arousal ratings.) Another redundancy value can be calculated as the variance in the second set accounted for by the first. Average redundancy is the mean of these two.

Results of this analysis are given in Table 3. Each solution accounted for between 94 and 95% of the variance in the other solution. Clearly, the three scaling solutions yielded nearly identical results, as had been suggested by visual inspection.

Discussion

Despite differences in the measurement models and in the procedures employed in collecting data, three different scaling techniques yielded a remarkably consistent picture. Multidimensional scaling is an explora-

![Figure 4. Unidimensional scaling of 28 affect words on pleasure-displeasure (horizontal axis) and degree of arousal (vertical axis).](image-url)
Table 3
Average Redundancy Among Three Scalings of 28 Affect Terms

<table>
<thead>
<tr>
<th>Scaling</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Direct circular scaling (Figure 2)</td>
<td>.95</td>
<td>.95</td>
<td></td>
</tr>
<tr>
<td>2. SSA-1 solution (Figure 3)</td>
<td>.94</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Pleasure and arousal unidimensional scaling (Figure 4)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. All values reported are significant at \( p \leq .01 \) \((N = 28)\). Coordinates used for direct circular scaling were the sine and cosine of the angle for each term.

tory procedure that puts few restraints on the resulting geometrical configuration. This technique nevertheless yielded results closely fitting the proposed theoretical structure. Ross' (1938) direct circular scaling did assume a two-dimensional space with maximum values falling on the perimeter of a circle. The actual data easily fit this model, with stimulus words spread fairly evenly around and very close to the perimeter (results that need not have occurred). Finally, the same conclusions were obtained in the unidimensional scaling with measures of pleasure–displeasure and degree of arousal. Again, the words were seen to spread out in a circular order around the perimeter of the space.

These results thus support the more general view that laymen have a mental map of affective life on which they rely in a variety of situations. Although the subjects in the first study may never before have been asked to place emotion names on the circumference of a circle, on reflection most could do so quite easily. Subjects have therefore not learned how to perform specific tasks, but have organized their knowledge of affect into a structure that is then utilized in whatever task is presented.

One characterization of the obtained structure (or mental map) is provided by the pleasure–displeasure and degree-of-arousal concepts, and the unidimensional scaling of the 28 words along these two dimensions clearly supported such an interpretation. On the other hand, it is also clear that affect terms fall meaningfully in the quadrants of the space as well as near the axes. A 45° rotation of the pleasure and arousal axes would thus result in two bipolar axes that could be labeled excitement–depression and distress–contentment (see Figure 1). Lundberg and Devine (1975) carried out a multi-dimensional scaling of emotion nouns, and two of their obtained dimensions were labeled depression–elation and contented–discontented. As the present results show, Lundberg and Devine's (1975) results should not be interpreted as conflicting with, but as rotational variants of, the pleasure and arousal dimensions. Rotation of the axes yields complementary, not contradictory, characterizations of the cognitive structure for affect.

The circular ordering of the variables provides still another characterization of affective space, although a disk or a wheel would be a better image, since affective states of moderate intensity would fall toward the middle of the space, with the origin presumably corresponding to adaptation level or a neutral feeling. Again, however, the circular ordering of variables complements rather than contradicts the dimensional characterization of affect.

Structure of Self-Reported Affect

As already mentioned, factor-analytic studies of self-reported affect, beginning with Nowlis' pioneering work (e.g., Nowlis & Nowlis, 1956), have resulted in a view of affect as a set of between 6 and 12 independent monopolar factors (e.g., Borgatta, 1961; Clyde, 1963; Curran & Cattell, 1975; Izard, 1977; Lorr, Daston, & Smith, 1967; McNair & Lorr, 1964; Nowlis, 1965). This commonly accepted representation is clearly at odds with the layman's representation of affect as presented in Figure 1. Of course, it could be that the layman's is a mistaken, perhaps overly simplified, view. Nonetheless, evidence is starting to accumulate against the view of affect as a large number of independent monopolar factors. The new evidence favors a view much like that of Figure 1.
When measures of the individual monopolar factors have been developed, they have been found to be moderately, or occasionally highly, intercorrelated rather than independent as had been assumed (Russell & Mehrabian, 1977). Meddis (1972) pointed out that the emergence of monopolar rather than bipolar factors may have been due to the rating formats typically used to gather affect data. Meddis (1972) showed that using a different rating format resulted in evidence of a smaller number of bipolar dimensions, and Svensson (1978) replicated this result. Russell (1979) suggested further methodological refinements and obtained even stronger evidence favoring bipolar dimensions in self-report data. Thayer (1978), who had provided some of the earlier evidence favoring monopolar factors (Thayer, 1967), more recently has himself obtained evidence more consistent with bipolar dimensions. Mackay, Cox, Burrows, and Lazzerini (1978) also found that methodological changes in Thayer's (1967) scales resulted in bipolar dimensions.

Once bipolar affect factors are assumed, it appears that the major proportion of variance is accounted for by only two factors. Various interpretations of the two factors have been offered, but one interpretation is pleasure-displeasure and degree of arousal (Russell, 1979; Russell & Mehrabian, 1974, 1977; Svensson, 1978). Moreover, Thayer's (1978) two bipolar dimensions appear to be interpretable as approximately rotational variants of pleasure and arousal, namely, excitement-depression and distress-relaxation (see Russell, 1979). In these same studies, additional factors were found to account for only small proportions of variance and, more importantly, to be interpretable as not directly describing emotions but rather cognitive or social correlates of emotion. For example, Russell and Mehrabian's (1974, 1977) factor of dominance-submission and Svensson's (1978) factor of social orientation can be interpreted as referring to perceived aspects of the antecedents or consequences of the emotion rather than to the emotion per se.

Finally, there is some evidence that affect states fall in a circular order around the perimeter of the two-dimensional bipolar space. Plutchik (1962) proposed a circular ordering of emotions, and Schaefer and Plutchik (1966) showed that the affective states associated with various psychiatric diagnostic categories form a circular order when these variables are plotted against their first two principal components. Russell (1979) gathered self-report data with scales of pleasure, displeasure, arousal, sleepiness, depression, and (based on a reinterpretation of Thayer's, 1967, scales) excitement, relaxation, and distress. These eight scales were also found to fall in a circular order when plotted against their first two principal components.

In short, there is no longer a consensus on the structure of affective experience, in part because of investigators' prior choices regarding methodology and definition of the domain of affect. Methodological bias such as rating format may account for the early factor-analytic results of a large set of monopolar factors. Moreover, there is growing evidence that a large part of the variance in self-report data can be accounted for by a structure very similar to the one illustrated in Figure 1. That is, self-report data can be characterized by the same three properties that are geometrically represented in that model: affective states falling in a circular order in a two-dimensional bipolar space, the axes of which are interpretable as pleasure-displeasure and degree of arousal.

If so, the structure proposed in Figure 1 appears to be more suitable to a circumplex measurement model than to the more commonly considered factor-analytic one. In a circumplex (Guttman, 1954), variables are arrayed in a circular order in a geometrical space of small dimensionality. One interesting property of a circumplex is that since rotation of the axes leaves the circular configuration of the variables intact, rotation is considered arbitrary. Plutchik (1970) and Sjoberg and Svensson (1976) argued that common factor analysis is an erroneous model for representing affect data and suggested exploring the utility of Guttman's (1954) model in this domain. If the circumplex is appropriate for affect, it can provide a very precise nomological network of testable propositions whereby every variable is related to every
other variable within the domain (Wiggins, 1979).

The following study was carried out to test the ability of the model proposed in Figure 1 to represent the structure of self-reported affect. The analyses were thus considered hypothesis-testing rather than exploratory.1

Method

The data used in the present analyses were gathered as part of a joint project with Jerry S. Wiggins on a taxonomy of affect-descriptive terms. The subject sample consisted of 343 male and female students enrolled at the University of British Columbia during the 1976–77, 1977–78, and 1978–79 academic years. Approximately half of the subjects were enrolled in some type of psychology course and received course credit for their participation. The remaining subjects were recruited from the campus at large and were paid for their participation.

The data were collected in group sessions lasting between 11 and 7 hours. Subjects first responded to Mehrabian and Russell’s (1974) state affect scales of pleasure–displeasure, degree of arousal, and dominance–submission. Each subject then rated how accurately each of 518 affect-descriptive adjectives described how he or she felt so far that day. The list of 518 adjectives was presented to each subject in a different random order. Accuracy ratings were made on an eight-place scale ranging from 1 (extremely inaccurate) to 8 (extremely accurate).

Previous work on response to single adjectives (as opposed to, for example, bipolar rating scales) has indicated the existence of a stylistic individual difference variable in the use of the rating scale (Bentler, 1969, 1973; Russell & Mehrabian, 1977; Russell, 1979). Since response style variance of this sort can influence the correlation matrix and factor structure, it was minimized in the present data by ipsatization: All subjects’ data were thus equated with respect to mean and variance across the 518 words, and all analyses were carried out on the ipsatized data. Ipsatization removes variance on a general factor on which all items load equally and positively. Because of the heterogeneity within the sample of affect terms, including many antonym pairs, a general factor in these data could only be interpreted as individual differences in response to the rating format rather than in response to the content of the items.

For this sample, the means on Mehrabian and Russell’s (1974) scales were 36.63 (SD = 8.04) for pleasure–displeasure and 29.96 (SD = 7.12) for degree of arousal. Both scales were reasonably reliable, with coefficient alpha values of .86 and .74, respectively. These two scales intercorrelated .36 (.45 when corrected for attenuation).

Twenty-six of the 28 words analyzed in the first three studies were included in the list of 518. For the two missing words, the nearest synonyms were chosen: drowsy was used in place of droopy, and tranquil in place of serene.

Results

A first test of the proposed structure of affect was carried out by placing the 28 affect variables into the two-dimensional space defined by Mehrabian and Russell’s (1974) scales of pleasure–displeasure and degree of arousal. Each of the 28 variables was regressed onto the two bipolar scales, and Figure 5 shows the resulting pattern of relationships, using beta weights from the regression analyses as coordinates. Its resemblance to Figures 2, 3, and 4 is striking, although the cluster of variables depressed, sad, and gloomy is somewhat rotated toward the horizontal axis.

Multiple correlations of each of the 28 words with the pleasure and arousal scales showed a significant (p < .001) relationship in every case, although the Rs ranged from only .22 to .62. The smallness of magnitude can in large part, but not wholly, be accounted for by unreliability of responses to single items. The reliability of these single-item scales cannot be estimated directly from the data available, but some indication comes from examining the maximum correlation of each item with the other 27 items: These values ranged in magnitude from .28 to .65. Another indication of the reliability of single items for this sample of subjects can be seen in the reliability of one-item versions of the Mehrabian and Russell (1974) pleasure and arousal scales, as estimated by the Spearman–Brown prophecy formula from the reliability of the full scale; these were .51 for pleasure and .32 for arousal.

Analysis as a circumplex. The fit of data to a circumplex model is generally tested by examining a plot of the variables as functions of their loadings on two principal components

---

1While this article was under editorial review, Plutchik (1986) published a specific circumplex model of affect and reported previously unpublished data from H. R. Conte’s (1975) doctoral dissertation supporting the model. Their model is similar, but not identical, to the one proposed here.
A CIRCUMPLEX MODEL OF AFFECT

Figure 5. Regression weights for 28 affect words as a function of pleasure-displeasure (horizontal axis) and degree of arousal (vertical axis).

(Wiggins, 1979). The first two components extracted from the intercorrelation matrix (with unities in the diagonal) of the 28 affect variables accounted for 45.8% of the total variance. Because of the earlier evidence of a correlation between the pleasure and arousal axes in these data, an oblique (direct oblimin) rotation of the two components was used, but the resulting correlation between components was only .12. This solution is shown in Figure 6 (where for convenience the axes are drawn as if orthogonal). The resemblance of Figure 6 to Figure 5 is self-evident, including the shift of depressed, sad, and gloomy toward the horizontal pleasure-displeasure axis.

Second-order principal-components analysis. Besides these first two principal components, there were three additional components with eigenvalues greater than unity, although a scree test showed a clear elbow at two components and the next three components together accounted for only an additional 13.1% of the total variance. (The five-component solution thus accounted for a total of 58.9% of the total variance.)

The five-component solution was orthogonally rotated to a Varimax criterion, and the components were labeled by their highest loading items. The first two components contained most of the items and were clearly bipolar. The first component was labeled happy-sad, contrasting words such as happy, delighted, and pleased with sad, depressed, and miserable. The second component was labeled tense-relaxed, contrasting tense and frustrated with relaxed, calm, tranquil, and at ease. The third component, labeled sleepy, was monopolar and consisted of sleepy, tired, and drowsy. The fourth component, labeled angry, was also monopolar and consisted of angry and annoyed. The fifth component, labeled alarmed, consisted of three positively loading items, alarmed, astonished, and afraid, and one negatively loading item, bored.

The emergence of monopolar factors is reminiscent of earlier factor-analytic work in
which it was concluded that a large number of factors of affect are independent. Independence of the five factors was therefore explored in a second-order factor analysis. Each item was assigned to the one component on which it loaded the highest, and component scores were then obtained by taking unweighted algebraic means on the set of items constituting each component. A principal-components analysis (with unities in the diagonal) of these five variables yielded two components with eigenvalues greater than 1.0 and which together accounted for 70.3% of the total variance. Figure 7 gives the plot of the factor loadings of the five variables after an orthogonal Varimax rotation. To clarify the structure, reflections of the first two components (sad for happy and relaxed for tense) are also shown and labeled, since both these components were clearly bipolar and thus could just as reasonably have been scored in the opposite direction. These results indicate that the five components were not independent. Indeed, Figure 7 roughly resembles the two-dimensional bipolar affective space seen earlier and thus again indicates that it can provide an adequate representation of self-report data.

**Discussion**

The present results showed that the circumplex model of affect illustrated in Figure 1 accounts for a substantial proportion, but not all, of the variance in self-reported affective states. Before discussing the success of the circumplex model, it may be helpful to ask what accounts for the remaining variance in

![Figure 7. Second-order principal components of five components of self-report data.](image-url)
self-report data, both for the present results and for similar results in general.

Some of this variance is simply unreliable, of course, and accounted for in terms of the inevitable errors of measurement.

A second source of variance can be seen in the various nonaffective but reliable content factors occasionally found in analyses of self-report data. Sjoberg and Svensson (1976), for example, found factors interpretable as social orientation and control. Russell and Mehrabian (1977) found that a dominance—submissiveness dimension accounted for a small but significant proportion of variance in affect scales beyond that accounted for by pleasure and arousal dimensions.

Third, further variance may be accounted for by a more detailed account of the process of labeling an affective state. There is already ample evidence that some variance in self-report data is accounted for by individual differences in the use of the rating scale. There is a consistent individual difference in the tendency to acquiesce—to rate words as more versus less accurate regardless of content (Bentler, 1969, 1973; Russell, 1979; Russell & Mehrabian, 1977). There was an attempt to eliminate acquiescence from the self-report data analyzed here, but there are other stylistic factors not as easily eliminated. For example, Johnston and Hackmann's (1977) study of affect data showed consistent individual differences in the use of the extreme ends (both positive and negative) of the rating scale.

More importantly, the model underlying factor analysis assumes that observed scores on each affect rating scale are a linear combination of scores on the underlying factors. The same assumption was made of course in the multiple regression analysis carried out here, since each affect item was explored as a linear combination of the pleasure and arousal dimensions. Linear combination rules are remarkably robust in approximating various combination rules (Dawes, 1979), but the process by which a person actually arrives at a rating for each affect term might not be a linear combination of the values on the underlying dimensions. Specification of the subjects' actual process should then account for more of the variance in self-report data. Suggestive evidence in this direction comes from a study by Russell and Mehrabian (1977), in which some affective states were found to be definable not as linear combinations but as multiplicative combinations of the pleasure and arousal dimensions, and from a study by Bush (Note 1) in which it was found that some subjects appear to use a linear combination rule but that others (a minority) use a multiplicative combination rule.

A fourth reason that the representation of Figures 1, 2, 3 and 4 differs from the factor-analytic structure is that the latter includes additional information. In Figure 2 the position of the affect terms reflects their conceptual overlap. Happy is close to delighted because they overlap in meaning. Happy is approximately orthogonal to sleepy because they are conceptually distinct. In a factor-analytic study of self-report data, the correlation coefficient measures this degree of conceptual overlap, since a subject who checks happy also typically checks delighted not because these are two separate events that happen to co-occur but because both words often describe the same state of affairs. The correlation coefficient also measures the likelihood of co-occurrence, however. For example, the first set of studies here indicated that pleasure—displeasure was conceptually distinct from degree of arousal. These two dimensions may, however, still co-vary in the real world. (Mental age and weight, although conceptually distinct, covary empirically). Indeed, it was found that with self-report data, scales of pleasure and arousal were moderately intercorrelated.

Pleasure and arousal, even if conceptually distinct, will be positively correlated if the affective states in which these two components occur with equal sign, such as delighted (+ on pleasure and + on arousal) or bored (− on pleasure and − on arousal), occur more frequently than do the affective states in which the two components occur with opposite signs, such as angry (− on pleasure and + on arousal) or tranquil (+ on pleasure and − on arousal). It is surely conceivable that more persons are bored than are
angry or that more persons are delighted than are tranquil. (In our sample, the mean score on *bored* was greater than the mean score on *angry*, and the mean score on *delighted* was greater than the mean score on *tranquil*). These differences in the frequency of various affective states are reflected in the correlations observed among measures of those states and in turn are reflected in the factor structure of those states. The frequency with which the various affective states occur is a separate issue, however, from the conceptual components that make up the meaning of that state.

In short, exploring affect through self-report data from a large sample of subjects will inevitably result in some variance unique to that methodology, just as the judgment tasks employed in the first three studies here inevitably resulted in some unique variance. What is then of more interest than this unique variance is the present finding that four different scaling methods shared much variance in common.

### General Discussion

Up to now I have maintained the commonly accepted distinction between results from introspective self-report data on affective experience and results from judgment data on, for example, the similarity between emotion names. Analysis of self-report data is generally considered to reveal the structure of actual affective experience, whereas analysis of judgment data is generally considered to yield merely semantic structures and is not considered capable of revealing the structure of actual experience. Some explanation is therefore required for the present finding that both types of data yielded very much the same structure. In the domain of personality traits, a controversy arose when a similar correspondence was found between the structure derived from self-ratings of personality traits and the structure derived from the judged similarity among trait-descriptive terms. Although some psychologists claim that the structure yielded by the two types of data represents the actual relationships among personality traits (e.g., Block, Weiss, & Thorne, 1979), others claim that the structure is “primarily an artifact of the rater’s or the questionnaire taker’s cognitive structure, and not a reflection of the real world” (D’Andrade, 1974, p. 181).

In the domain of affect, I would like to argue that the structure in common to both types of data is a model both for the layman’s cognitive structure for affect and for the actual structure of affective experience. Stated briefly, my thesis is that affective experience itself is the end product of a cognitive process that has already utilized that same cognitive structure for affect. The affective state, as it is experienced, is already meaningful. It is meaningful because a cognitive process has already occurred that interpreted (gave meaning to) the emotion.

People are not typically aware of all the pieces of information they rely on in analyzing their own emotional states. If people were aware of this process, how would we account for the very discrepant theories of the antecedents of affect that have been suggested by psychologists? These views include James’ (1890) argument that we rely on internal feedback from visceral stirrings, Bem’s (1967) argument that we rely on our own external behavior, Schachter and Singer’s (1962) argument that we rely on the external social situation, and Tomkins’ (1962–1963) argument that we rely on feedback from the facial muscles. It seems certain that not all people know just what information they rely on in interpreting their own emotional states. Whatever information we do rely on (and there is no reason that we cannot utilize all these various sources), the information itself does not directly produce affective experience. Rather, the information is first interpreted and made meaningful, which is to say it is used to categorize the internal, emotional state. It is then the interpretation rather than the information that we become aware of as affective experience. The experience of an affective state thus occurs only as the final result of this cognitive process. If this is the case, then the cognitive structure that is utilized in interpreting the meaning of verbal messages or of facial expressions from others is the same structure utilized in the process of conceptualizing one’s own state, which precedes the affective experience.
From the perspective just outlined, there is a cognitive structure for affect, which consists of a set of interrelated cognitive categories. This structure is utilized in a variety of situations, including the experiencing of one's own emotional state and the judging of another's emotional state. Self-report of one's affective state is thus a task like the labeling of photographs of faces and can be taken as a means of revealing the way in which emotions are conceptualized. Each such task will likely have some unique properties. Those properties that can be derived from a variety of ostensibly different tasks, each involving affect, are therefore likely to be properties of a process common to all the tasks. Current evidence suggests this process to be the cognitive conceptualization of emotion. Current evidence also suggests that the known properties of that cognitive conceptual structure are conveniently summarized by a simple circumplex model.

Reference Note


References


Received November 26, 1979