A CROSS-CULTURAL STUDY OF FACIAL EXPRESSIONS OF EMOTION USING MULTIDIMENSIONAL SCALING

David Bimler and John Kirkland
Massey University
Shing-Jen Chen
Hokkaido University

Abstract

This paper searches for evidence of display rules, using multidimensional scaling (MDS). Groups of New Zealand and Japanese informants applied variations of the "method of scaling" to a particular set of 30 photographs of facial expressions. We constructed a map of "expression space": a two-dimensional spatial model in which each photograph is represented by a single point and where the distances between points correspond to the dissimilarities which the subjects perceived between expressions (as indicated by their sorting data). The map is a reasonable reflection of how people's perceptions of facial expressions are structured. As well as accounting for these sorting data, the map predicts how subjects respond to quite different questions. Combined with verbal descriptions of the expressions, it lends itself to interpretation in terms of polar coordinates, and also in terms of an x-y system of coordinates.

Key Words: display rule, facial expressions, cross-cultural study, multidimensional scaling

MAPPING DECODING RULES

There is evidence from a number of European and non-European cultures that people classify facial expressions into a similar set of emotional categories. The way that people perceive and recognise expressions follows a similar pattern, whatever their culture (Ekman, 1994). But it has been suggested that this underlying pattern can be modulated to a certain extent by culture-specific "display rules" and "decoding rules" (Ekman, 1972).

As an example of a display rule, when Japanese and American subjects were observed in the presence of an authority figure, the former showed a tendency to underplay their initial (negative) expressions (Ekman & Friesen, 1976). As an example of a decoding rule, Matsumoto and Ekman (1989) found that cultural background affected the level of intensity that informants read into expressions, depending on the emotion category

describing each expression. In their study, Japanese and American respondents rated the expression intensity of 48 photographs.

In the present study we adopted the Method of Sorting which is non-verbal: subjects are not asked to describe the photographs using a restricted set of emotion labels (the Forced-choice paradigm), or even without restrictions. This removes a potentially confounding source of cross-cultural differences, since the emotion labels in two languages are not necessarily exact equivalents of each other: "translations" can differ, either in fine nuances of meaning, or in the level of intensity they imply. From our Multidimensional Scaling (MDS) analyses maps are constructed; similarly perceived expressions are placed closer together.

Maps can be constructed for the New Zealand and Japanese informants separately. If the decoding rules differ between the two cultures, this should show up as systematic differences between the maps.

METHOD

Stimuli

The I-FEEL series of photographs are part of a projective test (Infant Facial Expressions of Emotion from Looking at Pictures). The version that is currently available, the I-30 series, consists of 30 polychromatic photographs (11 by 8.5 cm) of infants and young children, showing a variety of spontaneous, unposed expressions. When used as the test's designers intended, a subject describes the emotions displayed in each face; these descriptions are tabulated, and compared against base-line descriptions to assess whether the subject is projecting undue amounts of happiness, sadness, anger, etc. onto the faces and utilised for clinical intervention.

Subjects and procedure

Twenty two adult New Zealand participants (10 male, 12 female) sorted the 30 photographs into piles or groups, on the basis of similarity of expression, so as to place items which "looked similar" into a group together. This is "unconstrained" or "free" sorting (F-sorting for short); it leaves the subject free to decide on the criteria for assessing similarity, the number of items in each group, and the number of groups. F-sorting has been applied to facial expressions in a number of earlier studies (Emde, Kligman, Reich & Wade, 1978; Hulin & Katz, 1935; Nummenmaa, 1990; Russell & Bullock, 1986a, 1986b; Russell, Lewicka & Niit, 1989; Stringer, 1967).

Twenty New Zealand subjects also provided "successive" or "hierarchical"sorting data (H-sorting, for short). H-sorting provides substantially more information. This time the subject's task is to arrange the cards, step by step, into progressively larger groups: selecting the two most similar groups at each step and merging them into a single group, repeating this procedure until one group is left. To start with, the items are all separate, so the subject's first step is to select the two most similar photographs and combine them into a group of two. Data are recorded at every step in this procedure. Since in pilot testing it turned out that 30 items were too many to be conveniently H-sorted the cards were shuffled and split into two subsets of 15, which subjects H-sorted separately. This random splitting was repeated so that the two subsets were different for each

subject.

Twenty six non-English speaking Japanese subjects sorted the I-FEEL photographs for this study. Japanese subjects were all female: mothers of children at a kindergarten in Sapporo. They followed Additive sorting (A-sorting), a procedure which combines features of F- and H-sorting. As a first step, each subject arranges the items into piles. In each subsequent step, the subject merges the two most similar groups, thereby arranging the items into larger and larger groups.

Analyses

Sorting data are not suitable for analysis with the standard form of MDS, since the subjects do not directly indicate the inter-item proximities which they perceive. A common approach is to convert their responses into indirect estimates of the proximities ("co-occurences"). However, this approach suffers from a tendency to produce distortions and artefacts (Bimler & Kirkland, 1997).

By making certain assumptions, it is possible to calculate the probability that a subject sorts the items in a specified way, given a table of the proximities perceived between them. The "Method of reconstructed dyads" produces a spatial model (two-dimensional in this case, i.e. a map) which maximises this probability, for the sorting sequences which the subjects actually provided (Bimler & Kirkland, 1996).

Results

The sorting data from New Zealand and Japanese respondents were combined, and analysed using the method of reconstructed dyads. Figure 1 is the result. The points are numbered from 100 to 130, these being the labels assigned to the I-30 series of I-FEEL items.

To make sense of the 30 scattered points of Figure 1, we can incorporate information about the emotional labels used to describe each face. To do so we used a table of base-line descriptions which is circulated as part of the I-FEEL Test documentation. That table summarises the words used by 145 American subjects to describe each face.

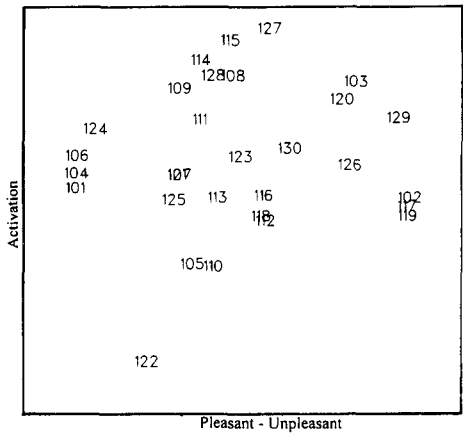
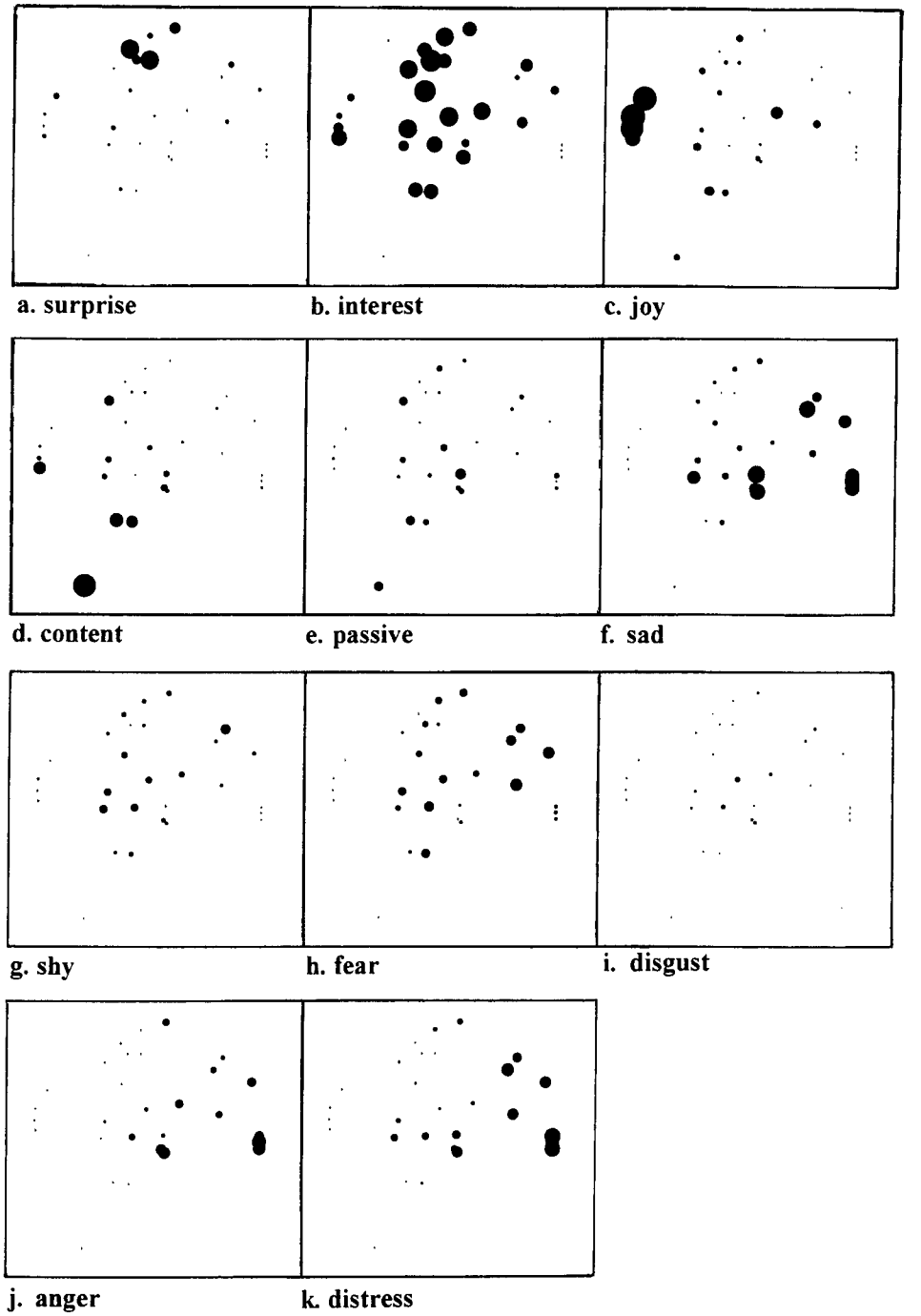


Figure 1 Map of 30 I-FEEL expressions derived from combined sorting data from New Zealand and Japanese informants.



Figures 2a-k Emotion-attribution data used to interpret Figure 1. Items are indicated by circles; the area of each circle is proportional to the percentage of people who described that face as: a. surprise; b. interest; c. joy; d. content; e. passive; f. sad; g. shy; h. fear; i. disgust; j. anger; k. distress. (crosses correspond to zero percentages.)

Their responses were sorted into 13 broad emotion categories (Surprise, Interest, Joy, Content, Passive, Sad, Shy, Shame, Disgust, Anger, Distresss, Fear, Other). For each photograph, the categories are listed as a percentage of the total responses, across a column of the table.

" Shame " occurs very seldomly (peaking at 2.1% of responses for Item 121). Ignoring that column, and the " Other " column, provided us the 11 panels of Figure 2. In each panel, the area of each dot is proportional to the percentages of subjects whose descriptions of the corresponding expression fell into that category.

To highlight differences and similarities between Japanese and New Zealand sorters we analysed their results separately and these are presented as the maps of Figures 3a and 3b.

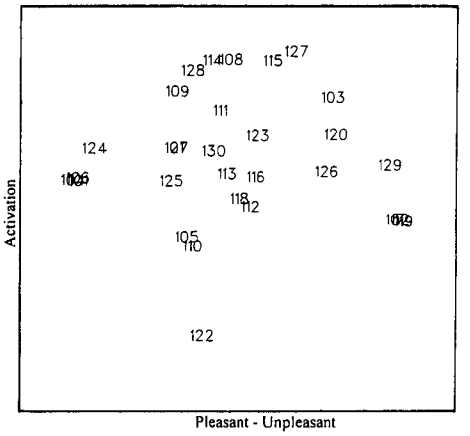


Figure 3(a) Map of 30 expressions derived from data from New Zealand informants in isolation.

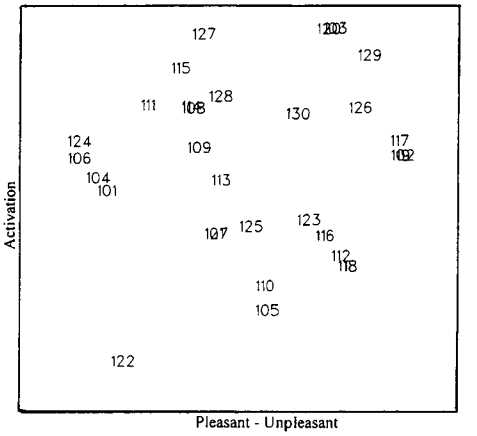


Figure 3(b) Map derived from Japanese informants in isolation.

DISCUSSION

General

One way to interpret Figure 1 is as a circle, divided by the emotion categories into overlapping sectors. The centre of the circle, where the sectors converge, is located roughly between points 105, 110, 112, 118, forming the origin of a system of polar coordinates. The radial coordinate represents intensity ; the angular coordinate corresponds to the emotion. The sequence of emotions around the circle is not arbitrary, but is determined by the pattern of similarities. As Woodworth and Schlosberg pointed out (1954), the same sequence also accounts for the pattern of misidentifications (to misidentify emotions, they must be sufficiently similar). For example, Matsumoto & Ekman note that expressions intended to convey " fear " were often called " surprise", a phenomenon noted also by Russell, Suzuki & Ishida, 1993.

Figure 1 can also be interpreted in terms of a pair of orthogonal dimensions. The horizontal axis can be identified as a " pleasure-displeasure " or " affective tone " dimension, since it runs from a cluster of smiling faces (101, 104, 106, 124), generally labelled " joy ", across to the clearly unhappy faces 102, 117, 119. The vertical axis runs from 122

(a sleeping child, generally labelled " content ") up to a cluster of startled expressions epitomised by 115, 127, and can plausibly be identified as an " activation " or " arousal " dimension. The same two dimensions emerged in a study of emotion words, applying MDS and factor analysis (Russell and Ridgeway, 1983). These are the dimensions used in the I-FEEL projective test, to quantify subjects' responses (the " Ridgeway technique ").

The dots in the panels of Figure 2 are definitely not distributed randomly. Expressions rated highly on each of the emotion categories are grouped together, forming zones or sectors. However, there are no sharp boundaries; the categories overlap, and fade into one another, in the way of " fuzzy concepts " (Russell & Bullock, 1996b). For example " Surprise " shades continuously into " Fear ", so that there are photographs for which either description is apt. Some adjacent categories overlap more than others - " anger " and " distress " are practically synonymous.

Cross-cultural comments

Matsumoto and Ekman (1989) found cultural variations in the intensity levels ascribed to different expressions of emotion. Japanese judges, rating 48 photographs on a 0-8 scale, rated " disgust " photographs as most intense; American judges rated " happiness " and " anger " as most intense.

Such variations can be investigated without making intensity judgments. We draw an analogy with colour vision. Like expressions, colours can be represented on a polar plot - the familiar colour wheel. One way to investigate colour-vision deficiencies (CVD) is to ask the subjects to assess the intensity of various primary colours, where any decrease in the intensity of red or green indicates a case of protanomaly or deteranomaly. An alternative non-verbal approach is to investigate the overall structure of colour similarities: a CVD flattens the colour circle into an ellipse and brings red and green closer to the neutral hues at the centre. Similarly, if variations in the decoding rules reduce the intensity of a given expression, for judges from a particular culture, we expect this to show up as a distorted polar-coordinate plot. Examples of that expression should be seen as more similar to central (neutral) expressions, bringing them closer to the centre. More salient expressions belong further out.

Russell, Lewicka and Niit (1989) collected cross-cultural sorting data and found no significant differences in the MDS solutions. But in that study, there were no opportunities for people to indicate whether items were closer to or further away from the centre, there being no " neutral " items to link them with - only adjacent items, around the outer edge of a polar plot (the expressions were picked to epitomise the emotions as clearly as possible, i.e. expressions of maximal intensity).

With the I-FEEL series, there are neutral items as well. This makes the sorting data more sensitive to differences in the saliences of the various emotional sectors. Figures 3(a), 3(b) result from separate analyses of the New Zealand and Japanese data, respectively. Clearly they are topologically the same; the general pattern of similarities and the sequence of emotions around the circle are not affected by cultural background. The location of item 122 (the face of a sleeping child) is poorly defined by sorting data: many people sort it into a group by itself, giving no indication how it related to other items.

Its varying positions in Figures 3(a) and 3(b) can be discounted. Figures 3a and 3b also differ in the exact placement of two extreme clusters, (101, 104, 106, 124) and (102, 117, 119), but again, these are relatively distinct expressions, which weakens the constraints that sorting data places on their locations.

Otherwise, these two Figures are remarkably similar. We conclude that cultural differences, if they exist at all with these stimuli, are smaller than the differences between age groups, sorting emotion words (Russell & Ridgeway, 1983). The variations in that study took the form of overall compressions and elongations of the solution - distorting a circumplex into an ellipse.

Recall that the New Zealand and Japanese groups were not matched for age. The Japanese subjects were the mothers of pre-school children, increasing their familiarity with the fluctuating moods of small children. There are also differences in the data-collection procedure. But additive sorting is intermediate in nature between H-sorting and F-sorting. We analysed the NZ F-sorts in isolation and compared them to the NZ split-design H-sorts; they are very similar.

Multidimensional scaling

It is difficult to make cross-cultural comparisons of perception of facial expressions by asking people to verbally identify expressions, since the verbal labels in the different languages may not be strictly synonymous. Words for emotions can vary in the levels of intensity they imply, and this can affect how applicable a word is to a given expression. Ekman and Matsumoto asked people to rate 48 expressions on how intensely they conveyed each of 7 emotions, on a 0-8 scale. But this confounds perceptual and linguistic variations.

MDS is a way of bypassing such possible queries. MDS also avoids the problems of context dependence which can affect how expressions are identified and rated.

Verbal descriptions of expressions and a MDS map of "expression space" are complementary. In a comparison across languages of the emotion labels used to describe a set of facial expressions (e.g. Russell, Suzuki & Ishida, 1993) the degree to which words are synonymous in the emotion lexicons of the contrasted cultures can be judged by creating counterparts to Figure 2, and comparing the "zones" or "sectors". This approach also allows one to specify how emotion terms differ: the extent and the way that the zones fail to overlap. If extending the I-FEEL projective test to other cultures - specifying the coordinates for each term in an emotion lexicon, in order to quantify a subject's overall response - again, we need a modified Figure 2 for this purpose.

We agree that the different expressions corresponding to different underlying emotions involve quite distinct patterns of facial-muscle contractions. But stimuli can fall into categories, and still fit into a spatial model. For example with colour we can represent "red" and "green" as sectors of a colour wheel, without denying that red and green stimuli produce quantitatively different experiences. There are continua, bridging the gaps between extreme, prototypal colours.

One approach in the study of facial expressions is to use stimuli of minimum ambiguity and maximum intensity: expressions with a single "right" identification. But in the real world one often encounters ambiguous expressions, open to more than one

interpretation, either because their intensity is low, or because signs are present of more than one underlying emotion. Among children, such blended emotions are the rule rather than the exception (Hiatt, Campos & Emde, 1979). Unblended emotions are an abstraction, akin to Platonic ideals.

These blended stimuli - expressions which straddle the boundaries between emotional categories - are less consistently pigeon-holed, when subjects are asked to categorise them into a restricted set of emotions or " primary affects " provided by the researcher (the forced-choice methodology). For the projective-testing purpose which the I-FEEL stimuli were selected for, this quality is desirable. It also makes them easier to map with MDS.

REFERENCES

Bimler, D. L. and Kirkland, J. (1997, in press). Multidimensional scaling of hierarchical sorting data applied to facial expressions. Scandinavian Journal of Psychology, *38*, 349-357.

Bimler, D. L. and Kirkland, J. (1996). A reconstructed-dyads approach to multidimensional scaling with sorting and hierarchical sorting data. Manuscript submitted for publication.

Ekman, P. (1972). Universals and cultural differences in facial expressions of emotion. Nebraska Symposium on Motivation (ed. J. K. Cole), *4*, 207-283.

Ekman, P. (1994). Strong evidence for universals in facial expressions. Psychological Bulletin, *115*, 268-297.

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

Emde, R. N., Kligman, D. H., Reich, J. H. and Wade, T. D. (1978). Emotional expression in infancy : I. Initial studies of social signaling and an emergent model. In : M. Lewis and L. A. Rosenblum (Eds.), The Development of Affect. New York : Plenum.

Hiatt, S. W., Campos, J. J. and Emde, R. N. (1979). Facial patterning and infant emotional expression : Happiness, surprise, and fear. Child Development, *50*, 1020-1035.

Hulin, W. S. and Katz, D. (1935). The Frois-Wittmann pictures of facial expression. Journal of Experimental Psychology, *18*, 482-498.

Matsumoto, D. and Ekman, P. (1989). American-Japanese cultural differences of intensity ratings of facial expressions of emotion. Motivation and Emotion, *13*, 143-157.

Nummenmaa, T. (1990). Sender repertoires of pure and blended facial expressions of emotion. Scandinavian Journal of Psychology, *31*, 161-180.

Russell, J. A. and Bullock, M. (1986a). On the dimensions pre-schoolers use to interpret facial expressions of emotion. Developmental Psychology, *22*, 97-102.

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

Russell, J. A., Lewicka, M. and Niit, M. (1989). A cross-cultural study of a circumplex model of affect. Journal of Personality and Social Psychology, *57*, 848-856.

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

Russell, J. A., Suzuki, N. and Ishida, N. (1993). Canadian, Greek and Japanese freely produced emotion labels for facial expressions. Motivation and Emotion, *17*, 337-351.

Stringer, P. (1967). Cluster analysis of non-verbal judgments of facial expressions. British Journal of Mathematical and Statistical Psychology, *20*, 71-79.

Woodworth, R. S. and Schlosberg, H. (1954). Experimental Psychology. New York : Holt, Rinehart & Winston.