JOURNAL OF EXPERIMENTAL SOCIAL PSYCHOLOGY **30**, 1–38 (1994)

Context Effects on Similarity Judgments of Multidimensional Stimuli: Inferring the Structure of the Emotion Space

JAMES S. ROBERTS AND DOUGLAS H. WEDELL

University of South Carolina

Received November 12, 1993

Four experiments were conducted to assess the effects of stimulus context on similarity judgments of emotion words. In each experiment, the context either uniformly represented locations on Russell's (1980) two-dimensional circumplex emotion space or emphasized emotions lying within particular quadrants of that space. The manipulation of context led to systematic changes in the similarity relations among the emotion words. Multidimensional scaling of the similarity data modeled these changes in terms of an emerging third dimension that teased apart distinctions among the emotion words added in each context. Contextual effects were strongest when similarities were assessed via a sorting task and were greatly attenuated when similarities were assessed via direct pairwise ratings of similarity. The contextual dependence of similarity measures provides one possible explanation for reported discrepancies in the dimensionality of the emotion space.

Psychologists have used many different methods to uncover the basic cognitive representation of emotions. Among these have been factor analytic techniques (e.g., Izard, 1972; Watson & Tellegen, 1985; Zevon & Tellegen, 1982), cluster analysis (e.g., Gehm & Scherer, 1988; Shaver, Schwartz, Kirson, & O'Connor, 1987), and multidimensional scaling (e.g., Averill, 1975; Bush, 1973; Russell, 1980; Shaver *et al.*, 1987). In many respects, it is not surprising that the representations emerging from these different methods have not always been consistent. However, it is disconcerting when the same method yields different representations. Our investigation focuses on the use of multidimensional scaling (MDS) techniques to analyze similarity judgments of emotion words. Prior MDS research on the similarity of emotion words has led to alternative conclusions regarding the structure and, in particular, the dimensionality of the emotion space. The primary purpose of our research was to identify and test explanations of these discrepant results. In particular, we focused on the hypothesis that the set of emotion words included in a study provides the context for judging similarity relations among them. Differ-

0022-1031/94 $6.00

ences in the stimulus context can thereby lead to changes in the judged similarity relations among the stimuli, which are then reflected in different multidimensional representations. We also investigated the possibility that alternative data collection procedures show differential sensitivity to the effects of judgment context.

Before presenting our experimental design, we first review the literature on the structure of the emotion space, concentrating on the strengths and weaknesses of the circumplex representation. We then propose two basic hypotheses concerning how differences in the sets of emotions being judged might lead to differences in the corresponding MDS representations. The more interesting of these is that the similarity relations themselves are context dependent, and we review evidence that supports this hypothesis.

THE CIRCUMPLEX MODEL OF EMOTION

Many researchers have suggested that relations among emotions can be represented geometrically, with each emotion constituting a point in an *n*-dimensional space. When judgments of emotional stimuli (e.g., emotion words, facial expressions, etc.) or ratings of self-reported emotions are analyzed using spatial techniques like MDS or factor analysis, the resulting structure often takes a special form known as a circumplex (Larsen & Diener, 1992; Russell, 1989). The emotion circumplex is a two-dimensional, geometric structure in which emotion stimuli are located approximately along the perimeter of a circle. The two orthogonal dimensions of the circumplex are bipolar and have frequently been interpreted as activation and valence (Russell, 1978, 1980, 1989; Russell & Bullock, 1985; Schlosberg, 1952), although other interpretations that involve alternative rotations of the two-dimensional plane are also popular (Meyer & Shack, 1989; Watson & Tellegen, 1985; Zevon & Tellegen, 1982). Because the circumplex is a spatial model, the similarities among emotional stimuli are represented by the interpoint distances of stimuli in the two-dimensional space. Emotional stimuli that are close together are presumably more similar than stimuli that are more distant from each other.

The circumplex model of emotion is an attractive representation primarily because it accounts for a large proportion of variation in subjects' responses to emotional stimuli in a parsimonious fashion. The circumplex has been found in studies that used either MDS (Russell, 1980; Russell & Bullock, 1985) or factor analytic techniques (Watson & Tellegen, 1985; Zevon & Tellegen, 1982). It has emerged in studies in which similarity judgments of emotion words (Russell, 1980), similarity judgments of faces (Russell & Bullock, 1985), or attribute ratings of self-reported emotion were collected (Zevon & Tellegen, 1982). Thus, the circumplex model appears to have some general applicability to the study of human emotions.

PROBLEMS WITH THE CIRCUMPLEX MODEL OF EMOTION

Although the circumplex model of emotion is both parsimonious and popular, it has its share of problems. For example, some researchers have searched for "basic dimensions" within the circumplex structure. These searches have been guided by efforts to maximize the interpretability of the configuration (Russell, 1978, 1980), the simple structure exhibited by the configuration (Watson & Tellegen, 1985), or the correlations among the dimensions of the configuration and external variables (Meyer & Shack, 1989). Larsen and Diener (1992) have suggested that the search for basic dimensions within the circumplex is futile. Specifically, they note that a geometric circumplex will lack a definitive rotation based on simple structure unless certain areas within the circumplex are unequally sampled. Therefore, orientations of the emotion circumplex that are consistently found using rotations designed to maximize simple structure are the result of consistent sampling inequities. Likewise, the use of external correlates to locate basic dimensions within the emotion circumplex does not yield a single orientation. Larsen (1989) has found meaningful relationships between each octant of the emotion circumplex and at least one distinct personality factor. Because distinct external correlates are found for all octants of the emotion circumplex, Larsen and Diener (1992) have concluded that such correlations cannot offer guidance as to which rotation is most basic. In short, they recommended that the orientation of the circumplex be chosen to maximize its utility with regard to the research question at hand.

Another problem with the circumplex model of emotion can arise when specific emotions are compared to each other. For example, the emotions anger and fear are typically located close to each other in the high-activation, negative-valence portion of the circumplex (Larsen & Diener, 1992; Russell, 1980; Watson & Tellegen, 1985). The close proximity of these emotions appears reasonable because each emotion involves both high arousal and unpleasantness. The circumplex model of emotion would thus lead us to believe that anger and fear are quite similar. Although the concepts of anger and fear may indeed share qualities of high arousal and unpleasantness, psychologists typically view these two emotions as distinct, and in many ways opposing, affective states (Ekman & Friesen, 1975; Ekman, Levenson & Friesen, 1983; Frijda, 1986; Frijda, Kuipers & ter Schure, 1989; Osgood, 1966). However, conceptually important differences between these (and other) emotions may be obscured when using only the two dimensions of the circumplex.

Indeed, many researchers have found more than two dimensions when attempting to represent spatially the structure of the emotions (Averill, 1975; Bush, 1973; Daly, Lancee & Polivy, 1983; Russell, 1978; Russell & Mehrabian, 1977; Schlosberg, 1954; Shaver *et al.*, 1987; Smith & Ellsworth, 1985). These additional dimensions have served to better differentiate emotions, including the differences noted between anger and fear.

STIMULUS CONTEXT AND THE DIMENSIONALITY OF THE EMOTION SPACE

Divergent reports of the dimensionality and structure of the emotion space may be due to a variety of differences between studies designed to reveal that structure. Differences in the emotional stimuli employed, the data collection procedures implemented, the particular responses collected, and the methods of data analysis used may all have contributed to the discrepant findings. However, divergent results have emerged in studies that have been conducted in a fairly similar fashion. For example, both Russell (1980) and Shaver *et al.* (1987) used a sorting procedure to collect subjects' similarity judgments of emotion words. The data from the two studies were analyzed with comparable MDS techniques; yet the studies produced disparate configurations. Russell's results revealed a two-dimensional circumplex structure of emotion in which the dimensions were oriented to reflect arousal and valence. Moreover, the emotion words *angry* and *afraid* were located relatively close together in that part of the space indicative of both high arousal and unpleasant affect. In addition to the dimensions of arousal and valence, Shaver *et al.* provided evidence for a third dimension that reflected potency (in the sense used by Osgood, Suci, & Tannenbaum, 1957). In order to gain a better understanding of the Shaver *et al.* solution, we rotated the first two dimensions of their configuration to the orientation described by Russell (1980; Fig. 3). From this perspective, the third dimension of the Shaver *et al.* solution served to distinguish between emotions related to anger and fear. It also differentiated passionate arousal from nonpassionate arousal (e.g., lust from astonishment).

Although there is a great deal of communality between the Russell (1980) and the Shaver *et al.* (1987) studies, they differed in a number of ways. Procedurally, the studies differed in the number of words sorted (28 vs 135), the number of times they were sorted (4 vs 1), and the way in which the number of sorting categories was specified (by the experimenter vs by the subject). Any of these differences may have played a role in producing the divergent findings. However, a crucial factor from our perspective lies in the distributional properties of the emotion words sampled in each study. Specifically, we hypothesized that the context of the judgment situation (i.e., the set of emotion stimuli considered by a subject within the judgment situation) can substantially affect the emergent structure of the emotion space.

Differences in the distributional representation of emotion words might

lead to differences in MDS solutions in either (or both) of two ways. The first of these we refer to as the *scaling artifact hypothesis*, which argues that the differences in stimulus contexts do not affect the relations among judged similarities themselves, but only the ability of the scaling procedure to detect and model these differences. Specifically, one could speculate that a MDS procedure may fail to provide evidence for certain dimensions if there are relatively few comparisons involving emotions that differ along those dimensions. In such cases the data do not provide sufficient information to define underrepresented dimensions as distinct from mere "error" in the data. Implicit in this hypothesis is the presumption that subjects' similarity judgments may differentiate emotions like anger and fear, but when there are only a few comparisons among these two types of emotions, the scaling procedure cannot reliably incorporate the judged differences into the solution.

A second and substantively more interesting explanation may be referred to as the *contextual similarity hypothesis*. It posits that the evaluation of stimulus qualities are not independent of the contextual stimuli presented within the judgment setting. An implication of this hypothesis is that judgments of similarity among emotion words may differ due to stimulus context. For example, one stimulus context may cause subjects to differentiate among emotion words using a particular set of dimensions, whereas another context may lead subjects to neglect one or more of those same dimensions.

The scaling artifact hypothesis suggests that the MDS procedure will fail to uncover the structure underlying the emotion space when that space is not adequately represented. This explanation implies that the results of the MDS procedure, rather than the actual similarity judgments, are dependent on stimulus context. On the other hand, the contextual similarity hypothesis argues that judgments of similarity may themselves be dependent on the set of contextual stimuli being compared. In the next section we explore evidence and theoretical models relating to this second hypothesis.

MODELS OF JUDGED SIMILARITY AND THE IMPLICATIONS OF CONTEXT

Two general frameworks that have been proposed in the psychological literature for conceptualizing similarity are geometric models and feature-based models. In this section, we describe theoretical models and empirical evidence relevant to the contextual similarity hypothesis.

Geometric models of similarity (Torgerson, 1958; Shepard, 1962a,b) assume that the judged similarity between two objects is a function of the distance between them in the perceptual (psychological) space. The function relating judged similarity and psychological distance can presumably be linear (Torgerson, 1958) or monotonic (Shepard, 1962a,b). However, as Tversky (1977) has pointed out, the traditional geometric model does not consider the effects of context and cannot explain a variety of experimental results which include violations of symmetry, minimality, and the triangle inequality.

To explain these violations, Tversky (1977) developed a feature model of similarity, according to which the judged similarity between two objects is a weighted function of common and unique features. The weight given a particular feature is dependent on the intensity of the feature (i.e., loudness, brightness, etc.) and the degree to which the feature facilitates the development of a stimulus classification rule (i.e., the diagnosticity of a feature). Tversky has argued that the diagnosticity of a feature is highly sensitive to context. For example, he showed that the judged similarity between a target country ("Austria") and two prespecified countries in a three-element set ("Sweden," "Hungary," and a third country) could be substantially altered by varying the third element of the set. Tversky concluded that the third stimulus modified the diagnosticity of the features inherent in the task, which led to different classification rules for determining similarity. The usefulness of Tversky's model in explaining similarity judgments has been demonstrated in subsequent work (Gati & Tversky, 1982, 1984; Tversky & Gati, 1982).

Although Tversky's (1977) feature-based representation is attractive, it does not map very well onto geometric descriptions of similarity that have proven both predictive and coherent (e.g., the circumplex structure of color, Shepard, 1962b). It also fails to capture potential context effects that may be linked to manipulations of the relative frequency or density of stimuli at different locations along an attribute dimension. There is an extensive literature documenting density effect on attribute judgments such as the attractiveness of faces (Wedell, Parducci, & Geiselman, 1987), happiness of life events (Wedell & Parducci, 1988), or magnitude of numbers (Birnbaum, 1974). These contextual density effects are captured nicely by the frequency principle of Parducci's (1963) range–frequency theory of judgment, according to which increasing the density between two stimuli along a dimension increases the difference in the rated values of those stimuli. If these differences in rated values serve as a basis for similarity judgment, then increasing contextual density within a region should lead to decreased similarity of stimuli within that region.

Building on these documented density effects for attribute judgments, Krumhansl (1978) developed an extension of the geometric model that incorporated a density principle to explain violations of the traditional geometric model. Her distance–density model proposes that similarity is a function of both the interpoint distance between stimuli and the local densities surrounding them. Specifically, the model suggests that the judged similarity between two objects, x and y, is a function of the psychological distance between x and y and the density of the psychological space around both x and y. Other things being equal, if the neighborhood(s) of the psychological space containing either stimulus (or both) is dense, then the judged similarity between the stimuli will be relatively less than when the neighborhood(s) is sparse.

Because the density of the psychological space is largely determined by the stimulus context, the model suggests that altering the contextual densities of the stimulus space will lead to differences in the perceived similarity among the stimuli within that space. This may be accomplished in one of two ways. First, increasing the dispersion of stimuli along a dimension may result in that dimension receiving greater weight when similarity is assessed. Thus, when there is little dispersion of stimuli along a dimension such as potency, differences along the potency dimension receive little if any weight. It is as if a three-dimensional structure of emotions is effectively collapsed to just two dimensions in assessing similarity. Second, increasing local densities of stimuli may lead to increased dissimilarity while the dimensional weights remain constant following the frequency principle. These types of context effects have been demonstrated many times within the attribute judgment domain. Indeed, Russell and Fehr (1987) have shown similar effects on ratings of the arousal and pleasure dimensions underlying the representation of emotions.

CONTEXTUAL SIMILARITY AS A FUNCTION OF TASK CONSTRAINTS

Krumhansl (1978) provided initial support for the distance–density model by fitting data from past nonexperimental studies that were typically based on analyzing confusion matrices generated from identification and discrimination tasks. However, Corter (1987) questioned whether density effects could be produced experimentally and whether they would impact direct ratings of similarity. In six experiments, he found little empirical support for the density hypothesis. Only when the similarity was indirectly assessed through confusion matrices based on identification tasks did he find significant effects of density. Furthermore, these effects were in the opposite direction than postulated by the distance–density model in that similarity was increased by increasing local density. He concluded that there was little evidence for density effects on direct pairwise ratings but that there may be some effects of density when a more indirect measure of similarity is used.

In developing his experimental tests, Corter cited work by Mellers and Birnbaum (1982) which had demonstrated that pairwise difference judgments along the dimension of lightness/darkness of dot patterns were insensitive to manipulations of contextual densities. Density had no effects on pairwise difference ratings even though the usual density effects were obtained on direct attribute ratings in accordance with Parducci's (1963) range–frequency theory. Pursuing this line of investigation, Wedell (1991) argued that density effects may operate through associative or constructive mechanisms that depend on both the nature of the task and the stimuli. He replicated the Mellers and Birnbaum null effects of context on pairwise similarity ratings of dot patterns, but demonstrated that significant density effects occur when attribute ratings precede the similarity rating task. These effects were attributed to an associative mechanism where by the contextually dependent attribute ratings were encoded into memory and later used to mediate similarity judgments.

Wedell (1991) also postulated a constructive mechanism according to which subjects use density information to construct scale values upon which they will base similarity judgments. Assuming that the construction process requires time and is most likely to be contextually dependent when each stimulus must be processed holistically rather than in strictly a pairwise fashion, he introduced a 3-s delay between presentation of members of the pair being rated. Strong density effects on pairwise similarity ratings emerged under the delay condition, and these findings were replicated using direct similarity ratings of squares.

Wedell's (1991) results present the strongest evidence to date that manipulation of contextual densities can affect direct ratings of similarity. However, the generality of these results is limited because they are based on stimuli that varied along just one dimension. Jones and Wedell (1987) failed to find strong effects of density on pairwise similarity ratings of schematic faces that were manipulated along two dimensions. However, they did find significant density effects for stimuli when densities were derived from a sorting task.

OVERVIEW OF EXPERIMENTS

The literature we have reviewed provides limited evidence that similarity relations among stimuli may be affected by manipulating the contextual set of stimuli being judged. We tested this hypothesis in a series of four experiments that directly assessed whether manipulating the contextual set of emotion words leads to (a) differences in the emotion space inferred by MDS techniques and (b) differences in the similarity judgments themselves. Because the literature has demonstrated that the different tasks for gathering similarity data may be differentially sensitive to manipulations of context, we explored variations on two basic tasks, (a) similarity sorting and (b) direct pairwise rating.

The same basic design was used in the four experiments and is described here to provide the reader with a concrete sense of the nature of the contextual manipulation. In each experiment, the set of emotion words presented to subjects was varied between subjects in order to reflect three different contextual conditions. In each condition, there were a total of 25 emotion words presented, of which 13 emotion words (the core set) appeared in all three conditions. As shown in the upper left of Fig. 1, these core stimuli were chosen to be uniformly distributed around the

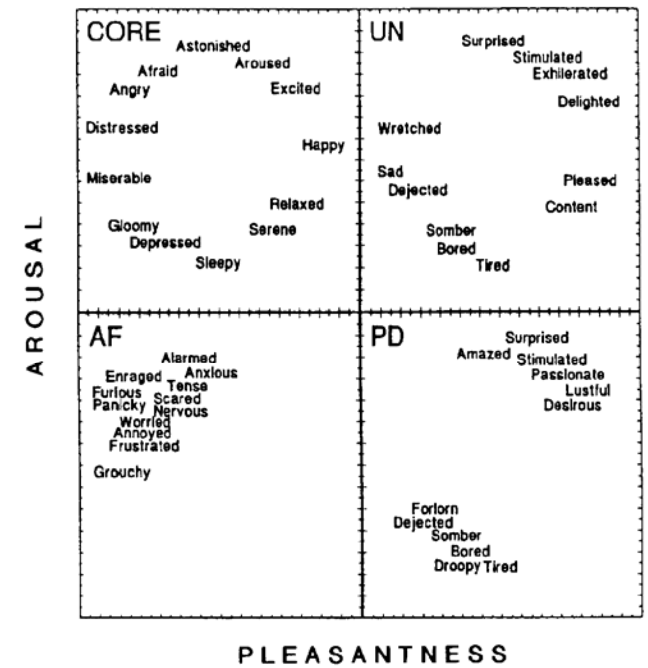

FIG. 1. Illustration of the contextual manipulation used in all four experiments. Emotion words are placed at approximate locations in the two-dimensional emotion space.

circumplex described by Russell (1980, Fig. 3). Thus, the core stimuli represented each of the four quadrants from Russell's two-dimensional solution to an approximately equal degree. Indeed, Russell (1980, Fig. 1) described the ideal locations of eight affect concepts that were equally spaced around the circumplex (i.e., arousal, excitement, pleasure, contentment, sleepiness, depression, misery, and distress). Our core set included at least one emotion word for each of these eight basic affect concepts.

The remaining 12 stimuli composing a given contextual condition were chosen to differentially manipulate stimulus density within the two-dimensional emotion circumplex. The 12 contextual stimuli from the *Uniform* (UN) stimulus condition are shown in the upper right of Fig. 1. These contextual stimuli were combined with the core stimuli to produce a set of 25 emotion words that formed the UN stimulus condition. The contextual stimuli were chosen to maintain a roughly uniform representation of Russell's stimulus space when presented with the core stimuli. Half of these contextual stimuli were actually used by Russell (1980), and thus their locations in his two-dimensional space were (approximately) known. The locations for the other 6 contextual stimuli were based on speculation.[1]

The contextual stimuli for the *Anger–Fear* (AF) stimulus condition (shown in the lower left of Fig. 1) consisted of emotions related to either anger or fear and were thought to correspond to the second quadrant of Russell's (1980, Fig. 3) stimulus space. When these stimuli were combined with the 13 core stimuli, the resulting AF stimulus condition represented all quadrants of Russell's circumplex, although the quadrant associated with unpleasant, high-arousal emotions was overrepresented.

In the *Passion–Depression* (PD) stimulus condition, the contextual stimuli (shown in the lower right of Fig. 1) presumably exemplified quadrants one and threee of Russell's space, and thus, were thought to emphasize pleasant arousal and depression, respectively. Of the six contextual stimuli chosen to exemplify positive arousal, half were typical of general arousal (e.g., *stimulated, surprised,* and *amazed*) whereas the remaining three stimuli reflected passion (*passionate, lustful,* and *desirous*). The sets of stimuli shown in Fig. 1 were used in all four experiments. In each experiment, the basic question was whether or not the contextual stimuli would differentially affect judgments of similarity among certain core stimuli and their corresponding locations in the MDS space.

EXPERIMENT 1

The purpose of Experiment 1 was to assess if alternative stimulus contexts would result in different structural representations of the emotion space, and if so, would these effects be consistent with the scaling artifact hypothesis or with the contextual similarity hypothesis (i.e., due to differences in how interstimulus similarity was judged by subjects). The typical distinction that psychologists make between anger and fear led us to speculate that the judged similarity between the emotion words *angry* and *afraid* would vary between subjects in the UN and AF contexts. Specifically, Krumhansl's (1978) distance–density model predicts that the emotion words *angry* and *afraid* will be judged as less similar when the stimulus space is dense in the neighborhood of these emotions. Because the emotions that reflect anger and fear were more densely represented

[1] The eight affect locations of Russell (1980, Fig. 1) differ slightly from the octants more recently described by Larsen and Diener (1992) and Watson and Tellegen (1985). The major difference is a shift in locations for some low activation (i.e., low arousal) and low pleasantness emotions. The latter two sets of authors describe low activation with words such as *quiet, still,* and *inactive,* and they place *sleepy* in the unactivated–unpleasant octant. Using the locations described by Larsen and Diener (1992), our core set includes at least one emotion word for seven of the eight octants, but excludes the low-activation octant. Because we do not include emotion words from the low activation octant specified by Larsen and Diener (1992) in any contextual condition, the absence of these words is an experimental constant. Moreover, this experimental constant cannot explain any differences in the results that might emerge between experimental conditions.

in the AF condition relative to the UN condition, the distance–density model predicts that the words *angry* and *afraid* will be judged as less similar in the former context. Moreover, this model demands that such differences be localized in the similarity judgments themselves, rather than solely within the scale values resulting from a MDS procedure.

We also speculated that the increased stimulus density in the first quadrant of the PD condition would produce context effects. Krumhansl's (1978) distance–density model suggests that emotion words such as *aroused* and *astonished* will be judged as less similar when the stimulus space surrounding them is dense. Tversky's (1977) feature model suggests a more specific mechanism whereby the classification rule upon which judgments of similarity are based will be altered by inclusion of stimuli implying sexual arousal. In particular, the words *aroused* and *excited* have connotations relating to both general and sexual levels of arousal. Conversely, the word *astonished* lacks sexual connotations but does reflect general arousal or high activation. The differences in connotative meaning among these core stimuli led us to speculate that the introduction of passion related emotions would cause subjects in the PD condition to develop a classification rule that was partially based on the level of sexual arousal associated with each stimulus. Because the UN stimulus set contained no passion related emotion words, we conjectured that subjects presented with this condition would judge both *aroused–astonished* and *excited–astonished* as more similar than their counterparts in the PD contextual condition.

This same prediction may also be generated from application of the principles involved in social priming. The literature on priming effects in social judgment provides abundant evidence that the interpretation of ambiguous stimuli is likely to be shifted toward a related and recently activated concept (Herr, Sherman, & Fazio, 1983; Martin, 1986; Srull & Wyer, 1980). Assuming that the words *aroused* and *excited* are relatively ambiguous with regard to a sexual or nonsexual interpretation, the PD context may prime the sex-related concept so that these words are then interpreted in line with their sexual connotations.

The speculations developed above were tested in Experiment 1 using an indirect measure of similarity. Specifically, subjects in this experiment were asked to sort the 25 emotion words from a given contextual condition based on the perceived similarity among the words. Unlike the many past studies, the sorting data in this experiment were gathered through a paper and pencil instrument rather than the more traditional index-card administration procedure.

Method

Subjects and Design

Subjects were 90 University of South Carolina undergraduates who received credit toward psychology course requirements. A 3 × 3 factorial design provided the basic structure for the experiment, with the 3 levels of stimulus context (UN, AF, and PD) manipulated between subjects and the number of available sort categories (4, 7, or 11) manipulated within subjects. To control for the order of appearance of stimuli on a page, three random orderings of stimuli were used in each contextual condition. This produced nine between-subject cells (3 stimulus orders nested within each of 3 contextual conditions) in the design. Subjects were randomly assigned to one of nine between-subject conditions. All subjects repeatedly sorted the 25 emotion words associated with their respective conditions on three successive trials. On each successive trial, the number of sort categories provided to subjects was increased from 4 to 7 to 11 categories, respectively.

The order in which stimuli appeared on the page associated with a given trial was obtained using a Latin square technique. Three random sequences of 25 stimuli were generated for each contextual condition. Within each condition, three orders of stimulus appearance were constructed by associating each random sequence with a particular trial. The sequence/trial pairs were such that, within a given contextual condition, each sequence was associated with each trial exactly once.

Procedure

The stimuli and the number of available sorting categories were presented to subjects in a booklet. Each subject received a test booklet consisting of one page of instructions followed by three trial pages. There were nine different packets corresponding to the nine between-subjects cells in the design. The 25 stimuli were listed in a 5 × 5 table at the top of each trial page. A small blank line appeared to the left of each stimulus in the table. The context and order of stimuli within the table varied as specified in the preceding section. The area of the page below the set of emotion words was divided into either 4, 7, or 11 boxes, corresponding to trials 1, 2, and 3, respectively. Beside each box was a box number that uniquely identified all boxes on a page.

Subjects were tested in two large groups. After obtaining a randomly distributed packet, subjects read the instructions and were given an opportunity to ask questions. The instructions requested subjects to read through all 25 words at the top of a trial page and then write each word in one of the boxes provided below the table of words. Subjects were asked to sort the stimuli so that words written in a particular box were more similar to each other than to words written in other boxes. Following the placement of a word into a box, subjects were to record the corresponding box number on the blank line beside the word in the stimulus table. It was thought that this mechanism would help subjects determine which stimuli had been classified at any given point in time.

Results

Scoring of Cooccurrence Data

The data from the sorting tasks were scored in a manner similar to that of Russell (1980). First, three 25 × 25 cooccurrence matrices, one for each trial, were constructed for each subject. The stimuli served as the basis for both columns and rows in these matrices and the elements were either 0 or 1. A particular element was set to 1 if the two stimuli referenced by the given row and column were placed in the same category in the sorting task. Otherwise, the element was set to zero. The cooccurrence matrices for each trial were then weighted according to the number of categories available on that trial. Thus, the matrices corresponding to trials 1, 2, and 3 were multiplied by 4, 7, or 11, respectively. This weighting scheme results in greater similarity for cooccurring stimuli when the number of categories is large (and thus, provides for finer discriminations) than when the number of categories is small (and thus, provides for coarser discriminations). The three weighted cooccurrence matrices were then summed for each subject and the resulting matrix was used as a measure of judged similarity among stimuli.

TABLE 1

STRESS VALUES FOR *n*-DIMENSIONAL MDS SOLUTIONS BY CONTEXT (EXPERIMENT 1)

Contextual condition	Dimension 1	2	3	4	5
UN	.275	<u>.093</u>	.078	.067	.052
AF	.434	.184	<u>.093</u>	.076	.053
PD	.297	.123	<u>.084</u>	.061	.054

Note. Stress values for selected solutions are underlined.

Multidimensional Scaling of Judged Similarity

The judged similarity matrices were averaged within contextual condition for the 67 subjects who provided complete data (i.e., those subjects who categorized all stimuli on each trial and who used all categories on each trial).[2] The average similarity matrices corresponding to the three contextual conditions were scaled separately using the nonmetric multidimensional algorithm outlined by Kruskal (1964a,b) and implemented by the ALSCAL program (SAS Institute, 1983).

Dimensionality. Two major criteria for determining the dimensionality of MDS solutions are change in stress (an index of badness of fit) and interpretability (Shepard, 1974). Table 1 shows the values of stress associated with solutions based on successively higher dimensions ranging from 1 to 5. The stress values for the UN condition were indicative of a two-dimensional solution, whereas those for the AF condition suggested a three-dimensional solution. The stress values corresponding to the PD condition were less clear as to the appropriateness of a two-dimensional versus a three-dimensional solution.

Examination of the configurations associated with solutions for both the PD and AF conditions suggested that a three-dimensional solution was most appropriate, because the third dimension was interpretable.

[2] In all four experiments reported in this article, subjects who failed to follow the instructions of the sorting task or those who failed to provide reliable data in the pairwise comparison task were initially removed from the analyses. To ensure that the removal of these subjects did not bias the findings, the analyses in each experiment were rerun including any nonmissing data provided by these individuals. The conclusions resulting from these additional analyses were, for the most part, the same with the exception of two slight changes which will be appropriately noted.

However, we did not find the third dimension for the UN condition readily interpretable and therefore we retained only two dimensions.

Configurations. The initial configurations associated with each solution were rotated using the Procrustes method outlined by Schonemann and Carroll (1970). This method applies a series of similarity transformations to a configuration (i.e., stretching, shrinking, reflecting, and/or rotating the configuration) such that the resulting configuration has a least-squares fit to a prespecified target. The same target or design matrix was used to rotate solutions from all three contexts and was developed by specifying the approximate locations of core stimuli from Russell's (1980) two-dimensional solution. (The third dimension of the target matrix was simply a vector of zeros.) The locations of core stimuli from a given contextual condition were rotated to the target, and the same similarity transformations that produced a least-squares fit of the core stimuli to the target were then applied to the 12 remaining contextual stimuli in that condition.

The first two dimensions of the configurations associated with each solution are shown in Fig. 2. The scale was arbitrarily set to reflect the maximum and minimum values in the design matrix, 1.0 and -1.0, respectively. The configurations for each contextual condition were strikingly similar to that obtained by Russell (1980). In each configuration, the stimuli tended to fall along the perimeter of an ellipse, although this was less pronounced in the PD condition because of the concentration of stimuli in the upper right quadrant of the configuration. Furthermore, the locations of stimuli are consistent with the vertical axis representing the arousal level associated with the stimuli and the horizontal dimension distinguishing stimuli in terms of their pleasantness.[3]

The first and third dimensions for the AF and PD solutions are shown in Fig. 3. In the AF context, the third dimension primarily serves to distinguish those emotions that represent fear and anxiety from those indicative of anger and rage. In contrast, the third dimension of the PD context appears to differentiate emotions in terms of their sexual versus asexual connotations. Thus, the emotion words *passionate* and *lustful* are distinguished from *amazed* and *astonished* on the third dimension, even though they have similar positions in the two-dimensional space. Consistent with this interpretation, the emotions *angry* and *afraid* (which lack sexual connotations) do not differ much on the third dimension in the PD configuration.

[3] Consistent with previously reported MDS solutions (Russell, 1980), the configurations of Fig. 2 place *sleepy* directly in the low arousal (low activation) octant. This location differs from that reported by Larsen and Diener (1992) and others, which place *sleepy* in the bottom left octant (unactivated, unpleasant octant). One possibility is that the location of words like *sleepy* may depend on the judgment method (self reports versus similarity judgments) or analytic method (factor analysis versus MDS) used by researchers.

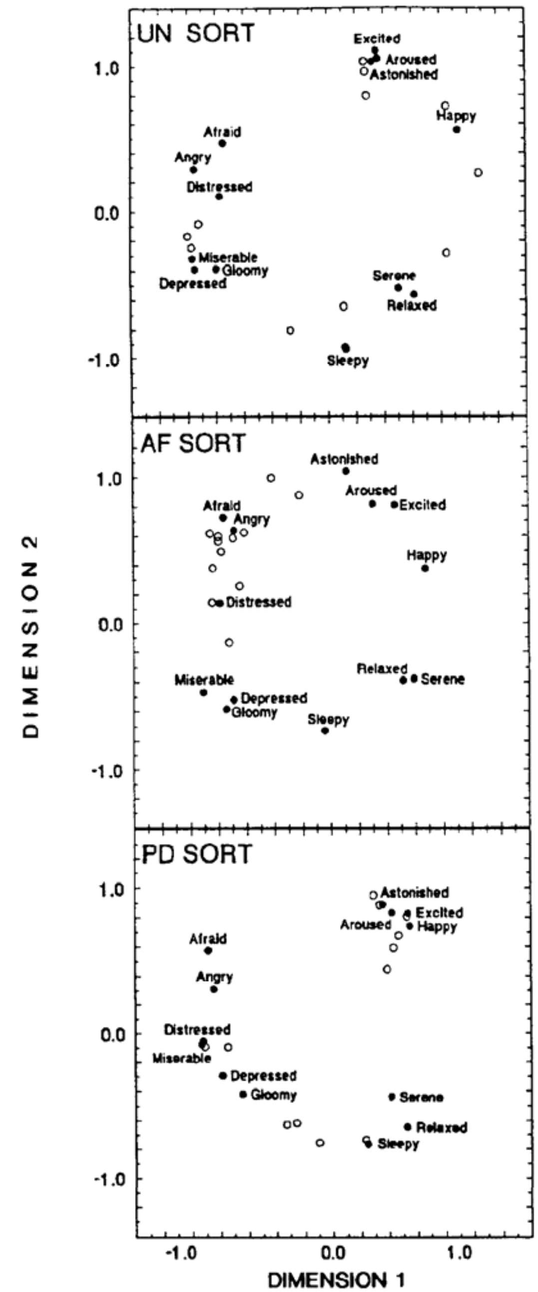

FIG. 2. Locations of core stimuli in the first two dimensions of the MDS space for Experiment 1. A similar circumplex structure is evident in all three contextual conditions.

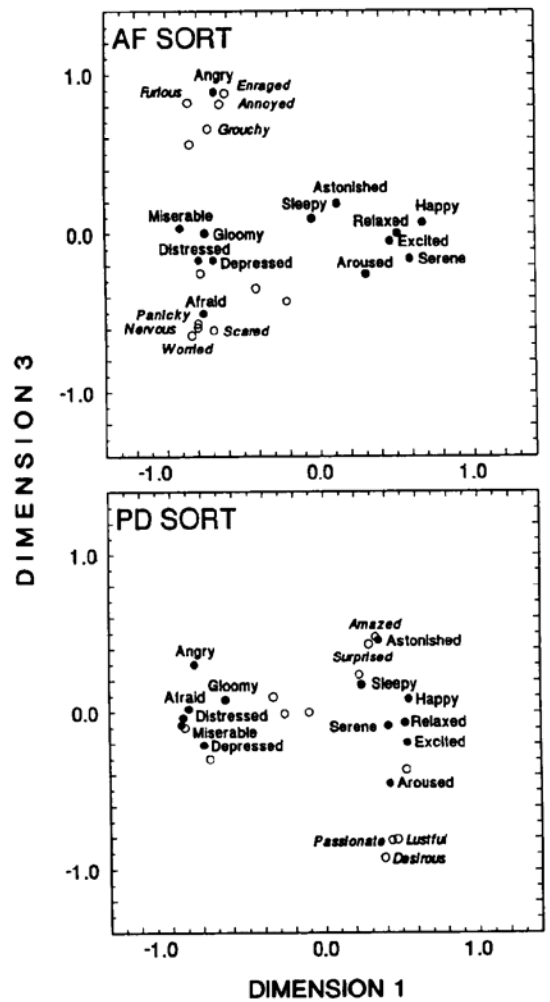

FIG. 3. Locations of core and selected contextual stimuli in the first and third dimensions of the MDS space for AF and PD conditions of Experiment 1.

Modeled distances. To add support to these visual observations made about Figs. 2 and 3, the modeled distances among core stimuli were investigated for the three contextual conditions. Specifically, the distance between each pair of the 13 core stimuli was computed based on the solution associated with each contextual condition. The percentile ranks corresponding to these 78 distances were then computed.

Selected percentile comparisons are given in Table 2. The distance

TABLE 2
PERCENTILE RANKS OF CORE DISTANCES FOR SELECTED EMOTION WORD PAIRS FROM EXPERIMENT 1

Context/ dimensionality	Emotion word pair		
	Angry–Afraid	Aroused–Astonished	Excited–Astonished
UN/2D	11	1	5
UN/3D	12	2	6
AF/3D	51	—	—
PD/3D	—	29	24

between *angry* and *afraid* in the UN condition was ranked at the 11th percentile relative to all pairs of core stimuli distances. In contrast, the corresponding percentile rank was 51 in the AF condition. Thus, relative to modeled distances between other core stimuli, the modeled distance between *angry* and *afraid* was much larger in the AF contextual condition than in the UN condition. The percentile-ranked distance of these words for the three-dimensional UN solution is also presented in Table 2 to demonstrate that the difference in distances between contexts was not due to the fact that two rather than three dimensions were extracted in the UN solution.

The percentile ranks of modeled distances between *aroused* and *astonished* as well as *excited* and *astonished* were also compared for the UN and PD conditions. As shown in Table 2, the difference in the percentile rank of *aroused–astonished* between the UN and PD conditions was substantial. Similar findings emerged with regard to the rank of modeled distance between *excited* and *astonished.* Thus, the relative distances between *aroused* and *astonished* as well as *excited* and *astonished* were larger in the PD condition than in the UN condition. Examination of corresponding percentile ranks for the three-dimensional UN solution again indicates that these differences are not dependent on whether two or three dimensions are extracted for the UN condition.

Analysis of Similarity Judgments

The 300 lower diagonal elements of each subject's similarity matrix were standardized with a mean of zero and a variance of one. The judged similarity between *angry* and *afraid, aroused* and *astonished,* and *excited* and *astonished* were then compared across relevant contextual conditions using those subjects who provided complete data. Each of the three similarity judgments were analyzed separately using an analysis of variance (ANOVA) procedure with contextual condition as a fixed, between-subjects factor and order of appearance (with three appearance orders) as a

TABLE 3
MEAN STANDARDIZED SIMILARITY JUDGMENTS FOR SELECTED COMPARISONS BY CONTEXT

	Context		
Emotion word pair	UN	AF	PD
Angry–Afraid			
Mean	.87	−.26	—
SD	1.48	.43	—
N	20	24	—
Aroused–Astonished			
Mean	.93	—	−.01
SD	.95	—	.92
N	20	—	23
Excited–Astonished			
Mean	1.66	—	.57
SD	1.23	—	1.22
N	20	—	23

random, between-subjects factor nested within context.[4] These analyses revealed significant differences due to context but no reliable differences due to order of appearance. Therefore, to obtain a more sensitive test, order of appearance was dropped from the analysis of each comparison and a simple one-way ANOVA was performed using contextual condition as the sole factor. The main effect of context was highly significant for the *angry–afraid* comparison, $F(1, 42) = 12.70$, $p < .001$, $MS_e = 1.09$. The mean values of the standardized similarity judgments are shown in Table 3, with larger values indicating greater similarity between the given pair of stimuli. Thus, *angry* and *afraid* were judged to be much more similar to each other in the UN condition than in the AF condition. However, the large difference in standard deviations between the two conditions is worth noting. Apparently subjects in the AF condition were much more consistent in their responses than were subjects in the UN condition.

Similar analyses were run to assess contextual differences in judged similarity between *aroused* and *astonished* and between *excited* and *astonished*. In these comparisons, the contrast was made between the UN and PD contexts. As seen in Table 3, subjects in the UN context perceived *aroused* and *astonished* as significantly more similar than those subjects in the PD context, $F(1, 41) = 10.78$, $p < .003$, $MS_e = .878$. This same feeling was also observed for the judged similarity between *excited* and *astonished*, $F(1, 41) = 8.36$, $p < .007$, $MS_e = 1.51$.

[4] The alpha levels for all ANOVAs reported in this article were equal to .05/3 = .01667 since, in each analysis, there were three comparisons under study,.

Discussion

The results of Experiment 1 support the hypothesis that context can dramatically affect the evaluations of similarity among emotion words. The two-dimensional MDS solution for the UN contextual condition replicated Russell's (1980) findings, substituting a paper and pencil sorting task for the traditional task of sorting index cards. Like Russell's solution, the configuration resulting from the UN condition possessed a circumplex quality with emotion words located along the periphery of an ellipse. Our orientation of the two axes reflected the degree of arousal and the level of pleasure associated with the emotion words. When the distribution of contextual stimuli was varied to emphasize anger and fear emotions as in the AF condition, a three-dimensional structure emerged in which the third dimension teased apart differences between anger and fear. The emergence of this third dimension is similar to the MDS findings of Shaver *et al.* (1987). The differences found in the MDS configurations from the UN and AF conditions suggest that the spatial representation of emotions, as reflected by similarity judgments of emotion words, is quite dependent on stimulus context. Moreover, the fact that similar differences between the UN and AF contexts existed in the similarity judgments themselves indicates that such context effects are not simply anomalies of the MDS procedure. Instead, they appear to result from context-dependent processing of stimulus information.

The contextual differences noted between the UN and PD conditions provided additional support for these conclusions. When the stimulus context emphasized two different facets of arousal (sexual and asexual) as in the PD condition, the judged similarity of emotion word pairs *aroused–astonished* and *excited–astonished* was less than when the context reflected a uniform distribution of emotions (in which no words had explicit sexual connotations). Again, because these differences were realized in both the MDS scale values and the similarity judgments themselves, they are more consistent with the contextual similarity hypothesis than the scaling artifact hypothesis.

A number of different theories can explain the observed contextual effects. Krumhansl's (1978) distance–density model predicts greater dissimilarity within dense regions of the psychological space. The emergence of a third dimension under different contextual conditions (i.e,. AF and PD contextual conditions) is consistent with the distance–density model. Specifically, the degree of variation of contextual stimuli along different dimensions of the full *n*-dimensional emotion space may determine which dimensions are utilized by subjects and which are under utilized or collapsed. In the AF condition, increased variation along the third dimension differentiating anger from fear evoked its use by a majority of subjects in the similarity judgment task. Parallel arguments can explain the emergence of the third dimension differentiating passionate from nonpassionate arousal in the PD condition.

Tversky's (1977) feature model provides a similar explanation of these effects. In particular, the set of contextual stimuli might determine the weight of different features of the emotions. In the UN condition, few emotion words shared the features that distinguish anger from fear and so these features received little weight. In the AF condition, many words shared these features and hence the features received greater weight in determining similarity. Again, a parallel argument can be made to account for the difference in similarity judgments in the UN and PD conditions.

The AF and PD conditions produced similar types of effects, but these effects may have resulted from different mechanisms. It is unlikely that subjects thinking about anger and fear in the UN condition failed to recognize how these emotions differed. Instead, they appear to have simply given these differences little weight in judging similarity. Although the increased dissimilarity of word pairs *aroused–astonished* and *excited–astonished* in the PD condition may have also resulted from differential weighting, the priming hypothesis suggests that these differences may have alternatively arisen from a change in the meaning ascribed to the emotion words *aroused* and *excited.* In particular, the context words *desirous, lustful,* and *passionate* may have primed a sexual concept node that was used to interpret the target words *aroused* and *excited.* The present experiment does not distinguish among these theoretical interpretations.

The emergence of a third dimension for two of our contextual sets might be interpreted as empirical support for a three-dimensional emotion space. However, we believe that a closer examination of the results suggests that at least four, and possibly more, dimensions are required to reflect adequately the entire perceptual space of emotion words. This is because the emergent third dimension in the MDS configuration was not the same in the two contexts. In the AF condition, the third dimension strongly distinguished anger from fear emotions but did little to distinguish astonishment from arousal and excitement. The opposite was true in the PD condition in which the third dimension strongly distinguished astonishment from arousal and excitement, but did little to distinguish anger from fear. Moreover, the notion of a "true" dimensionality of the emotion space appears very tenuous given that the effective dimensionality itself is subject to context effects.

EXPERIMENT 2

Experiment 1 demonstrated that measures of the similarity of emotions derived from a sorting task can be strongly dependent on the stimulus context. However, it is possible that this dependency was partially a function of the manner in which the sorting task was conducted. Specifically, subjects sorted emotion words into 4, 7, and 11 categories on three

successive and independent trials. Because only four categories were used on the first trial, no more than three dimensions could be consistently reflected in judgments made under those conditions. Moreover, with only four response categories, it seems intuitively easier to distinguish stimuli along two orthogonal bipolar dimensions than to triangulate them along three dimensions. If subjects indeed based their judgments on just two dimensions during the first trial, this strategy may have persisted to some degree on subsequent trials, especially in the UN condition where only two dimensions were emphasized by the context. Thus, although subjects were in no way constrained to use grouping distinctions from the 4-category sort in the later 7- and 11-category sorts, a tendency to do so in the UN condition could have led to lower dimensionality of the MDS solution.

In addition to constraints on dimensionality, the initial use of only a few categories may have resulted in greater contextual dependence of similarity judgments. Unidimensional judgments on psychophysical and social dimensions are more sensitive to manipulations of the distributional context when the number of rating categories is small rather than large (Parducci & Wedell, 1986; Wedell & Parducci, 1988, Wedell *et al.*, 1987; Wedell *et al.*, 1990). Parallel effects may occur with similarity judgments, with the initial contextual dependency of judgment persisting through subsequent trials.

Given these potential limits on the generality of Experiment 1, a second experiment was conducted to determine whether the use of a relatively few number of categories during the initial similarity sort was responsible for the observed contextual dependence. This experiment was identical to the first with the exception that the number of response categories available on each successive trial was reversed. The number of categories on the first trial was 11, followed by 7 and 4 categories on subsequent trials, respectively. If the context effects found in Experiment 1 were due to the persistence of context dependent response constraints associated with the initial use of only a few categories, then initially sorting emotion words into many categories might reduce or eliminate these effects.

Method

Subjects were 74 University of South Carolina undergraduates who received credit toward psychology course requirements. The design and procedure were identical to that for Experiment 1 with the exception that the trial pages of the test packets were arranged in reverse order. Therefore, on the first trial, subjects sorted the 25 adjectives into 11 categories, followed by 7 categories on the second trial, and 4 categories on the third trial.

Results

Scoring of Cooccurrence Data

Matrices of derived similarity judgments were constructed from the cooccurrence data provided by each subject using the same procedure

TABLE 4
STRESS VALUES FOR *n*-DIMENSIONAL MDS SOLUTIONS BY CONTEXT (EXPERIMENT 2)

Contextual condition	Dimension 1	2	3	4	5
UN	.303	.100	.072	.061	.045
AF	.191	.125	.069	.058	.044
PD	.375	.141	.066	.053	.034

Note. Stress values for selected solutions are underlined.

described in Experiment 1. The derived similarity matrices were averaged separately for subjects in each contextual condition. Only those 52 subjects who provided complete data were used to calculate the average similarity matrices.

Multidimensional Scaling of Judged Similarity

Dimensionality. A separate MDS analysis of the average similarity judgments was performed for each contextual condition using the ALSCAL program. The stress values corresponding to solutions based on 1 to 5 dimensions are given in Table 4. The stress values along with the interpretability of the resulting configurations suggested that three-dimensional solutions were most appropriate for the AF and PD conditions. Once again, both the slight change in stress values and the lack of interpretability for the third dimension led us to select a two-dimensional solution for the UN context.

Configurations. The three configurations appeared very similar to those found in Experiment 1. To assess this similarity, the 300 distances between all possible pairs of stimuli were computed for each context and correlated with the corresponding distances from Experiment 1. The resulting correlations were high (.97 in the UN context, .96 in the AF context, and .95 in the PD context). Given the high concordance between the configurations of this experiment and those shown in Figs. 2 and 3, the configurations for Experiment 2 are not presented here.

Modeled distances. A subsequent investigation of the ranked distances among core stimuli in each configuration is summarized in Table 5. To guard against the possibility of inducing differences by retaining a two-dimensional rather than a three-dimensional solution for the UN context, the ranked distances from both solutions were reviewed, and both are given in Table 5. The percentile ranks found in Experiment 1 are also shown in Table 5 to facilitate comparison.

As in Experiment 1, the ranked distance of the core words *angry* and *afraid* was much greater in the AF condition relative to the UN condition. Similarly, the ranked distance between the words *aroused* and *astonished*

TABLE 5
PERCENTILE RANKS OF DISTANCES FOR SELECTED CORE EMOTION WORD PAIRS IN EXPERIMENTS 1 AND 2

Emotion word pair	Context/ dimensionality	Experiment 1	Experiment 2
Angry–Afraid			
	UN/2D	11	5
	UN/3D	12	2
	AF/3D	51	58
Aroused–Astonished			
	UN/2D	1	6
	UN/3D	2	3
	PD/3D	29	27
Excited–Astonished			
	UN/2D	5	12
	UN/3D	6	13
	PD/3D	24	12

was much greater in the PD context than in the UN context. However, the ranked distance between the words *excited* and *astonished* were essentially the same across the PD and UN conditions. Thus, context effects on modeled distances were replicated for two of the three comparisons.

Analysis of Similarity Judgments

As in Experiment 1, each subject's judgments of the similarity between all possible pairs of emotion words were standardized with a mean of zero and a variance of one. The standardized judgments for the *angry–afraid, aroused–astonished,* and *excited–astonished* pairs were then analyzed separately using three ANOVAs. As in the first experiment, an analysis that included the fixed effect of context and the random effect of appearance order was first conducted for each of the three emotion pairs. Again, the effect of appearance order was not significant either at any level of context or pooled across contexts. Therefore, a more powerful test of context was conducted in which contextual condition served as the sole factor in a one-way ANOVA. This analysis was repeated for each emotion word pair and the mean similarity judgments resulting from each analysis are given in Table 6.

The mean judgments in Table 6 are all in the same direction found in Experiment 1. On average, the subjects in the AF condition judged the *angry–afraid* emotion pair to be substantially less similar than did the subjects in the UN context, $F(1, 34) = 13.68$, $p < .0008$, $MS_e = .660$. Additionally, those subjects in the PD condition perceived the emotion pairs *aroused–astonished* and *excited–astonished* to be less similar than did their counterparts in the UN condition. However, these differences

TABLE 6
MEAN STANDARDIZED SIMILARITY JUDGMENTS FOR SELECTED COMPARISONS BY CONTEXT (EXPERIMENT 2)

	Context		
Emotion word pair	UN	AF	PD
Angry–Afraid			
Mean	.62	−.39	—
SD	1.04	.26	—
N	21	15	—
Aroused–Astonished			
Mean	.94	—	.36
SD	.88	—	1.05
N	21	—	21
Excited–Astonished			
Mean	.85	—	.63
SD	1.10	—	.88
N	21	—	21

were not statistically significant, $F(1, 40) = 3.75$, $p < .06$, $MS_e = .933$ for *aroused–astonished*, and $F(1, 40) = .51$, $p < .48$, $MS_e = .988$ for *excited–astonished*.[5]

Discussion

The above results did not provide much support for the hypothesis that the context effects of Experiment 1 were dependent on initially sorting stimuli into only a few categories. Although the context effects noted in Experiment 1 for the word pair *excited–astonished* were reduced, the overall pattern of context effects from Experiment 1 was replicated. Configurations for PD and AF included a third dimension that distinguished among the contextual stimuli. The judged similarity of *aroused–astonished* decreased and its corresponding model distance increased in the PD condition relative to the UN condition. Moreover, the effects of context on the *angry–afraid* comparison were equal to, if not greater than, those found in Experiment 1. Therefore, the combined results of Experiments 1 and 2 indicate that similarity measures among emotion words derived from a sorting task are strongly dependent on the contextual set of words presented to the subject.

EXPERIMENTS 3 AND 4

Wedell (1990) has suggested that the locus of context effects can be determined somewhat by the degree to which similar effects of context are exhibited across different experimental procedures. A substantial change in the nature of context effects resulting from changes in the procedures used to obtain judgments implies that these effects are localized in the latter stages of the information processing sequence that are concerned with the generation of overt responses. On the other hand, generalization of context effects across a variety of response procedures supports their location at a deeper level of processing resonsible for the cognitive representation of stimuli.

Experiments 3 and 4 were conducted to determine whether the contextual dependence of similarity judgments found in Experiments 1 and 2 using sorting tasks would occur when similarity was assessed using a direct pairwise similarity rating procedure. There are at least two reasons to believe that the sorting task may be more prone to context effects than the rating task. First, the nature of the sorting task necessitates that the set of contextual stimuli is more cognitively available to subjects than in the pairwise task. This is because all stimuli are simultaneously present for the subject to refer to as he or she sorts them into different groups. In the pairwise rating task, typically only the two stimuli currently being judged are present at the time of judgment.

Second, the sorting task should be more susceptible to context effects arising from the frequency principle of Parducci's (1963) range–frequency theory. According to the frequency principle, subjects tend to use each category equally often. Thus, stimuli that are perceived as dissimilar from one another may not be sorted into different groups if the frequency of stimuli in the resulting groups would be too low. For example, although subjects may have wished to distinguish *angry* from *afraid* in the UN condition, these words would have likely been placed into groups of very low frequency (essentially just one stimulus each), thereby violating the frequency principle. Inclusion of many anger- and fear-related words in the AF condition would then facilitate forming separate groups for these emotion categories, because these groups would have many members. These considerations do not arise in pairwise rating.

Experimental research generally supports the distinctions made above. Using schematic faces, Jones and Wedell (1987) reported significant contextual effects when similarity was assessed using a sorting procedure similar to that used in Experiments 1 and 2, but not when similarity was assessed by direct pairwise rating. Corter (1987) found no significant effects of context on pairwise similarity ratings for several sets of stimuli, but he did find some context effects when similarity was calculated from a confusion matrix generated by an identification task (which may be considered analogous to sorting stimuli into unique groups). Finally, Mellers and Birnbaum (1982) have demonstrated that context effects on attribute ratings for psychophysical stimuli do not generalize to a pairwise difference rating procedure.

[5] When the analysis was rerun using the data from all subjects, the effect of context became significant for *aroused–astonished* but remained nonsignificant for *excited–astonished*.

Experiments 3 and 4 replicated the contextual manipulations of Experiments 1 and 2 but used a pairwise similarity rating task rather than a sorting task. In Experiment 3 the words making up each pair were presented simultaneously on the screen, whereas in Experiment 4, a 3-s delay was introduced between the presentation of the first and second words of the pair. The delay manipulation was introduced because Wedell (1991) has shown that context effects on pairwise dissimilarity judgments of unidimensional stimuli are significantly enhanced by introducing a delay between the presentation of elements in a stimulus pair. Wedell has suggested that subjects use the delay period to compare the first stimulus to contextual stimuli that reside in memory, thereby increasing the contextual dependency of subsequent similarity ratings.

Method

Design and Subjects

Subjects were University of South Carolina undergraduates who participated in partial fulfillment of psychology course requirements. There were 71 and 85 subjects assigned to Experiments 3 and 4, respectively. A one-way design was implemented for both experiments in which stimulus context served as a between-subjects factor. The contextual conditions and associated stimuli in both studies were identical to those in Experiments 1 and 2.

Apparatus

Microprocessors (IBM Model PS2 50Z's with VGA color displays) were used to present instructions, display stimuli, and collect responses.

Procedure (Experiment 3)

Up to five subjects participated at a time, with each subject seated at a separate terminal. Subjects were randomly assigned to contextual conditions. Instructions were identical across contextual conditions and generally informed subjects that their task was to judge the similarity between various pairs of emotion words. Judgments of similarity were obtained using a 9-point rating scale.

Stimuli were presented in stages. During the first stage, all subjects in a given context were presented with the same 15 pairs of emotion words which were randomly chosen from the 300 unique pairs ([25 × 24/2] pairs) in that contextual condition. During this preview stage, subjects were instructed to become familiar with the task, but did not rate the stimulus pairs. This same set of 15 pairs was presented during the second stage of the experiment (i.e., the pretest stage). During the pretest stage, subjects viewed the two emotion words in each stimulus pair and then provided a judgment of the similarity between them using the 9-point rating scale. In the next stage of the experiment, the test stage, subjects were presented with all 300 stimulus pairs, and similarity ratings were obtained after each pair was presented. Thus, 15 of the 300 pairs were rated in both pretest and test stages.

In each stage of the experiment, one element of each stimulus pair was presented on the left side of the computer screen while the other was presented on the right side. A 9-point scale was printed at the bottom of the screen with 9 evenly spaced hash marks labeled 1 through 9. The endpoints of the scale, 1 and 9, were anchored with the terms "Very Dissimilar" and "Very Similar," respectively. The ordering of the 300 stimulus pairs presented during the test stage was determined using the method described by Ross (1934), in which each stimulus occurred equally often on the left and right of the screen and the average number of trials between repetitions of a stimulus were maximized. Each subject received a different stimulus order by creating a random sequence of 300 stimulus pairs prior to computing the Ross ordering.

Procedure (Experiment 4)

The procedure of Experiment 4 was the same as that of Experiment 3, with the exception of the timing of stimulus presentation. For each trial, the first word in the pair appeared on the left side of the computer screen for 1 s and was then erased. After a 3-s delay, the second word in the pair appeared on the right side of the screen for 1 s and was then erased. The response scale remained visible to the subject over the entire duration of the trial. Subjects were instructed to respond after the erasure of the second word of the stimulus pair.

TABLE 7

STRESS VALUES FOR *n*-DIMENSIONAL MDS SOLUTIONS BY CONTEXT (EXPERIMENTS 3 AND 4)

Contextual condition/experiment	Dimension 1	2	3	4	5
UN/Exp. 3	.293	.119	<u>.091</u>	.067	.056
UN/Exp. 4	.316	.127	<u>.089</u>	.072	.057
AF/Exp. 3	.360	.198	<u>.103</u>	.087	.065
AF/Exp. 4	.257	.139	<u>.098</u>	.074	.062
PD/Exp. 3	.262	.128	<u>.094</u>	.076	.062
PD/Exp. 4	.370	.188	<u>.109</u>	.074	.064

Note. Stress values for selected solutions are underlined.

Results

Reliability Analysis of Judgments

The dual judgments for the 15 stimuli presented in both the pretest and the test stages were correlated within each subject. The average correlation between judgments was .68 in Experiment 3, and the correlations spanned from −.27 to .94. In Experiment 4, the average correlation between judgments was .69, and correlations ranged between −.25 and .96. Data from the 53 subjects in Experiment 3 and the 57 subjects in Experiment 4 who exhibited correlations greater than or equal to .65 were studied in subsequent analyses.

Multidimensional Scaling of Judged Similarity

The similarity judgments were averaged across subjects within each contextual condition for each experiment, and these were then scaled separately by context using the same MDS and rotational procedures described in Experiment 1.

Dimensionality. The stress values associated with solutions based on 1 through 5 dimensions are given in Table 7. In both experiments, the stress values for the AF and PD conditions indicated that a three-dimensional solution was most appropriate and thus, the appropriate dimensionality of these solutions paralleled those from Experiments 1 and 2. However, unlike the results from the sorting experiments, the stress values for the UN conditions in the pairwise experiments also suggested a three-dimensional solution.

Configurations. The configurations associated with the corresponding three-dimensional solutions from Experiment 3 are shown in Figs. 4 and 5. The configuration for each solution was rotated to the same target matrix used in the previous experiments. The first two dimensions from each solution are shown in Fig. 4. Dimensions 1 and 2 of each configuration are, again, suggestive of Russell's (1980) circumplex pattern with the vertical dimension corresponding to arousal (low to high) and the horizontal dimension indicative of pleasure (unpleasant to pleasant). The stimuli in all three contexts are located on the boundary of this circumplex, although the circumplex is better defined in the UN condition where stimuli are not densely packed in any particular quadrant.

Figure 5 shows the first and third dimensions of the configurations from each context. As in Experiment 1, the third dimension in the AF condition primarily distinguished between anger- and fear-related emotions. However, the separation between these two types of emotions was less pronounced in the current configuration.

Of equal interest was the emergence of a third dimension in the UN condition. Like the AF condition, the third dimension distinguished between emotion words *angry* and *afraid.* There was no hint of an interpretable third dimension for UN conditions in the sorting experiments (1 and 2). Thus the major difference between sorting and rating techniques is in the emergence of this third dimension for UN conditions. The UN configuration from Experiment 4 replicated this same distinction, thereby establishing the stability of this dimension.

The third dimension resulting from the PD condition was also similar to that from Experiments 1 and 2 in that it distinguished between sexual and asexual forms of arousal. However, this distinction was again less pronounced in the current study. The appearance of a third dimension in the UN condition also provided more separation between the *aroused–astonished* and *excited–astonished* emotion word pairs than observed in the previous study. Thus, a visual inspection of the three configurations in Figs. 4 and 5 suggests that the differential context effects found in Experiments 1 and 2 were attenuated when a standard pairwise comparison procedure was used.

The configurations associated with delay condition of Experiment 4 were visually quite similar to those found under no delay conditions. This similarity was confirmed by correlating the 300 distances associated with all stimulus pairs within a given context in Experiment 4 with those from

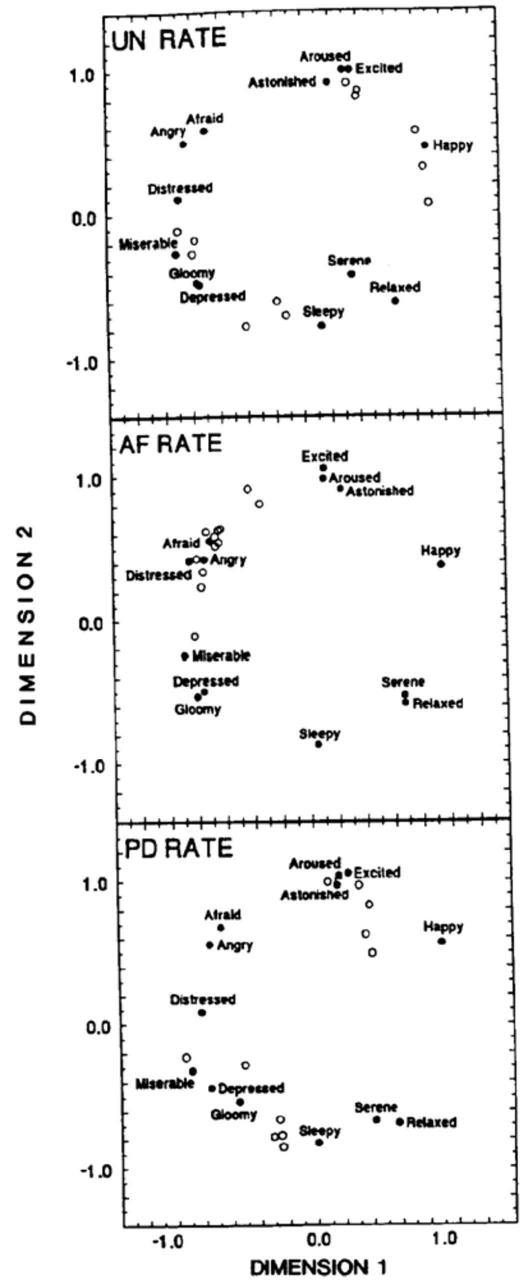

FIG. 4. Locations of core stimuli in the first two dimensions of the MDS space for Experiment 3.

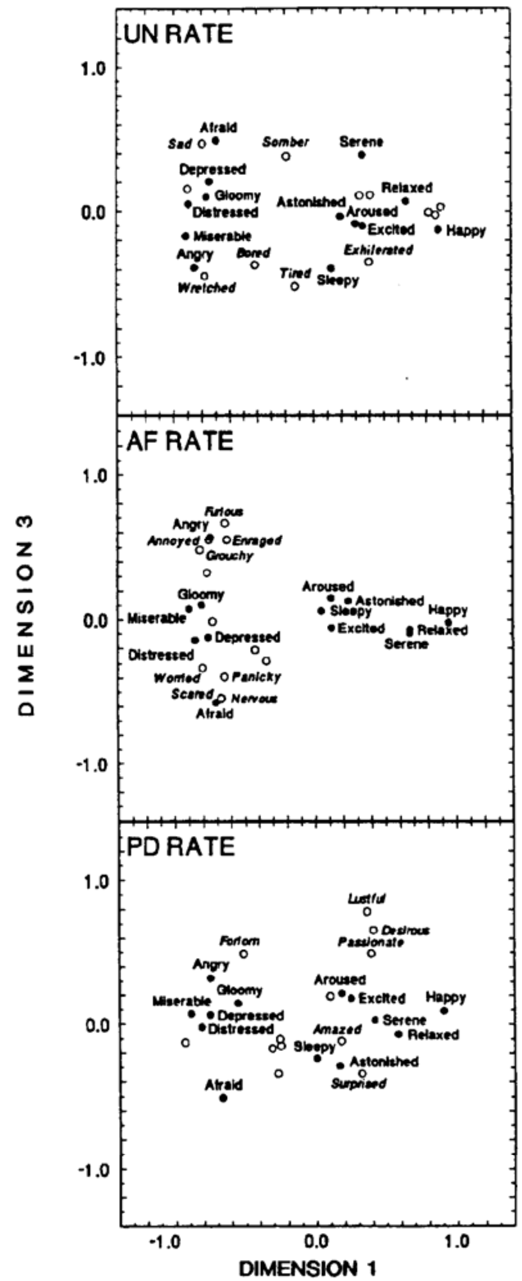

FIG. 5. Locations of core and selected contextual stimuli in the first and third dimensions of the MDS space for Experiment 3.

TABLE 8

PERCENTILE RANKS OF CORE DISTANCES FOR SELECTED EMOTION WORD PAIRS BY EXPERIMENT

Emotion word pair	Context/ dimensionality	Sorting		Rating	
		Exp. 1	Exp. 2	Exp. 3	Exp. 4
Angry–Afraid					
	UN/3D	12	2	13	24
	AF/3D	51	58	30	36
Aroused–Astonished					
	UN/3D	2	3	5	3
	PD/3D	29	27	8	10
Excited–Astonished					
	UN/3D	6	13	8	5
	PD/3D	24	12	10	8

Experiment 3. The resulting correlations were .98, .98, and .97 for the UN, AF, and PD contexts, respectively. Given the close corresepondence of configurations across these two experiments, those for Experiment 4 are not shown here.

Modeled distances. The attenuation of context effects was also supported by an inspection of the ranked distances of core stimuli in each context. These ranks are given in Table 8 and are summarized for all experiments for easy comparison. The ranked distance between the emotion words *angry* and *afraid* in the AF contexts was substantially smaller in the two rating experiments than in the two sorting experiments. In contrast, the distance between the words *angry* and *afraid* increased slightly in the UN conditions of the pairwise comparison experiments relative to the sorting experiments. Thus, the contextual differences in ranked distance for *angry* versus *afraid* were reduced when similarity was assessed via pairwise rating. In the PD context, the ranked distance between *aroused* and *astonished* decreased in both pairwise comparison experiments relative to the sorting experiments. Moreover, the magnitude of this decrease was similar for Experiments 3 and 4. Again, this suggested that the context effects found with the sorting procedure had attenuated in the experiments using a pairwise comparison approach to data collection. The ranked distance between the emotion words *excited* and *astonished* was quite similar in the PD and UN contexts for both pairwise comparison experiments. This finding parallels that for Experiment 2.

Analysis of Similarity Judgments

The similarity judgments made by those subjects with adequate pretest–test correlations were standardized within subject to a mean of zero and

TABLE 9
MEAN STANDARDIZED SIMILARITY JUDGMENTS FOR SELECTED COMPARISONS BY CONTEXT (EXPERIMENTS 3 AND 4)

	Experiment 3 context			Experiment 4 context		
Emotion word pair	UN	AF	PD	UN	AF	PD
Angry–Afraid						
Mean	.03	−.05	—	−.15	−.23	—
SD	.55	.73	—	.68	.76	—
N	20	16	—	23	18	—
Aroused–Astonished						
Mean	.80	—	1.13	.56	—	.89
SD	.52	—	.49	.71	—	.48
N	20	—	17	23	—	24
Excited–Astonished						
Mean	.98	—	1.19	.92	—	1.03
SD	.56	—	.40	.50	—	.53
N	20	—	17	23	—	24

variance of one. The average standardized similarity judgments for *angry–afraid, aroused–astonished,* and *excited–astonished* are shown in Table 9.

Unlike the sorting experiments, subjects in the AF context did not differ significantly from those in the UN condition with regard to their judgments of similarity between the emotion words *angry* and *afraid* in either pairwise rating experiment, F's < 1.0. When comparing subjects from the PD and UN contexts, differences in judged similarity between the words *excited* and *astonished* were also nonsignificant, p's $> .20$. Similarly, context-based differences in the average judged similarity of *aroused* and *astonished* also failed to achieve significance, although there was a tendency in both pairwise experiments to judge these words as more similar in the PD condition, $F(1, 35) = 3.80, p < .06, MS_e = .255$ for Experiment 3 and $F(1, 45) = 3.39, p < .07, MS_e = .367$ for Experiment 4.[6]

Discussion

The context effects that were evident in the sorting experiments diminished substantially when judgments were gathered using a pairwise rating procedure. This attenuation was not only visible in the MDS configurations but was also confirmed in the similarity judgments themselves. The procedural dependence of these context effects suggests that the

[6] When the data from all subjects in Experiment 3 were reanalyzed, the trend initially found for the *aroused–astonished* comparison attenuated.

underlying cognitive representation of these emotions may be rather stable. The greater contextual sensitivity of the sorting procedure may result from the greater salience of contextual stimuli for that procedure. Furthermore, tendencies to equalize category usage may contribute to the contextual dependence of the sorting procedure. The issue of procedural sensitivity of these types of contextual effects is not addressed directly in either Tversky's (1977) or Krumhansl's (1978) theory.

The procedure-specific attenuation of context effects may be related to the more general finding of preference reversals in simple choice situations. For example, Slovic and Lichtenstein (1983) have reviewed many studies in which changes in the response mechanism available to subjects in a dual-alternative preference task (i.e., binary choice or alternative matching) actually changed the preferences for the objects in question. Several researchers have explained these preference reversals as changes in judgment strategy that are induced by differences in the task demands associated with alternative experimental procedures (Tversky, 1969; Tversky, Sattath, & Slovic, 1988). A similar argument could be constructed to explain the divergent findings of preceding experiments. Specifically, the number of dimensions subjects evaluate when making similarity judgments may be an interactive function of the variability of the stimuli along a given dimension and the response required in the task. Under this hypothesis, the subjects in a sorting task distinguish stimuli along only those dimensions that have some moderate degree of variation. On the other hand, in a pairwise comparison task, subjects may consider relevant dimensions with less regard to the variability of stimuli along those dimensions (a more compensatory approach).

Regardless of what explanation is offered to explain the findings of the sorting and pairwise comparison experiment, one conclusion seems quite clear. Similarity judgments derived from the implementation of a sorting procedure appear to be more susceptible to context effects than direct measures of similarity obtained from the pairwise rating procedure. These results did not depend on whether a delay was imposed in the pairwise procedure. Thus, Wedell's (1991) finding of significantly greater contextual density effects on pairwise similarity judgments with the introduction of a delay does not appear to generalize to the emotion words used in this study. This lack of generalization may be due to several factors. First, the stimuli used in Wedell (1991) were unidimensional rather than multidimensional. Second, his psychophysical stimuli lacked prior associative value. In contrast, the emotion words comprising our three contextual conditions were entangled with associative links that have developed over the verbal histories of our subjects. Finally, the nature of the density effects was somewhat different, with density increasing distance but not dimensionality.

SUMMARY

The experiments we have reported provide evidence that similarity judgments may vary due to differences in either the contextual stimuli or the data collection procedures used in a given experiment. These differences in judged similarity resulted in systematic differences among the corresponding MDS solutions. Specifically, manipulation of context affected whether the data were best explained by two or three dimensions and also affected the function of the third dimension when it was produced. These differences in MDS solutions were reliable, in that all solutions had acceptable stress levels, were interpretable, and were replicable within a given context and data collection method. These three qualities further suggest that an MDS representation provides a reasonable basis for understanding the structure of the emotion space. Although other researchers argue that emotions are best represented by a hierarchical category structure (Shaver *et al.*, 1987), the spatial representations of emotion provided here and in other studies (e.g., Russell, 1980; Russell & Bullock, 1985; Watson & Tellegen, 1985) also provide a coherent framework for understanding relations among different emotions.

The context effects underlying our results provide a possible explanation for the different perceptual spaces reported in past literature on similarity judgments of emotion words (Averill, 1975; Bush, 1973; Russell, 1980; Shaver *et al.*, 1987). Our results are clearly more consistent with a three-dimensional representation of emotions. However, these results pose a challenge to a simple three-dimensional representation of emotions because the nature of the third dimension was not always consistent. That is, the distinction between anger and fear may be primarily along a different dimension than the distinction between lust and astonishment. Although the general dimension of potency may capture some of these differences, it does not appear to capture them fully. Thus, our results suggest that at least four dimensions are needed to fully represent the emotion space. Moreover, had other contexts been considered, then perhaps other dimensions would have emerged. Our intention was not to ascertain the "true" dimensionality of the emotion space, but instead to show that stimulus context can influence the apparent dimensionality of that space. In short, although MDS may provide a reasonable basis for the study of the emotion space, the assessment of its dimensionality is highly dependent on the particular stimuli used by the researcher.

The contextual differences in the MDS solutions reported above may seem much like the structural differences often observed with factor analytic solutions when the variable set is expanded. Thus, some readers may equate the notions of contextual independence and factorial invariance. For example, one might view the first two dimensions of each MDS solution (i.e., arousal and pleasantness) as "higher order" dimensions that

consistently arise because of the basic structure inherent in the judgments of emotions. Additionally, the emergence of other dimensions that more finely discriminate among emotions might be thought of as "lower order" dimensions which arise due to the addition of emotion stimuli that reflect a particular distinction. The constructs of contextual independence and factorial invariance are, however, quite different and should not be equated. In these experiments, the notion of factorial lability corresponds to the scaling artifact hypothesis in which certain dimensions arise simply due to the expanded number of emotion stimuli that reflect those dimensions. According to this hypothesis, the appearance of alternative lower order dimensions is a function of the different contextual stimuli in which core stimuli are embedded rather than differences in the judgments of core stimuli themselves. In contrast, our results support the contextual similarity hypothesis in that the differences among MDS structures from the sorting experiments were linked to differences in the similarity judgments among core stimuli. Moreover, these differences in similarity judgments were a function of context and were relatively stable within Experiments 1 and 2. The notion of factorial lability (i.e., the scaling artifact hypothesis) cannot explain these results.

In general, our results suggest that similarity judgments obtained indirectly through a sorting task are more susceptible to context effects than are direct similarity judgments obtained from a pairwise rating procedure. Therefore, we advocate the use of sorting procedures only in those instances where the range of the stimulus domain and the distribution of stimuli within that range are adequately represented in the stimuli sampled for study. Naturally, it is sometimes difficult to determine the adequacy of the stimulus representation. If one is committed to the use of a sorting procedure, then we minimally recommend exploring whether manipulations of the contextual set of stimuli produce different stimulus representations. However, we suggest that researchers utilize a pairwise rating procedure to obtain similarity judgments whenever the situation permits.

REFERENCES

Averill, J. R. (1975). A semantic atlas of emotional concepts. *JSAS Catalog of Selected Documents in Psychology*, **5**, 330. (MS. 421)

Birnbaum, M. H. (1974). Using contextual effects to derive psychophysical scales. *Perception & Psychophysics*, **15**, 89–96.

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

Corterl, J. E. (1987). Similarity, confusability, and the density hypothesis. *Journal of Experimental Psychology: General*, **116**, 238–249.

Daly, E. M., Lancee, W. J., & Polivy, J. (1983). A conical model for the taxonomy of emotional experience. *Journal of Personality and Social Psychology*, **45**, 443–457.

Ekman, L. J., & Friesen, W. V. (1975). *Unmasking the face*. Englewood Cliffs, NJ: Prentice–Hall.

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

Frijda, N. H. (1986). *The emotions.* Cambridge: Cambridge Univ. Press.

Frijda, N. H., Kuipers, P., & ter Schure, E. (1989). Relations among emotion, appraisal, and emotion action readiness. *Journal of Personality and Social Psychology*, **57**, 212–228.

Gati, I., & Tversky, A. (1982). Representations of qualitative and quantitative dimensions. *Journal of Experimental Psychology: Human Perception and Performance*, **8**, 235–240.

Gati, I., & Tversky, A. (1984). Weighting common and distinctive features in perceptual and conceptual judgments. *Cognitive Psychology*, **16**, 341–370.

Gehm, T. L., & Scherer, K. R. (1988). Factors determining the dimensions of the subjective emotional space. In K. R. Scherer (Eds.) *Facets of emotion.* Hillsdale, NJ: Erlbaum.

Herr, P. M., Sherman, S. J., & Fazio, R. H. (1983). On the consequences of priming: Assimilation and contrast effects. *Journal of Experimental Social Psychology*, **19**, 323–340.

Izard, C. E. (1972). *Patterns of emotions.* New York: Academic Press.

Jones, L. E., & Wedell, D. H. (1987, July). *Contextual effects in multidimensional scaling: Variations in the stimulus densities.* Paper presented at the Fifth Annual European Meeting of the Psychometric Society, Enchede, Netherlands.

Krumhansl, C. L. (1978). Concerning the applicability of geometric models to similarity data: The interrelationship between similarity and spatial density. *Psychological Review*, **85**, 445–463.

Kruskal, J. B. (1964a). Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika*, **29**, 1–27.

Kruskal, J. B. (1964b). Nonmetric multidimensional scaling: A numerical method. *Psychometrika*, **29**, 115–129.

Larsen, R. J. (1989). Personality as an affect dispositional system. In L. A. Clark & D. Watson (Chair), *Emotional Bases of Personality.* Symposium conducted at the meeting of the American Psychological Association, New Orleans.

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

Martin, L. L. (1986). Set/reset: Use and disuse of concepts in impression formation. *Journal of Personality and Social Psychology*, **51**, 493–504.

Mellers, B. A., & Birnbaum, M. H. (1982). Loci of contextual effects in judgment. *Journal of Experimental Psychology: Human Perception and Performance*, **8**, 582–601.

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.

Osgood, C. E. (1966). Dimensionality of the semantic space for communication via facial expressions. *Scandinavian Journal of Psychology*, **7**, 1–30.

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

Parducci, A. (1963). Range-frequency compromise in judgment. *Psychological Monographs*, **77**, (92, Whole No. 565).

Parducci, A., & Wedell, D. H. (1986). The category effect with rating scales: Number of categories, number of stimuli, and method of presentation. *Journal of Experimental Psychology: Human Perception and Performance*, **12**, 496–516.

Ross, R. T. (1934). Optimum orders for the presentation of pairs in the method of paired comparisons. *Journal of Educational Psychology*, **25**, 375–382.

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). San Diego: Academic Press.

Russell, J. A., & M. Bullock (1985). Multidimensional scaling of emotional facial expressions: Similarity from preschoolers to adults. *Journal of Personality and Social Psychology*, **48**, 1290–1298.

Russell, J. A., & Fehr, B. (1987). Relativity of perception of emotion in facial expressions. *Journal of Experimental Psychology: General*, **116**, 223–237.

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

SAS Institute (1983). *SUGI supplemental library user's guide.* Cary, NC: SAS Institute, Inc.

Schlosberg, H. (1952). The description of facial expressions in terms of two dimensions. *Journal of Experimental Psychology*, **44**, 229–237.

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

Schonemann, P. H., & Carroll, R. M. (1970). Fitting one matrix to another under choice of a central dilation and a rigid motion. *Psychometrika*, **35**, 245–255.

Shaver, P., Schwartz, J., Kirson, D., & O'Connor, C. (1987). Emotion knowledge: Further exploration of a prototype approach. *Journal of Personality and Social Psychology*, **52**, 1061–1086.

Shepard, R. N. (1962a). The analysis of proximities: Multidimensional scaling with an unknown distance function. I. *Psychometrika*, **27**, 125–140.

Shepard, R. N. (1962b). The analysis of proximities: Multidimensional scaling with an unknown distance function. II. *Psychometrika*, **27**, 219–246.

Shepard, R. N. (1974). Representation of structure in similarity data: Problems and prospects. *Psychometrika*, **39**, 373–421.

Slovic, P., & Lichtenstein, S. (1983). Preference reversals: A broader perspective. *American Economic Review*, **73**, 596–605.

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

Srull, T. K., & Wyer, R. S. (1980). Category accessibility and social perception: Some implications for the study of person memory and interpersonal judgment. *Journal of Personality and Social Psychology*, **38**, 841–856.

Torgerson, W. (1958). *Theory and methods of scaling.* New York: Wiley.

Tversky, A. (1969). Intransitivity of preferences. *Psychological Review*, **76**, 31–48.

Tversky, A. (1977). Features of similarity. *Psychological Review*, **84**, 327–352.

Tversky, A., & Gati, I. (1982). Similarity, separability and the triangle inequality. *Psychological Review*, **89**, 123–154.

Tversky, A., Sattath, S., & Slovic, P. (1988). Contingent weighting in judgment and choice. *Psychological Review*, **95**, 371–384.

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

Wedell, D. H. (1990). Methods for determining the locus of context effects in judgment. In J. P. Caverni, J. M. Fabre & M. Gonzalez (Eds.), *Cognitive Biases* (pp. 285–302). New York: Elsevier Science.

Wedell, D. H. (1991). *A constructive–associative model of the contextual dependence of unidimensional similarity.* Unpublished manuscript.

Wedell, D. H., & Parducci, A. (1988). The category effect in social judgment: Experimental ratings of happiness. *Journal of Personality and Social Psychology*, **58**, 319–329.

Wedell, D. H., Parducci, A., & Geiselman, R. E. (1987). A formal analysis of ratings of physical attractiveness: Successive contrast and simultaneous assimilation. *Journal of Experimental Social Psychology*, **23**, 230–249.

Wedell, D. H., Parducci, A., & Lane, M. (1990). Reducing the dependence of clinical judgment on the immediate context: Effects of number of categories and type of anchors. *Journal of Personality and Social Psychology,* **58,** 319–329.

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