Developmental Psychology
1986, Vol. 22, No. 1, 97-102

0012-1649/86/$00.75

On the Dimensions Preschoolers Use to Interpret Facial Expressions of Emotion

James A. Russell and Merry Bullock
University of British Columbia, Vancouver, British Columbia, Canada

Dimensions of the meaning attributed to facial expressions of emotion were studied in preschoolers (nineteen 4-year-olds, twenty-one 3-year-olds, and thirty-eight 2-year-olds) plus thirty adults. Subjects indicated the similarity or dissimilarity between different emotions by placing photographs of emotional facial expressions into preordained numbers of groups. For each age group, multidimensional scaling of the pairwise similarities yielded a two-dimensional structure in which the expressions fell in a roughly similar circular order. Its dimensions could be interpreted as degree of pleasure and degree of arousal. Four-year-olds and adults also produced a third dimension, tentatively interpreted as assertiveness versus taken aback.

As adults, we are highly skilled at reading facial expressions of emotion. We can interpret expressions in terms of such basic categories as anger, fear, happiness, surprise, and the like. We can also interpret expressions in terms of such basic bipolar dimensions as pleasure–displeasure and arousal–sleepiness. But in what terms do children at various ages interpret facial expressions? Does the message derived from a particular expression vary with age? Knowing how children of different ages interpret facial expressions should give us an important clue about the development of the skills involved in the interpretation of emotion in general. In this article, we examine one aspect of this issue: whether 2-, 3-, and 4-year-olds interpret facial expressions in terms of dimensions of pleasure and arousal.

Degree of pleasure and degree of arousal are clearly the major dimensions (although not the only dimensions) underlying the way in which adults interpret emotions—their own and those of others. Research from several domains supports this conclusion (see Dittman, 1972). One source of evidence is multidimensional-scaling studies of the similarity perceived between emotions expressed in the face (Abelson & Sermat, 1962; Royal & Hays, 1959; Russell & Bullock, 1985; Shepard, 1962) or the voice (Cliff & Young, 1968). Another is multidimensional-scaling studies of similarity judgments about the meaning of words denoting emotions (Bush, 1973; Neufeld, 1975, 1976; Russell, 1978, 1980). A third source is factor-analytic studies of self-reports of emotion (Russell, 1979, 1980; Russell & Steiger, 1982). More important, the same two dimensions can be found in other languages and cultures (Russell, 1983). The present study was guided by a model that summarizes this body of literature by placing feeling states in a circular order around the perimeter of a two-dimensional space, the axes of which are pleasure–displeasure and arousal–sleepiness (Russell, 1980). An example of evidence fitting this model is shown in the top graph of Figure 1, in which each label refers to a facial expression posed to simulate that emotion.

Little research is available concerning what, or even if, dimensions are involved in children's interpretation of emotion. One study examined the meaning of emotion-related words and the dimensions of self-reported mood in school-aged children, from third through seventh grades (Russell & Ridgeway, 1983). The pleasure and arousal dimensions were clearly available to even the youngest children studied, and emotions fell in roughly the same circular order seen in Figure 1.

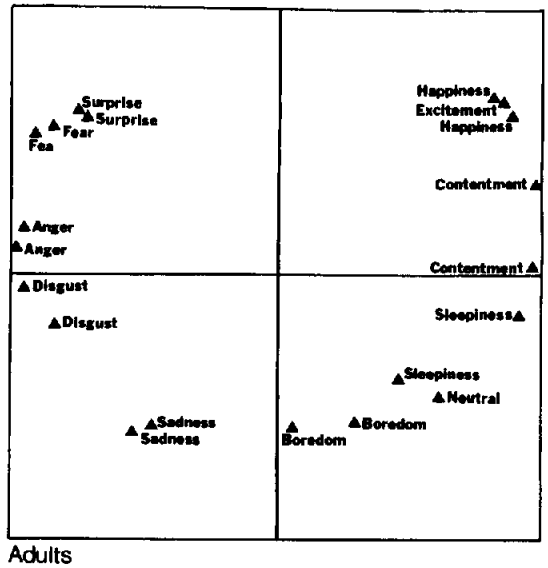
A recent study (Russell & Bullock, 1985) looked at the possibility of dimensions in 4- and 5-year-olds, the youngest sample studied to date concerning this question. The method employed was identical to that used with adults to produce the top graph in Figure 1: Each child was asked to examine 20 facial expressions of emotion. There were two prototypical expressions for each of Ekman and Friesen's (1976) proposed six "basic" emotions: fear, anger, surprise, happiness, sadness, and disgust (plus a neutral expression). Seven other expressions were included to allow the emergence of the complete circular arrangement found in adults—there were one or two expressions each for excitement, calm, sleepiness, and boredom. To indicate the degree of similarity between feelings expressed in each face, the child placed photographs of people who "felt more alike" into a preordained number of groups. The data were analyzed with a multidimensional scaling procedure, and plots of the results are shown in the bottom two graphs of Figure 1. Although the grouping procedure provided only a rough estimate of perceived similarity, the results suggested that the pleasure and arousal dimensions underlie the patterning of even the youngest children's groupings. Correlations with independently derived ratings of the pleasure and arousal conveyed by each facial expression supported this interpretation. The present study built directly on these results.

The present study used a simpler procedure than that used in the Russell and Bullock (1985) study, so that data could be gathered from 2- and 3- as well as 4-year-olds. The children were carefully guided through a procedure whereby they could judge the similarity between the feelings conveyed by 10 different facial

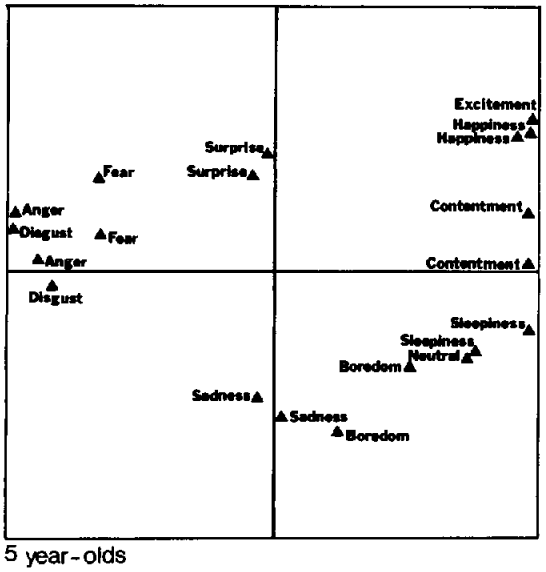
This study was supported by Grant 410-83-0054 from the Social Sciences and Humanities Research Council of Canada.

We thank Lisa Woudzia for serving as the experimenter in this study, and Liz McCririck for her patient help in preparing the manuscript.

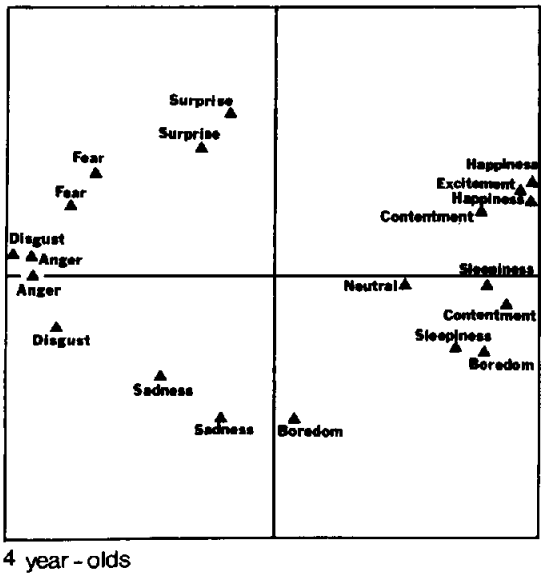
Correspondence concerning this article should be addressed to James A. Russell, Department of Psychology, University of British Columbia, Vancouver, B.C., Canada V6T 1W5.



Adults



5 year-olds



4 year-olds

expressions. Their judgments were analyzed through multidimensional scaling. Emergence of a structure similar to that seen in Figure 1 would provide evidence that young preschoolers have available to them the dimensions of pleasure and arousal by which similarity can be judged.

Method

Subjects

Subjects were thirty-eight 2-year-olds, twenty-one 3-year-olds, nineteen 4-year-olds, and thirty adults. Children were recruited from various day-care facilities serving neighborhoods of different socioeconomic backgrounds in or around Vancouver, Canada. All were native speakers of English, not because this study relied on emotion words but because it relied on verbal instructions. Adults were volunteer undergraduates at the University of British Columbia.

Facial Expressions

To simplify the task as much as possible, we chose only 10 photographs of different facial expressions of emotion. All were posed by women. We selected 6 from Ekman and Friesen's (1976) collection of prototypical expressions of pure "basic" emotions, 1 each for happiness, surprise, fear, anger, disgust, and sadness. Another 4 expressions were added to complete our structural model. For convenience, we labeled these 4 as excitement, calm, sleepiness, and boredom. (These were the states the actresses had been asked to pose.) But neither our labels nor those from Ekman and Friesen played any role in this study, and the subjects were given no labels for the 10 facial expressions.

Procedure

The procedure for measuring perceived similarity derives from Ward's (1977) method of multiple category sorts, by which subjects are asked to group stimuli. Pilot studies led us to divide Ward's procedure into four distinct phases, so that 2- and 3-year-olds could follow the instructions.

A female experimenter familiarized herself with the child and then began the experiment with a game of colors. Each aspect of the color game paralleled the steps in the experiment proper, with the exception that feedback was provided for the color trials. If a child made a clearly inappropriate choice with colors, the experimenter said something like "I don't think so. I think it goes in this pile." Relevant parts of the color game were intermixed with the experiment proper so that the child had mastered the instructions with colors immediately before performing the same task with emotional expressions.

To simplify the task for children, the sorting procedure was divided into four phases. Phase 1 always came first and served to provide the first photograph for each group in Phases 2 and 3. The next three phases gathered the actual similarity data and were in random order. Each phase was carried out on a separate day to maintain the interest of the children; exceptions were made for a few children who insisted on continuing, however.

The procedure for adults was similar to that for children, except that the color game was omitted.

Figure 1. Twenty facial expressions of emotion in a two-dimensional space. Separate solutions for adults, 5-year-olds, and 4-year-olds are shown. (From "Multidimensional Scaling of Emotional Facial Expressions: Similarity From Preschoolers to Adults" by James A. Russell and Merry Bullock, *Journal of Personality and Social Psychology, 48*, p. 1296, Figure 5. Copyright 1985 by the American Psychological Association. Reprinted by permission of publisher and author.)

Phase 1. The purpose of this phase was to choose two, and then three, pictures perceived as different from each other. The experimenter began with the relevant part of the color game. Because the procedure for the color game matched that for the facial expressions, we describe here only the procedure with the latter. After the color game was mastered, the experimenter introduced a deck of cards showing the 10 facial expressions. The child was shown the cards a card at a time, and the actresses were referred to as "mommies." All faces were then displayed in random order on the table. The child was asked to pick a card—any card. That one was placed on the table in front of the child. Referring back to the other 9, the experimenter said, "Look at how different all the mommies feel." The child looked them over and was then asked to pick the one "mommy who feels the most *different* from this mommy" [pointing to the one already chosen]. Once the child had chosen, all 10 cards were gathered and shuffled. The procedure was then repeated with the child picking three expressions as "most different."

The 3- and 4-year-olds appeared to do reasonably well at this task, but 2-year-olds tended to look bewildered. The experimenter maintained a nonevaluative tone, telling the child it was just a game and to pick whatever one he or she wanted to. Any response was praised. Because the 2-year-olds appeared to have difficulty with this phase, we increased the sample size of 2-year-olds to compensate for the error variance introduced here. By outward appearances, the 2-year-olds were not bewildered by the instructions for the next three phases.

Phase 2. The purpose of this phase was to have the child divide the pictures into *two* groups based on similarity. After the appropriate game with colors, the experimenter placed in front of the child the 2 photographs he or she had chosen as different in Phase 1. The child was told that "we're going to make piles of all the mommies who feel alike." The experimenter shuffled the cards with the remaining 8 faces and picked one at random, saying "Look at how this mommy feels. Does she feel most like this mommy [holding the photograph over one of the already chosen photographs]? Or, does she feel most like this mommy [holding it over the other]?" The photograph was added to the chosen group, and the experimenter repeated this procedure for each of the remaining photos until all had been sorted into one of the two groups. There was no restriction on the number that could go into a group. All photos already placed in a group remained visible throughout.

Phase 3. The purpose of this phase was to have the child divide the photographs into *three* groups based on similarity. After the appropriate game with colors, the experimenter placed in front of the child the 3 photographs chosen as different in Phase 1. The remaining procedure was identical to that for Phase 2, except that the child placed each photograph into one of three groups.

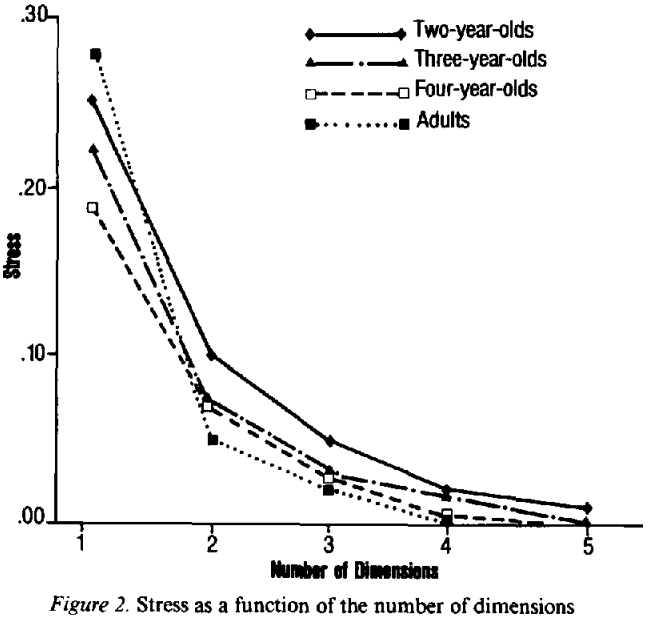


Figure 2. Stress as a function of the number of dimensions in multidimensional-scaling solutions.

Phase 4. The purpose of this phase was to have the child divide the photographs into *five* groups based on similarity. After the appropriate game with colors, the experimenter held all 10 photographs in her hand and asked the child to pick a card. The card chosen was placed in front of the child, who then was asked to find among the remaining 9 the "mommy who feels most like this mommy." The chosen pair was then set aside, and this procedure was repeated with the remaining 8 photographs, then with 6, and so on, until five pairs had been formed.

Similarity. Our measure of similarity for each pair of expressions was the sum of similarity scores derived from the last three phases. For each phase, similarity for a pair was scored 0 if the two expressions were placed in different groups and scored a positive integer if they were placed in the same group. The integer was either 2, 3, or 5, depending on the number of alternatives available during that phase.[1] For example, 2 photographs placed in the same group in Phase 2 (the sort into two groups), but in different groups in Phases 3 and 4, received a score of 2 + 0 + 0 = 2. Maximum similarity was 2 + 3 + 5 = 10. This procedure produced one pairwise similarity matrix of the 10 expressions for each subject. A mean score across subjects of each age group constituted a final similarity matrix.

Multidimensional Scaling

Multidimensional scaling is a technique too rarely exploited with young children. From relatively simple behaviors (such as asking children to put pictures in groups), one can infer aspects of the cognitive processes that determine those behaviors. Multidimensional scaling provides a geometric representation of the similarities the subjects perceive between the stimuli (Shepard, 1962). In this case, the facial expressions of emotion are placed in a "space" such that the ones perceived as more similar are placed closer together; ones perceived as more dissimilar are farther apart.

A matrix of the perceived similarities between all possible pairs is submitted to the multidimensional-scaling program. In this case, each matrix was analyzed by the Guttman-Lingoes (Lingoes, 1965, 1973) smallest space program, SSA-1. The program creates a one-dimensional "space" in which the stimuli are located in such a way that similarity is represented by closeness in the space. By examining the location of the stimuli along the dimension, the dimension can be interpreted. Next, a two-dimensional space is created, then a three-dimensional space, and so on. At each step, the fit between the input data and the space is assessed, often with a measure of poorness of fit called *stress*. The optimal solution is chosen on the basis of this assessment plus the interpretability of the dimensions.

The appropriate number of dimensions in the multidimensional space can be estimated by examining the change in stress as a function of the number of dimensions; these values are shown in Figure 2. In all four age groups, one-dimensional solutions resulted in large values of stress, indicating that a one-dimensional solution was insufficient. Stress was substantially reduced by adding a second dimension. Because relatively small improvements in stress were obtained with each additional dimension beyond the second or third, these improvements were likely due to fitting random error. It thus appears from Figure 2 that either a two- or three-dimensional solution would be warranted, depending on the interpretability of the dimensions.

[1] A separate analysis, in which we set the integer at one for all three phases, produced results highly similar to those reported in this article. The rationale for the weighting procedure used here is described by Ward (1977); the weights help capture the greater similarity implied between two pictures put in the same group when more groups are available.

Results

Separate two-dimensional solutions for each age group are shown in Figure 3. The overall similarity across ages is apparent, despite small discrepancies and irrelevant rotational differences (irrelevant because the "space" could be rotated clockwise or counterclockwise without altering the fit between the data and the solution). Even 2-year-olds produced the predicted results of a circular ordering and pleasure and arousal dimensions. A quantitative test of the apparent convergence is given in Table 1. Because of rotational differences among the solutions, Table 1 reports multiple correlations between a specified single dimension and each two-dimensional solution. For example, the first dimension extracted from the 3-year-olds' data achieved a multiple correlation of .99 with the two-dimensional solution of the 2-year-olds. One way of interpreting this figure is that .99 is the

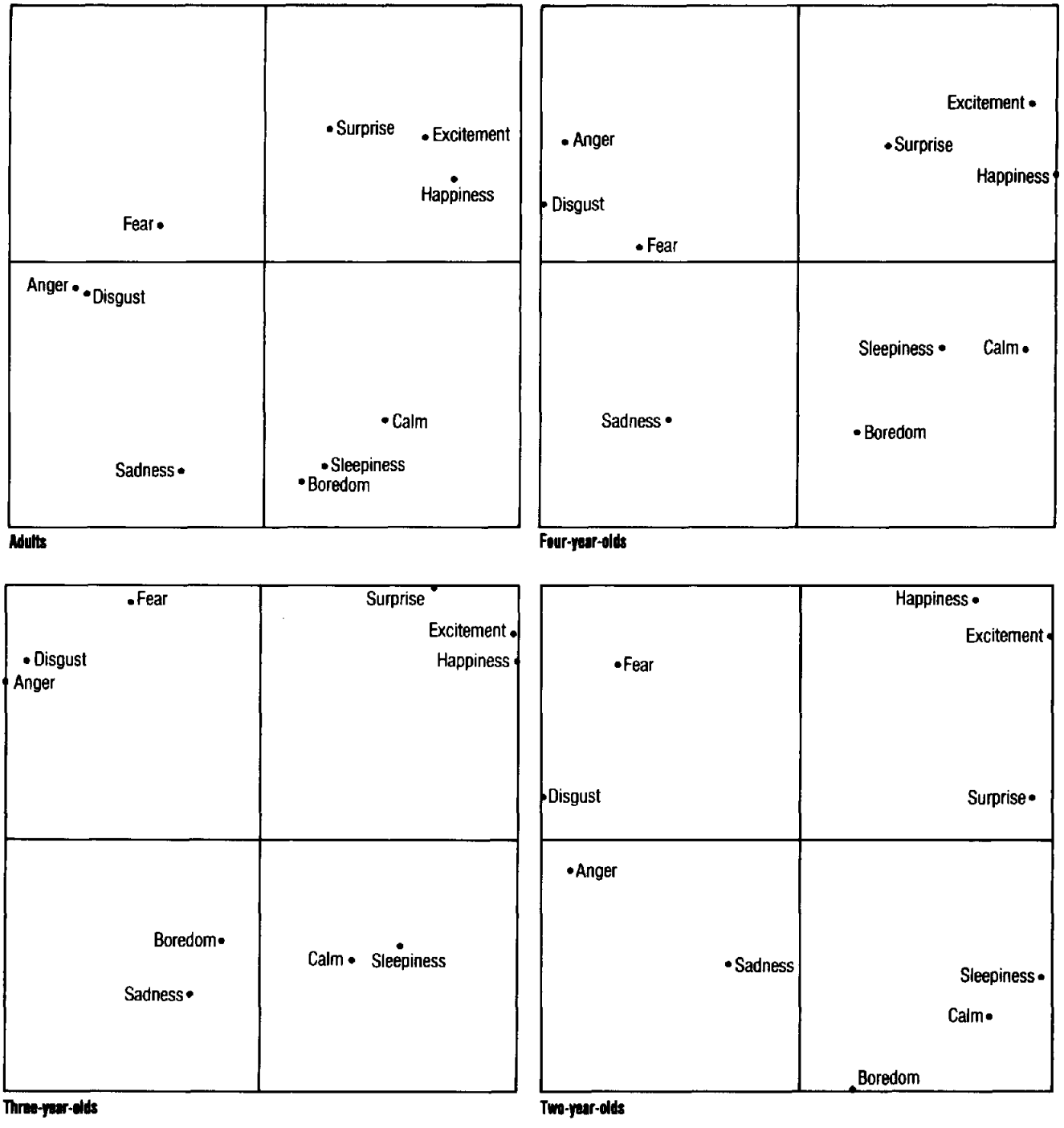


Figure 3. Two-dimensional scaling solutions of 10 facial expressions of emotion.

Table 1
Multiple Correlations Between Two-Dimensional Solutions and Single Dimensions

Two-dimensional solutions	Single dimensions: 2-year-olds	3-year-olds	4-year-olds	Adults	Pleasure/arousal
	Dimension 1				Pleasure
2-year-olds		.99	.95	.96	.89
3-year-olds	.97		.96	.98	.90
4-year-olds	.95	.96		.95	.95
Adults	.93	.98	.99		.96
	Dimension 2				Arousal
2-year-olds		.87	.80	.83	.91
3-year-olds	.89		.72	.96	.96
4-year-olds	.81	.73		.97	.93
Adults	.87	.97	.93		.96

Note. All multiple correlations given are statistically significant at the .05 level.

correlation of the first dimension from the 3-year-olds with one dimension from the 2-year-olds created by rotating the space. The figures reported in Table 1 for the first dimensions of multidimensional-scaling solutions are independent of those reported for the second dimensions.

Did the obtained dimensions correspond to pleasure–displeasure and degree of arousal? Inspection of Figure 3 suggests that the answer is yes, and this suggestion is supported by multiple correlations with independently obtained ratings for each of the facial expressions on pleasure and arousal. Separate groups of adults had previously been asked to rate the pleasure–displeasure and arousal–sleepiness conveyed by each expression (the details of the rating procedure are described by Russell & Bullock, 1985). These correlations are also reported in Table 1.

Was there a third dimension? Four-year-olds and adults produced a third dimension that appeared to put, at one extreme, expressions conveying assertiveness, boldness, and moving toward. At the opposite extreme were expressions conveying moving back, overwhelmed, taken aback, or cowering. (In terms of emotion categories, anger, calm, and happiness were contrasted with sadness, fear, and surprise.) The third dimension from 4-year-olds correlated .89 ($p < .01$) with that from adults. Third dimensions from 2- and 3-year-olds were not interpretable and failed to correlate significantly with each other or with the third dimension from either the 4-year-olds or the adults. A third dimension from a solution of only 10 stimuli must be interpreted with much caution, but the third dimension seen here is reminiscent of such third dimensions as potency, control, and dominance seen in previous studies (Russell, 1978). Our *failure* to obtain a third dimension in the younger children must be interpreted with even more caution. These results do not mean that our youngest subjects lacked a third dimension—only that with the resolution of our data, our evidence may have been insufficient to address the issue.

Discussion

Physical Similarity as an Alternative Explanation

In this study, the children had to rely on the physical features of the face to make whatever judgments they made. Might the children's judgments of similarity have been judgments of the similarity of physical features alone, rather than similarity of emotional meaning? After all, similarity of physical features is likely to be positively correlated with similarity of emotional meaning (Wiggers, 1982). Thus, if that correlation is high enough and if the children had been considering only physical similarity (rather than extracting an emotional meaning from those physical features), the resulting multidimensional-scaling solution might have looked approximately the same.[2]

For adults, this possibility can be ruled out because evidence from judgments of facial expressions fits a pattern of highly consistent results whereby emotions, whether conveyed verbally or nonverbally or experienced directly, are judged in terms of pleasure and arousal. Evidence for pleasure and arousal dimensions based on self-report data or judged similarity of emotion concepts exists for children as young as 8 years (Russell & Ridgeway, 1983), but we await such evidence for younger children. The similarity across age groups in the present study on facial expressions suggests that such evidence is likely to be positive for younger children, but we do not have conclusive evidence of this yet.

There is also reason to doubt that the alternative hypothesis of solely physical similarity can account for the entire circular structure of similarity across facial expressions. Even though similarity of physical features may be correlated with similarity of emotional meaning, the two are not identical. No one has shown that anything like the circular structure seen in Figures 1 and 3 can be produced from similarity of physical features alone. Most physical features in our set of 10 photographs were irrelevant to emotion. Whether children rely solely on physical features in judging similarity could thus be examined by systematically varying irrelevant physical features. Walden and Field (1982) found that irrelevant physical features had little effect on preschool children's ability to match emotion labels and faces, but again a definitive answer awaits more direct evidence.

Categories Versus Dimensions

Most investigators who write on the topic of children's interpretation of facial expressions focus on categories of emotion rather than dimensions of emotion. In fact, some writers ignore the possibility of dimensions altogether. What do the present results tell us about categories of emotion in preschoolers? We take our results as evidence *for* the existence of dimensions but not as evidence *against* the existence of categories. To distinguish

[2] Of course, whatever meaning the child found in a face was inferred from the physical characteristics of that face. The question raised here is whether the child's judgment of similarity between two faces was actually mediated by the *meaning* of each face or based merely on the similarity of their physical features. This worry does not arise in multidimensional scaling of words because similarity of meaning between two words is not correlated with similarity of physical features: *mad* is perceived to be more similar to *angry* than to *glad* even though the reverse is true if only physical similarity were taken into account.

categories from dimensions is not to set them up as mutually exclusive possibilities. Both categorical and dimensional accounts may be correct, as is the case with adults. (Or, neither may be. The child might attach to a face a different meaning altogether. Perhaps, to the preschooler, facial expressions are simply cues for different behaviors. Roughly speaking, the child may read such messages as "come here" and "go away.") Categorical and dimensional accounts of emotion perception may be in no more conflict than wave and particle theories of light. Each may be required to account for different properties of children's behavior.

We would like to argue that both accounts should be considered in future research. First, doing so may sharpen the questions experimenters put to children. We may find that a child's responses commonly taken to demonstrate the existence of categories can be interpreted as stemming from dimensions, or vice versa. It is not known if young preschool children can be said to possess the same basic categories for emotion that adults do. Although many studies have shown that young preschoolers can match category labels with the appropriate expressions with accuracy that is greater than chance (Izard, 1971), they are inaccurate enough that it is difficult to say precisely what underlies their responses. Recent evidence suggests that the way in which preschoolers categorize even prototypical expressions of emotion is not equivalent to the way adults do so (Bullock & Russell, 1984, 1985). Rather, preschoolers may possess precursors of adult categories, precursors dominated by the dimensions of pleasure and arousal. The present results give evidence for the hypothesis that pleasure and arousal play a role in how young preschoolers interpret emotion in the facial expressions of others.

Second, both dimensional and categorical accounts need not be true for children. Perhaps children develop categories very early and only later integrate them into the sort of circular structure seen here. In this case, dimensions like pleasure and arousal may be abstractions understood after categories are understood. On the other hand, the dimensions and circular ordering seen here may capture a global, relatively undifferentiated interpretation of emotion that develops very early (Bullock & Russell, in press). In this case, categories might represent further differentiations within these global perceptions and might be found later in development than dimensions. Alternatively, each system may develop independently.

References

Abelson, R. P., & Sermat, V. (1962). Multidimensional scaling of facial expressions. *Journal of Experimental Psychology, 63,* 546–554.

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

Bullock, M., & Russell, J. A. (in press-a). Further evidence on preschoolers' interpretation of facial expressions. *International Journal of Behavioral Development.*

Bullock, M., & Russell, J. A. (in press). Concepts of emotion in developmental psychology. In C. E. Izard & P. Read (Ed.), *Measuring emotions in infants and children: Vol. 2.* Cambridge, England: Cambridge University Press.

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

Cliff, N., & Young, F. W. (1968). On the relation between unidimensional judgments and multidimensional scaling. *Organizational Behavior and Human Performance, 3,* 269–285.

Dittman, A. T. (1972). *Interpersonal messages of emotion.* New York: Springer.

Ekman, P., & Friesen, W. V. (1976). *Pictures of facial affect.* Palo Alto, CA: Consulting Psychologists Press.

Izard, C. E. (1971). *The face of emotion.* New York: Appleton-Century-Crofts.

Lingoes, J. C. (1965). An IBM-7090 program for Guttman-Lingoes Smallest Space Analysis—I. *Behavioral Science, 10,* 183–184.

Lingoes, J. C. (1973). *Guttman-Lingoes Nonmetric Program Series.* Ann Arbor, MI: Mathesis.

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

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

Royal, D. C., & Hays, W. L. (1959). Empirical dimensions of emotional behavior. *Acta Psychologica, 15,* 419.

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

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

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

Russell, J. A. (1983). Pancultural aspects of human conceptual organization of emotion. *Journal of Personality and Social Psychology, 45,* 1281–1288.

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

Russell, J. A., & Ridgeway, D. (1983). Dimensions underlying children's emotion concepts. *Developmental Psychology, 19,* 795–804.

Russell, J. A., & Steiger, J. H. (1982). The structure in persons' implicit taxonomy of emotions. *Journal of Research in Personality, 16,* 447–469.

Shepard, R. N. (1962). The analysis of proximities: Multidimensional scaling with an unknown distance function. *Psychometrika, 27,* 125–139, 219–246.

Walden, T. A., & Field, T. M. (1982). Discrimination of facial expressions by preschool children. *Child Development, 53,* 1312–1319.

Ward, L. M. (1977). Multidimensional scaling of the molar physical environment. *Multivariate Behavioral Research, 12,* 23–42.

Wiggers, M. (1982). Judgments of facial expressions of emotion predicted from facial behavior. *Journal of Nonverbal Behavior, 7,* 101–116.

Received April 11, 1984
Revision received March 11, 1985 ■