SYMBOLS FOR DISPLAY OF MULTIVARIATE DATA:

THE FACE

Robert J. K. Jacob

Naval Research Laboratory
Washington, D. C.

People are well-known to be proficient at processing visual information (Entwisle and Huggins 1973). They can do sophisticated processing tasks, almost below the level of consciousness, when the data are presented graphically (Arnheim 1969). Until the advent of computer graphics, however, people were not nearly as good at generating graphical--or iconic--information as they were at assimilating it. Hence most data were actually communicated using symbols--in the symbolic mode, rather than the iconic mode. Now, the problem has become how best to use the iconic mode for communicating information (Huggins and Entwisle 1974). While there are some traditional iconic techniques, such as maps and cartesian graphs, given the new capabilities, it becomes worthwhile to look for other, perhaps better and richer ways to use the iconic mode for communicating information.

A novel iconic device for communicating multidimensional numerical information was proposed by Herman Chernoff (1971; 1973). This was the cartoon human face. People look at and process faces constantly. They have become well adapted to this task and are extremely good at performing it. Hence people would be expected to perform visual processing on faces better than on otherwise comparable visual stimuli. In fact, some evidence suggests that the perception of faces is a special visual process (Yin 1970).

To represent multidimensional numerical data facially, variation in each of the coordinates of the data is represented by variation in one characteristic of some feature of the cartoon face. For example, the first component of the data might be represented by the length of the nose. Other components would be represented by others of the 18 possible parameters, such as the curvature of the mouth, separation of the eyes, width of the nose, and so on. Then, the overall value of one multidimensional datum would be represented by a single face. Its overall expression--the observer's own synthesis of the various individual features--would constitute a single image depicting the overall position of the point in its multidimensional space. The variety of possible facial expressions would represent the variation possible in a set of numerical data. By looking at the faces and applying one's innate visual processing abilities to them, an observer could perform the visual equivalents of such tasks as multivariate clustering or identifying outliers as easily as he notices family resemblances between people, and by precisely the same, almost unconscious mental mechanism.

Figure 1 shows how the faces are used to represent data. Here, each face represents the value of an eight-dimensional datum chosen from an uncorrelated multivariate normal distribution. One datum differs significantly from the remaining 19 on several dimensions. It is rather clearly and rapidly identifiable (by a facial expression which differs from the remaining 19), despite the presence of considerable noise from the normal distribution. (It is Face 4.)

Several changes were made to Chernoff's original faces for these studies. The nose was changed from a line to a triangle, and its width became an additional variable. Chernoff's face height and width parameters were replaced by size and aspect ratio, which better match perceived dimensions. Some discontinuities in effects of changes in face outline parameters were remedied by providing a set of ratio parameters for the outline. Finally, it was found that reducing the range of variation on most parameters gave a more realistic set of faces; these were preferred because people are especially attuned to very small variations in realistic faces. (See Jacob (1976a; 1976b) for the computer program used to generate the faces.)

COMPARING FACES TO OTHER DISPLAYS

The first set of experiments (Jacob, Egeth, and Bevan 1976) was intended to ascertain whether subjects could perform common or useful tasks better with data displayed as faces or as traditional iconic or symbolic displays. In each of the two experiments, subjects performed a simple task involving a set of synthetic data. Performance was compared between subjects who were given the facial representation for the data and those who were given other representations (Chernoff and Rizvi 1975; Mezzich and Worthington 1978).

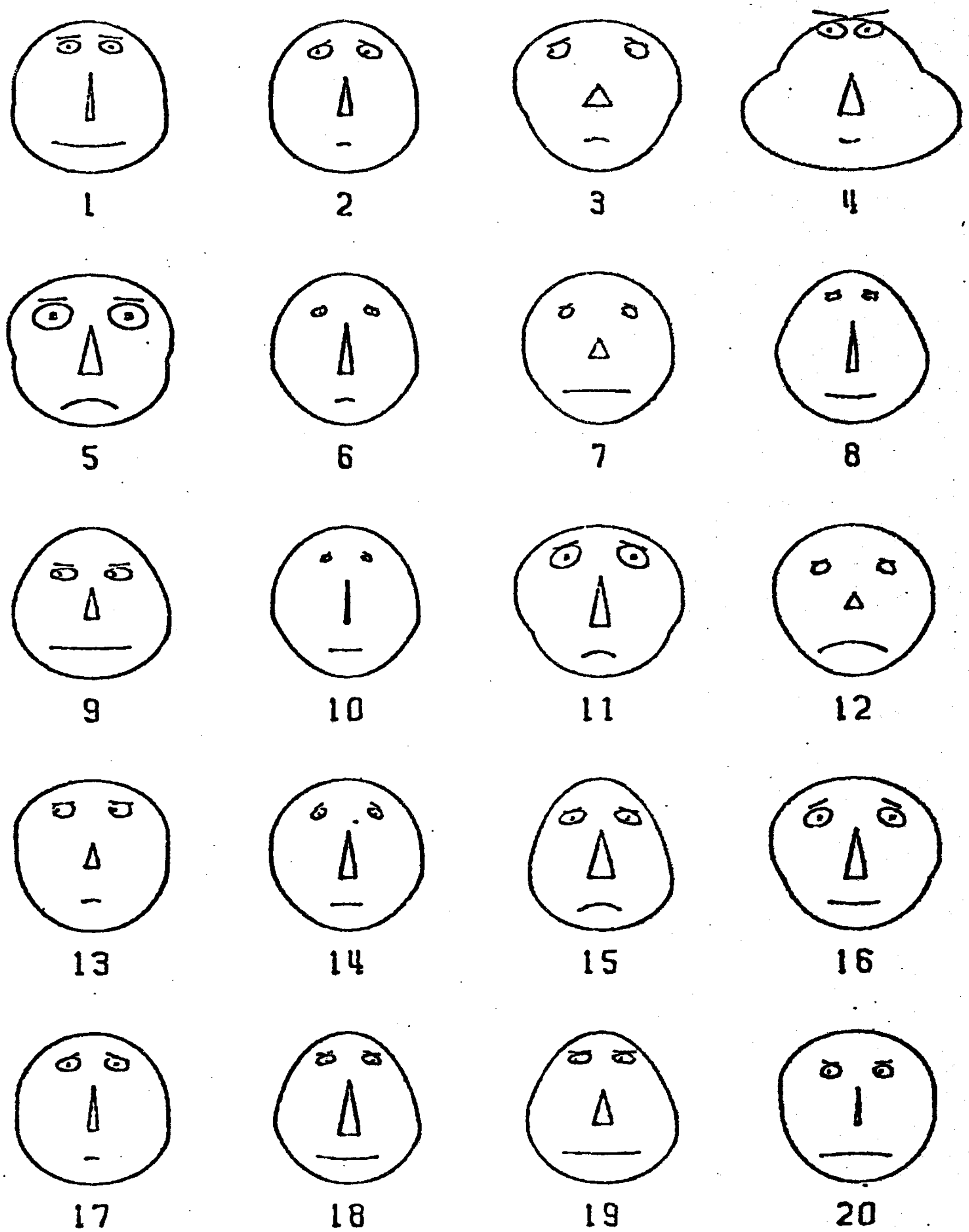

FIGURE 1. Example of a Facial Display of an Intensive Care Unit

Experiment 1

The task in the first experiment was paired-associate learning, a simple, standard task. It consists of asking subjects to learn to associate a name with each data point. Twelve such points were represented by digits, "glyphs" (Anderson 1960), polygons (Siegel, Goldwyn, and Friedman 1971), inverted faces, and faces. Each of these displays is illustrated in Figure 2. The entire procedure was repeated for three different dimensionalities. A total of 120 subjects were used.

Results revealed a variety of effects, some mutually confounding. There was a clear dimensionality effect as expected; subjects performed better on points in higher-dimensional spaces, since they contained a greater amount of memorizable information (Egeth 1966). Because the digit displays lent themselves to rote rehearsal, they induced rather good performance. The three-dimensional polygons gave particularly poor performance, probably because they are all perspective transforms of each other. The glyphs were difficult to organize perceptually and therefore to code, and they resulted in generally poor performance. The overall result, however, was that faces were at least as good as any of the other displays, and often better.

The most interesting observation was that the inverted faces tended to be difficult to learn to label, especially when they were of low dimensionality. Inverted faces provide a good control for perceptual factors such as complexity, symmetry, and integrality of elements, but they lack the familiarity of faces. Hence this finding suggests that familiarity may be an important factor in explaining the superiority of the face display. That is, the face does not appear to be simply one of a number of possible geometrically well-designed displays, but, rather, it has unique properties.

Experiment 2

While the paired-associate learning task was a standard psychological research task, it was not the sort of task to which the faces were intended to be applied in practice. The second experiment investigated a realistic and practically-useful task. This was clustering, or sorting into categories, or pattern recognition.

The task consisted of a set of 50 points in a nine-dimensional space, which were to be organized into five groups. They were generated in five clusters, each normally distributed around a center point, called the prototype.

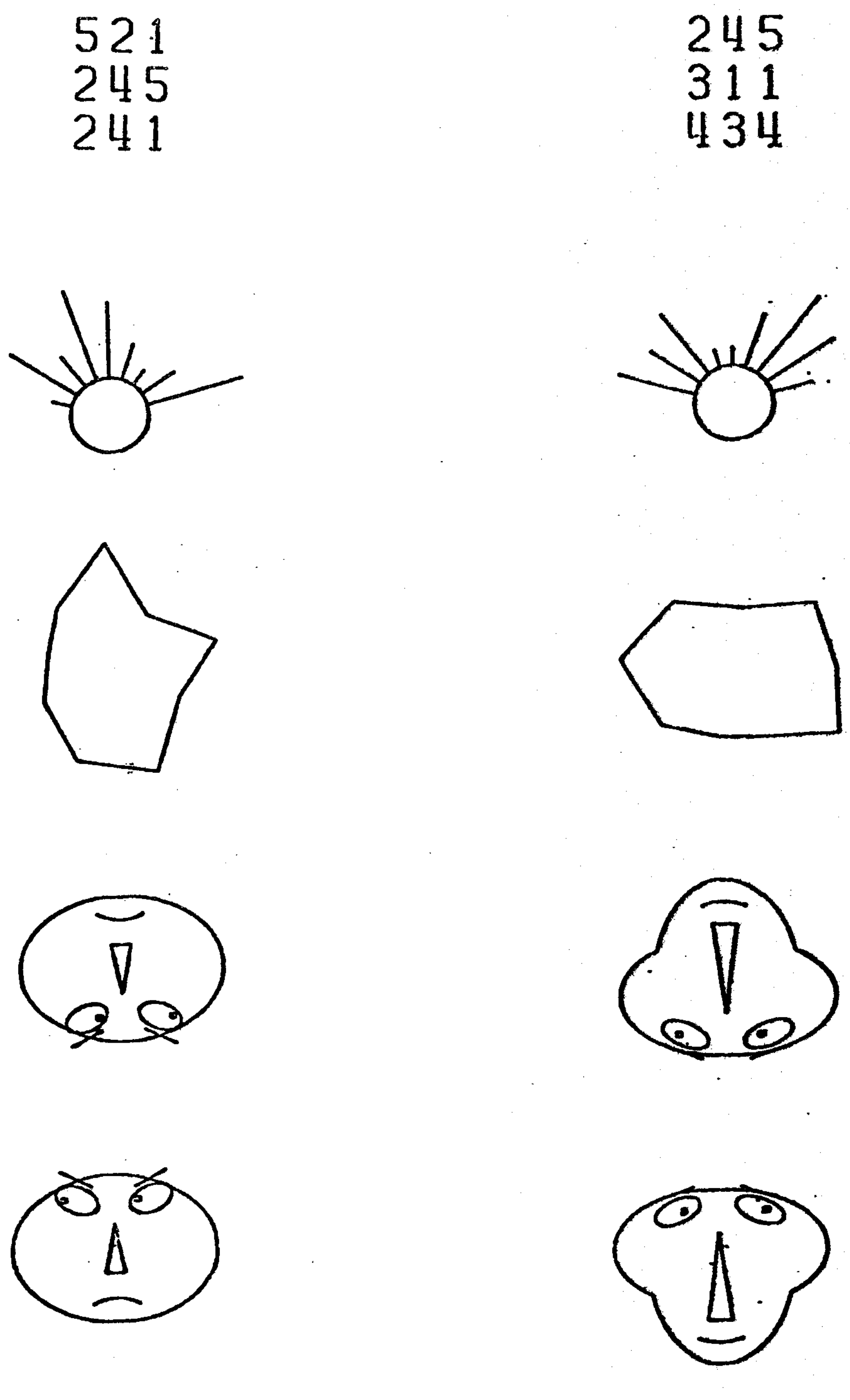

FIGURE 2. Example Stimuli from **Experiment 1**

The subject's task was to look at the five prototypes and then assign each of the 50 deviants to a cluster surrounding one of the five prototypes. The correct answers were those which put deviants with the prototypes from which they were generated, and to which they were closest in Euclidean distance. (Conventional cluster analysis procedures were also applied to the deviants, and they successfully grouped together those deviants that were associated with each prototype.) While this was a contrived task in that questions were derived from the answers, it was outwardly similar to many realistic tasks. In a real task, the subject would have the five prototypes in his mind, abstracted from his experience or training. He would look at a new data point and assign it to one of the groups he knew. For example, a doctor would examine the data on a patient and then assign him to a cluster that represents a particular disease.

As in the first experiment, the 55 data points were represented in several different ways, and subjects performed the same task with the different displays: faces, a second set of faces with the range of possible variation reduced to three-fourths that of the first, polygons as in the first experiment, and digits. Figures 3 through 6 present the prototypes (top row of each figure) and examples of their deviants (succeeding rows) for the four different display types respectively. Polygons were used here because they had been found to be the better of the two alternate graphic displays used in the first experiment, probably because their elements are better integrated (Garner 1974).

Results consisted of the number of errors subjects made in classifying the 50 points. Table I shows the mean number of errors (chance performance would give 40 errors) they made and the mean time (in seconds per card) they took in sorting the 50 cards. The two types of faces were found

TABLE I. Results of Experiment 2 — 24 Subjects

	Faces	Faces (3/4 range)	Polygons	Digits
Mean no. wrong	14.46	17.29	28.21	31.38
Standard dev.	5.50	5.95	5.27	7.50
Mean time (sec./card)	4.97	4.88	4.43	9.89
Standard dev.	1.96	1.75	1.72	3.86

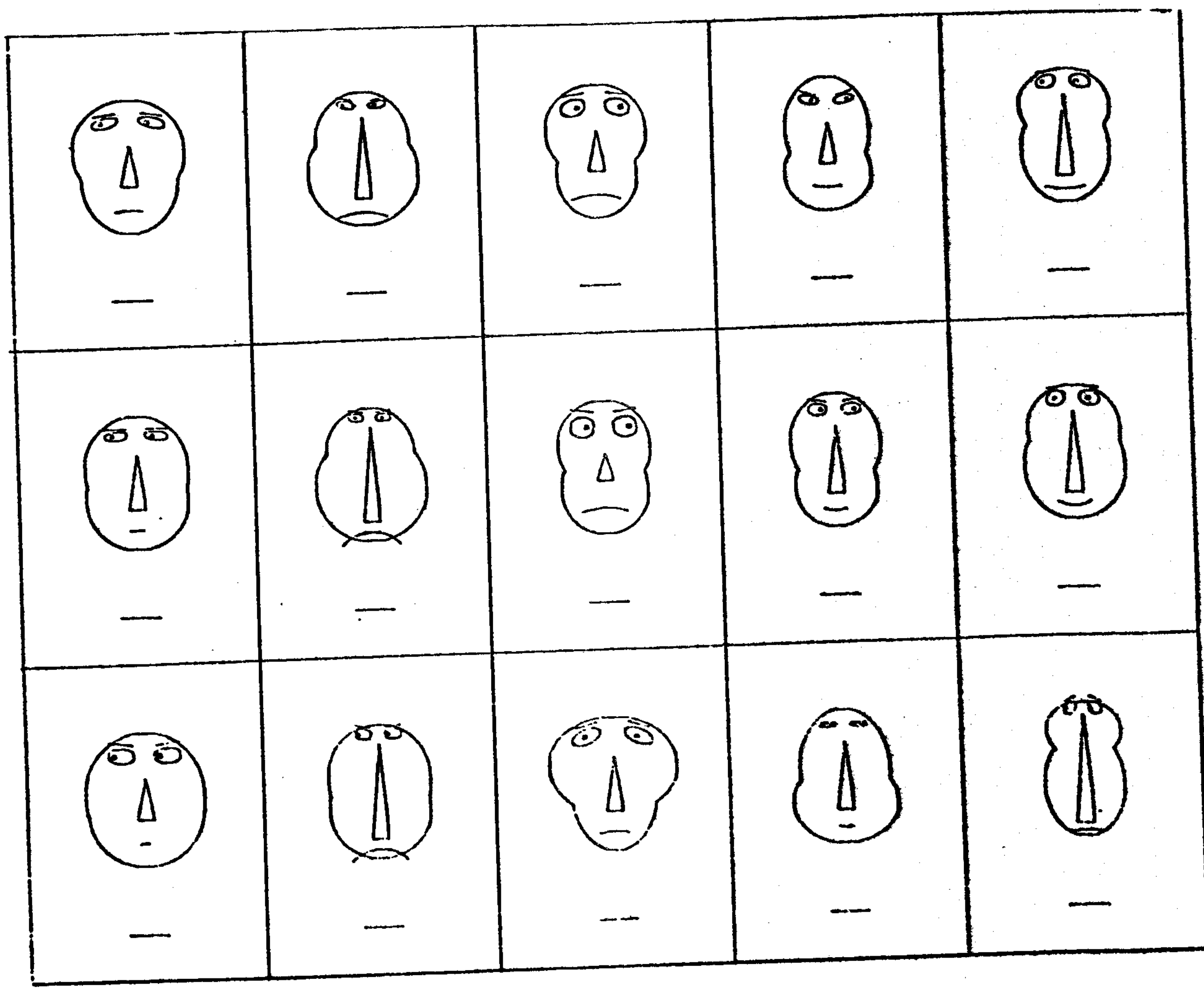

FIGURE 3. Example Stimuli from Experiment 2--Faces

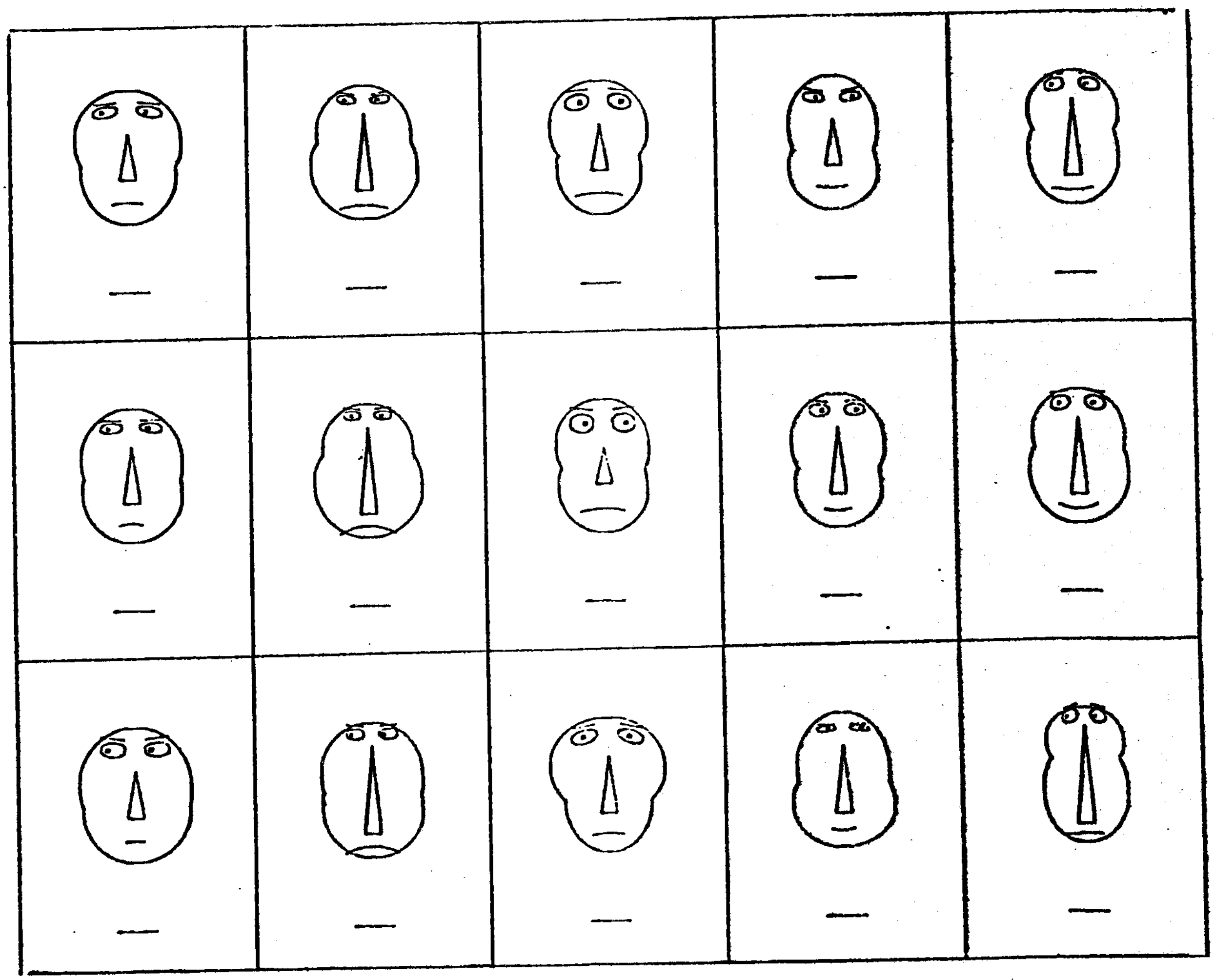

FIGURE 4. Example Stimuli from Experiment 2--3/4 Range Faces

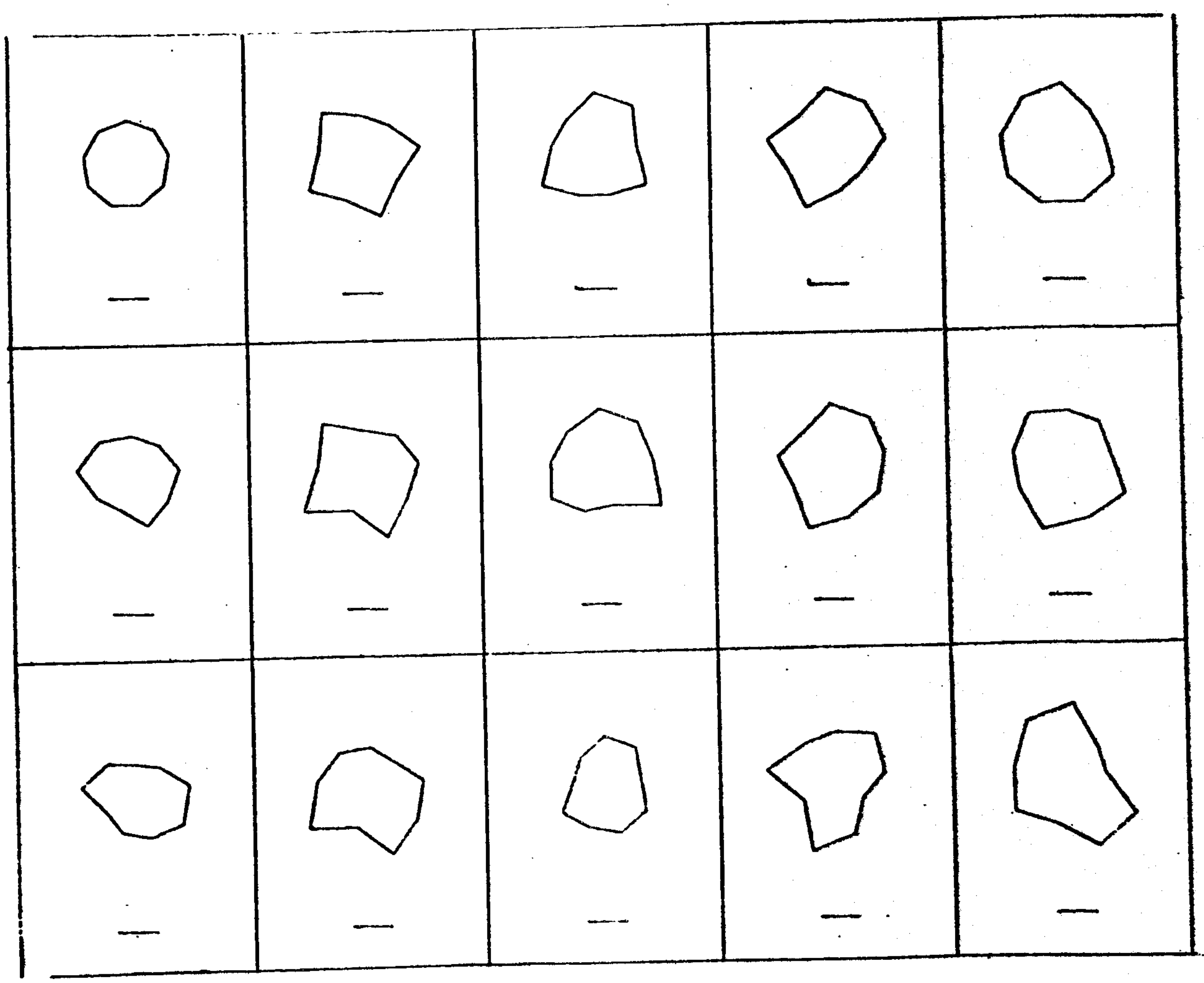

FIGURE 5. Example Stimuli from Experiment 2--Polygons

333 333 333 —	363 636 363 —	633 633 636 —	636 363 366 —	666 666 633 —
345 214 343 —	374 827 454 —	644 532 856 —	635 375 556 —	465 586 734 —
135 123 552 —	345 827 552 —	421 434 424 —	448 285 146 —	886 748 812 —

FIGURE 6. Example Stimuli from Experiment 2--Digits

clearly to be superior to both the polygons and the digits at $p < 0.001$. Significant differences were not found between the two face types or between the polygons and digits. While the polygons could be sorted as quickly as the faces, they were not sorted correctly.

The conclusion drawn was that subjects performed a realistic and useful task significantly better when the data were represented by faces than when they were represented by a conventional display (digits) or by a well-integrated graphic display (polygons). As the experimental task is a fairly general one, one underlying many specific data analysis tasks such as diagnosis, pattern recognition, and cluster analysis, it is claimed that faces provide a superior display for many multivariate applications. Subjects' comments on the experiment help explain this result. They reported that they put all the "happy" faces in one pile, "angry" ones in another, and so on; they found this easy to do. In fact, because of the representation, they were performing a fairly sophisticated multivariate clustering task accurately using only their visual processing abilities. For the other displays, they reported inventing more complicated strategies, which turned out to be self-defeating.

This synthesis by the observer himself of the various graphical elements of the facial display into a single gestalt is one of the principal advantages of this type of iconic display. Many other common types of displays contain several variable elements and could thus be used for graphing multivariate data, but often such displays predispose toward a piecemeal, sequential mode of processing, which obscures the recognition of relationships among elements. By contrast, faces induce their observer to integrate the display elements into a meaningful whole. Previous research with simple cartoon faces and with photographs of real faces has indicated that observers do indeed process these stimuli in such a wholistic fashion (Yin 1969; Smith and Nielsen 1970; Reed 1972).

USING THE FACE DISPLAYS

Having supported the initial supposition that the faces provide a demonstrably good display for Euclidean data of several dimensions, the problem of displaying a specific type of actual (rather than synthetic) data using faces was addressed. The data selected for this purpose were the results of a psychological test intended to determine a patient's psychological personality profile. It was thought that such a profile might possess a more natural facial representation than most other sorts of data.

The form the data took was the results of five particular scales of the Minnesota Multiphasic Personality Inventory (MMPI) (Hathaway and McKinley 1942). The U. S. Public Health Service Hospital in Baltimore administers this test to patients as part of a comprehensive health testing and evaluation and was interested in alternate ways to display the test results. The hospital uses five of the clinical scales of the MMPI: Hypochondriasis, Depression, Paranoia, Schizophrenia, and Hypomania.

Following the approach both of Chernoff and of the previous two experiments, the five components of an MMPI data point could simply have been assigned arbitrarily to five of the facial features (while the unused features were kept at constant values). The resulting facial expressions and the personality traits or disorders which each represents would then be learned by doctors, just as they have learned the meanings of the numerical data and the graphs presently in the patient reports. However, it has been widely observed (e.g. Secord, Dukes, and Bevan 1954; Harrison 1964) that particular facial expressions tend to signify particular personality traits to observers with great consistency. Therefore, if the face displays could be devised in such a way that the expression on the cartoon face suggested the same personality traits as those in a particular MMPI report, the resulting face displays would tend to communicate the meaning of the data they represent intuitively. To this extent, a self-explanatory display would have been constructed, somewhat like an hypothetical graph in which it is not necessary to label the axes, because the meaning of the curve is inherently obvious.

Consider, for example, a particularly unfortunate arbitrary assignment of MMPI scales to facial features, in which a smile on the face signified a patient suffering from severe depression. While this could certainly be learned, just as the letters depression or the shape of the personality profile are learned, such training would clearly be a poor utilization of the observer's skills.

Therefore experiments were undertaken to attempt to obtain a positive relationship between the five components of the MMPI score vector and the 18 variable parameters of the face construction. It was hoped that the resulting face displays would be highly intuitive and suggestive; unlike most computer output formats which require the human observer to learn to understand the computer's language, the power of the computer would here be used to tailor the display format to suit and exploit the person's intuition and preconceptions.

Experiment 3

The first experiment in this study attempted to measure a relationship between MMPI scores and face parameters based on one observer's preconceptions or stereotypes. This corresponded to a transformation between the five-space of MMPI scores and the 18-space of face parameters. Because of the imprecision in the process of perception of personality from faces, it was hoped that a linear model would be sufficiently accurate for useful results. Subsequent analysis of the experimental results for higher-order interactions showed this to be a reasonable choice. Moreover, the dimensionality of the problem made any other model very much more difficult to study. Thus a matrix (T) was proposed to define a linear transformation from the space of MMPI score vectors (d for diagnosis) into that of face parameters (p).

A set of 200 faces was generated using parameter (p) vectors chosen from an 18-variate uniform random distribution. Figure 7 shows a sample of these faces. Dr. Faith Gilroy, a research psychologist at the Public Health Service Hospital, then rated each of the faces on the five scales. She was, in effect, indicating what MMPI results each of the faces signified to her, or, more specifically, what MMPI score she thought a person who looked like each of the 200 faces would receive.

A multiple linear regression of the p vectors on the d's was computed from 200 pairs of such vectors, producing a T matrix of regression coefficients (Jacob 1976a). That matrix could then be used to estimate a p vector (or face drawing) for any given d vector (or personality score). Such estimated p vectors were computed and compared to the original (stimulus) p vectors; the mean squared error over all components of all the vectors was 0.07497 (components of the p vectors ranged between zero and one).

The T matrix was displayed graphically by computing the p vectors that correspond to equally-spaced points along the axes of the d space (that is, points that represent patients who have only one psychological disorder). Figure 8 shows the resulting display. In it, each row depicts a series of patients with increasing amounts of a single disorder. Because of the rating scale used, zero (the first column) represents an inverse amount of the disorder, one represents no disorder (the origin of the d space), and four represents a large, extrapolated amount of the disorder. It was thought that these faces (particularly those in the column labelled three) actually corresponded to common stereotypes of the personality traits they were claimed to represent. The subject had never seen these faces nor any resembling them; rather, they had been deduced from the linear

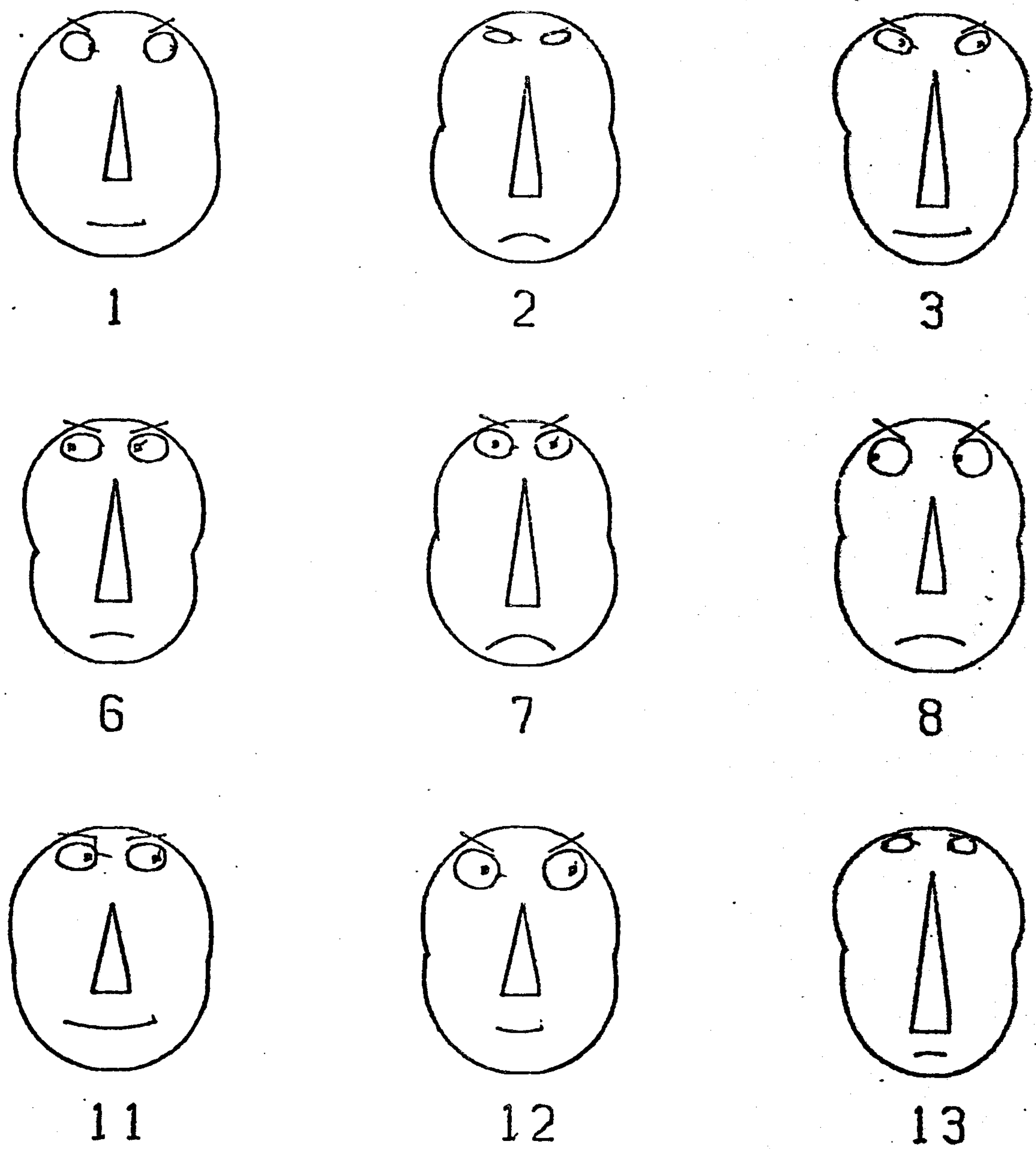

FIGURE 7. Example Stimuli from Experiment 3

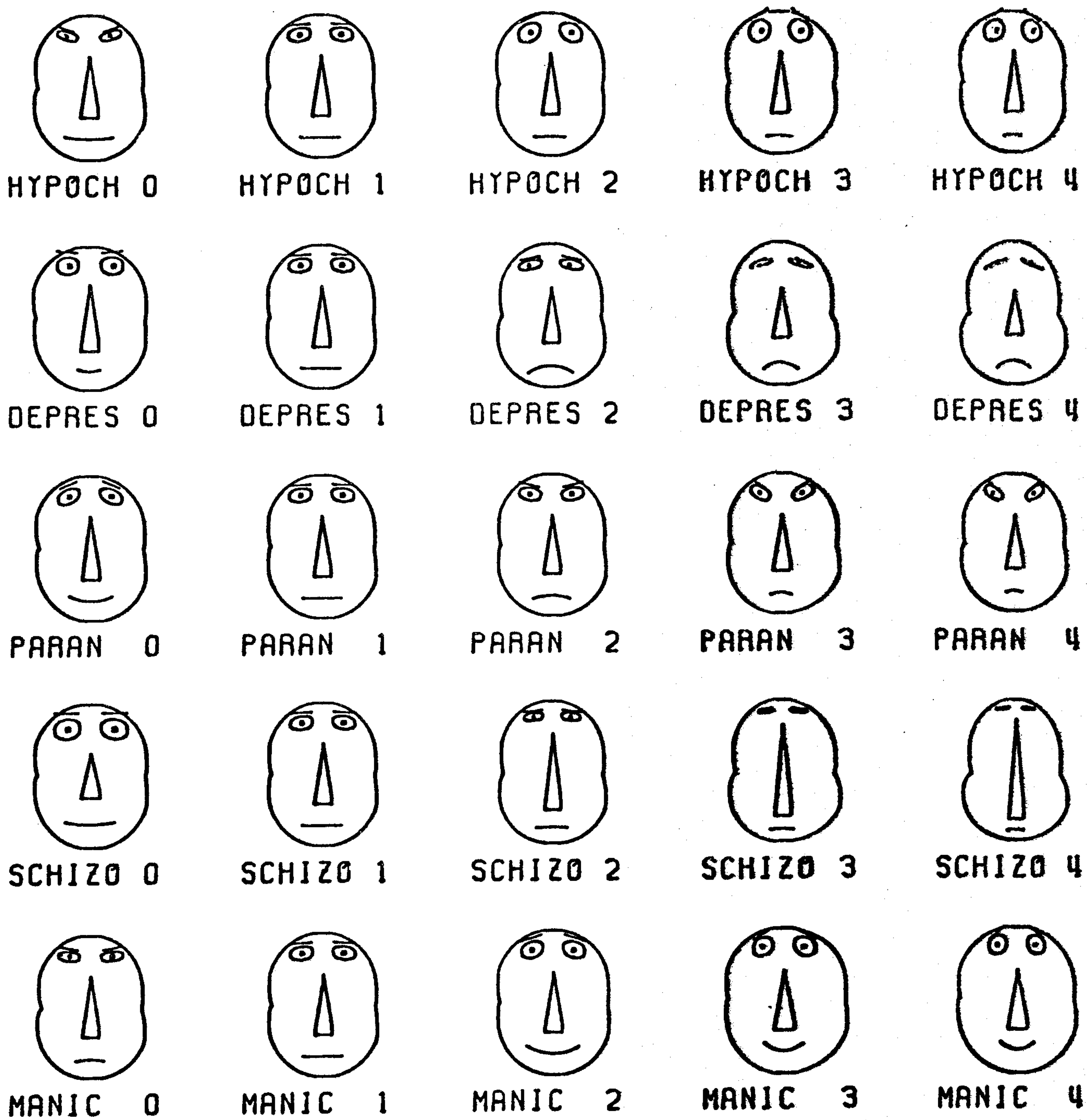

FIGURE 8. Facial Representation of the T Matrix

regression using faces reported to have more than one disorder.

Some comparisons were made between this T matrix and results obtained by previous investigators. While no studies had used stimuli of this complexity or the same rating scales, some of the observed relations between basic facial feature variations and basic emotions were confirmed. Comparison to the work of McKelvie (1973) and of Harrison (1964) corroborated both the major axes of facial variation found in the T matrix and their relationships to variation in emotional states. (As one might expect, these all suggest that the joint variation in the mouth and eyebrows are the major determinants of emotional content of the facial expression; that variation induces variations along axes comparable to Paranoia, Depression, and Hypochondriasis.)

An attempt was made to determine the important factors or axes incorporated in the transformation T by performing a canonical correlation analysis on the 200 pairs of p and d vectors (Morrison 1967; Tatsuoka 1971). This procedure finds pairs of axes in the p and d spaces in such a way that the correlation between each such pair is maximized and there is no correlation between any two pairs. Thus, all of the correlation between the two sets of data is contained in the correlations between pairs of corresponding axes; and the axes in each space are mutually orthogonal. The first three of the axes so discovered were found to be statistically significant ($p < 0.001$); they gave correlations of 0.747, 0.652, and 0.455. These axes are plotted facially in Figure 9, in the same manner as those in Figure 8, except that the origin is labelled zero here (center column), and the other columns show points lying .5 and one standard deviation away in either direction.

The first axis appears to be related to a "happy-sad" dimension; its counterpart in the d space had large and opposite loadings on the Depression and Hypomania scales. The second axis had a large loading on Paranoia in one direction and on Hypochondriasis and Hypomania in the other, suggesting that eye angle is communicative of a range of expressions from anger or intensity to a vacant or helpless look. The third axis is less suggestive; it seems to take up much of the remaining variability in the facial features. The first two axes correspond closely to the results of the previous investigators cited, and they are also intuitively plausible. These axes could be used to suggest a new set of coordinates for face parameters that match the perceptual space; but, strictly, they are only pertinent to the representation of the MMPI data under study.

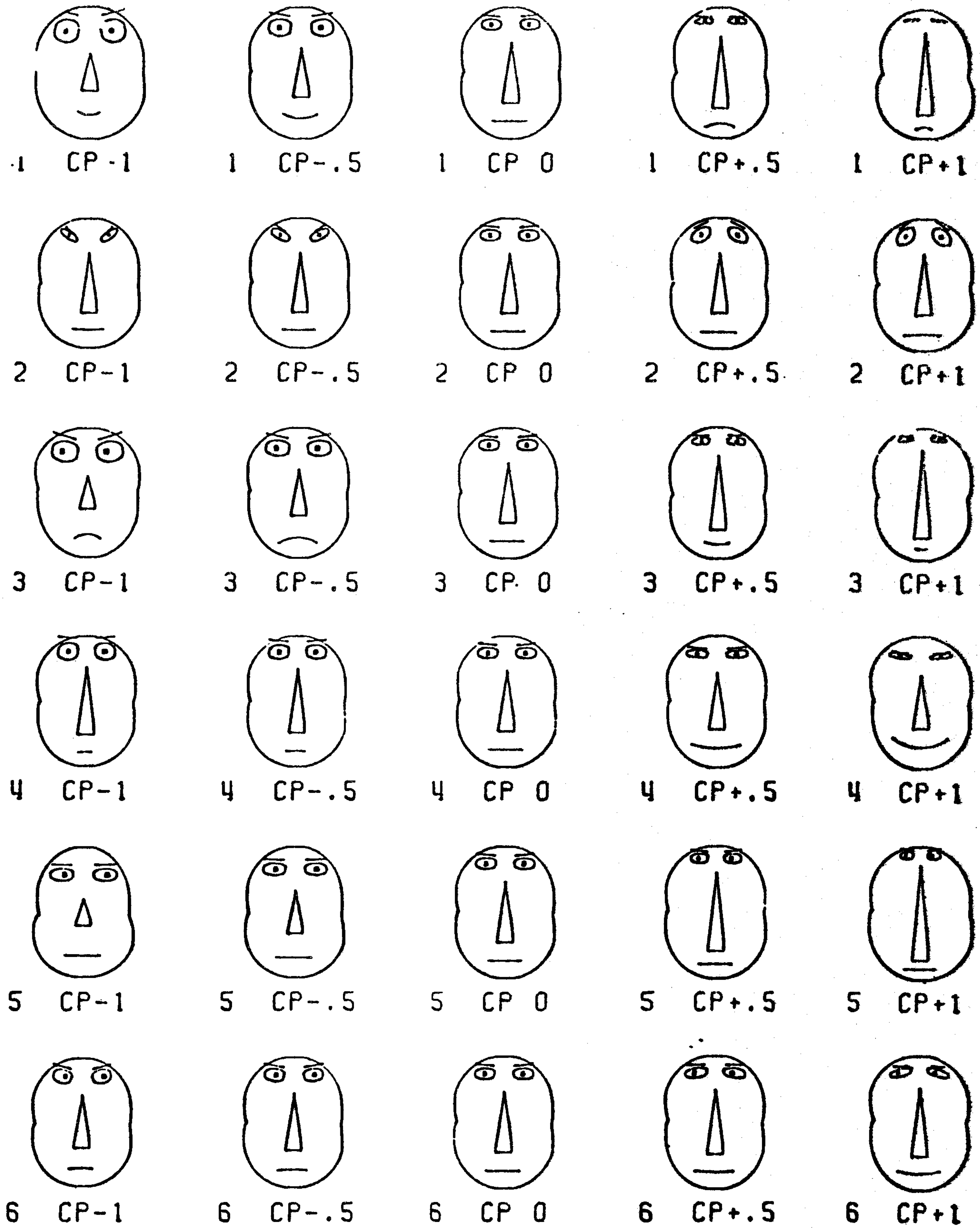

FIGURE 9. Facial Representation of the Canonical Axes in the p Space

An additional computation shows that the angles in the p space between the facial representations of the orthogonal axes of the d space ranged from 70 to 112 degrees, suggesting that the orthogonal d axes were indeed perceived as being related to orthogonal variations in their facial representations.

Thus, Experiment 3 provided a linear transformation from MMPI scores to faces which was both intuitively appealing and internally consistent. Further study was undertaken in order to validate and then apply this relationship.

Experiment 4

An attempt was made to replicate the previous experiment with the same and with another subject. A new set of random faces, generated similarly to the first set, was presented to two subjects who rated them as in the previous experiment. Actual responses were compared to those predicted using the T matrix of Experiment 3.

The comparison was confounded by the appearance of significant response bias. That is, subjects gave consistently higher or consistently lower ratings to the faces on certain scales. It could be determined that, in those cases where the response magnitudes matched the predictions (approximately half of the data), the present results supported the previous ones in direction as well. In the remaining cases, neither support nor contradiction could be asserted. This experiment could have been improved by embedding the stimulus faces in a larger group which would have induced subjects to attain the same mental set (and thus the same response bias) as that of the subject during Experiment 3. Instead, experience from this experiment was used to devise a new experiment which would provide a more powerful test of the transferability of the T matrix relationship.

Experiment 5

For the relationship T to be valid and transferable to other observers, it must appeal to intuitive stereotypes that are already present in the minds of most observers. Such stereotypes need not possess any absolute validity; they need only be widely and uniformly held in order to be exploitable in devising a facial display for MMPI scores. Thus, this experiment was designed to test the applicability of the stereotypes already discovered. Untrained subjects in the experiment were asked to match facial representations of random hypothetical MMPI scores to alternate representations of the same data. Since the numerical MMPI scores were not meaningful to the subjects (or to the intended final users of the display), an independently-developed textual representation for MMPI scores (Rome *et al.* 1962) was used in this study.

The 30 subjects were each given 50 stimuli, an example of which appears in Figure 10. In each, the subjects were asked to indicate which of the five faces given best corresponded to the given text description. In fact, that description was the textual representation of a particular point in the d space. One of the five faces was the facial (using the T matrix) representation of the same point, and the remaining faces were representations of other, randomly-selected points.

The principal result of interest was whether entirely naive subjects could select the face that was claimed (by the results of Experiment 3) to represent the same MMPI data as the text at better than chance performance. If the T matrix had no wider validity than for one subject at one time, the subjects would not perform the present task; if, however, the matrix relationship corresponded to widely-held stereotypes, the subjects would use such to perform this task better than a random guessing hypothesis would predict. Results were obtained by measuring the Euclidean distance in the five-dimensional d space between the expected answer and the answer a subject chose. Such a distance could range from zero (correct choice) to 4.5 (the maximum diagonal dimension of the hypercube). Table II presents these data. A matched *t* test on the data revealed that subjects were able, with highly significant ($p < 0.0005$) accuracy, to choose those faces which were designed to communicate the same information as the text items.

One concern was that measuring the Euclidean distance between the correct answer and the answer a subject chose tended to emphasize unimportant differences (between two answers that were both far from the correct one) while obscuring more important differences (between correct and nearly-correct answers). To correct for this, the

TABLE II. Results of Experiment 5 — 30 Subjects

Chance performance	1.571
Mean observation	1.226
Standard deviation	0.146
t(29)	12.975

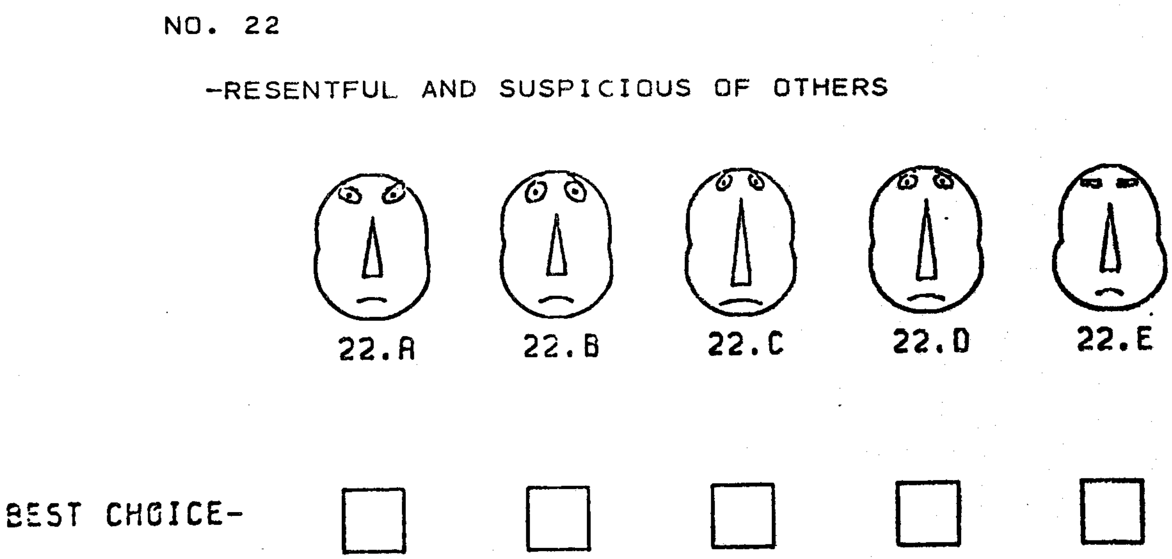

FIGURE 10. Example Stimulus from Experiment 5

transformation EXP(-DISTANCE) was used as an alternate measure for scoring the experimental results. Further, the scoring was repeated using the reciprocal of the rank-order distance, in order to eliminate the effect of the arbitrarily-varying distances between the randomly-generated stimulus points.* However all of the measures used in scoring gave similar results.

The results of this experiment suggest that the faces plus the T transformation obtained provide a data display that requires no training of the observer. Without any prior information other than their innate facial sterectypes, subjects were able correctly to perceive the data being displayed.

Experiment 6

Experiment 5, then, demonstrated that the faces could be used to communicate psychological data to naive subjects. Experiment 2 showed that a particular useful task could be performed better and more quickly with facially-represented data than with several other representations. Together, the experiments suggest that the face might be superior mode of displaying the MMPI data under consideration. The present experiment was intended to test this composite hypothesis by having subjects perform a meaningful and realistic task which requires apprehension of MMPI data. Various subjects would perform the same task using the facial and textual representations of the same MMPI data, and their performance would be compared.

---

*It was originally intended that the four wrong answers to each question in this "multiple-choice test" would lie at the same four distances from the correct answer for every stimulus. However, this meant that those four answers would be distributed near the surface of a hypersphere with its center at the correct answer. Even when the four distances differed substantially, the five faces imparted a strong sense of this geometric relationship, and, as a result, one could use this relationship rather easily to select the correct answer in each group without seeing the question or knowing the geometry explicitly. Skewing the distribution of the directions in which the wrong answers lay changed the configuration to an approximation of a hyper-cone with the correct answer at its vertex, but a comparable difficulty arose. Finally, randomly-varying distances were permitted, and the experimental results were scored using both actual and rank-order distances.

A truly realistic task would be the diagnosis and treatment of a real patient. The results would be measured by evaluating the patient's well-being at the conclusion of the treatment. Unfortunately, there would be far too many confounding variables in such an experiment (as well as practical problems). Instead, a crude task, analogous to psychological triage, was devised. Subjects were asked to rate the overall emotional well-being of an hypothetical patient, given his MMPI test scores presented in one of two ways. Their success would be measured by comparing their responses to the responses of a clinical psychologist who studied the unprocessed numerical MMPI scores. Thus, to the extent that a naive subject's responses, using the facial or textual representation, corresponded to this baseline, it could be claimed that, through the use of that representation for the data, he was able to perform the same task as the trained psychologist.

Thirty-two subjects were each given 50 stimuli, each of which resembled either Figure 11 or Figure 12. In each case, the subject was being asked to rate a random point in the d space (represented facially or textually) for emotional well-being.

Results were obtained by measuring the correlation coefficient between a subject's ratings and those of the psychologist. A chance hypothesis would have predicted zero correlation. The mean correlation scores over subjects are presented in Table III. First, one can observe that subjects' performance exceeded chance expectation significantly ($p < 0.005$) for both faces and text. Next, conventional and also paired-observation t tests were made to find the difference between the two display types. Both tests showed that subjects performed the task significantly ($p < 0.005$) better when given the text than when given the faces.

Some insight into this unexpected situation may be gained by studying the text displays in more detail. It appeared that the more disturbed a patient was, the longer his text description was. Hence subjects' responses to the text could have been based on this inadvertent iconic content of the text display; they could have been responding to the quantity of text rather than to its meaning. To test this, an algorithm that rated the emotional well-being of a patient based only on the quantity of text in the textual representation of his MMPI score was applied to the experimental stimuli. As shown in the table, the algorithm achieved slightly better performance than the subjects who used the text display. Thus the superior performance of the text displays could be explained by their unintentional iconic content; or, illiterate subjects could have produced the same responses from the text displays as

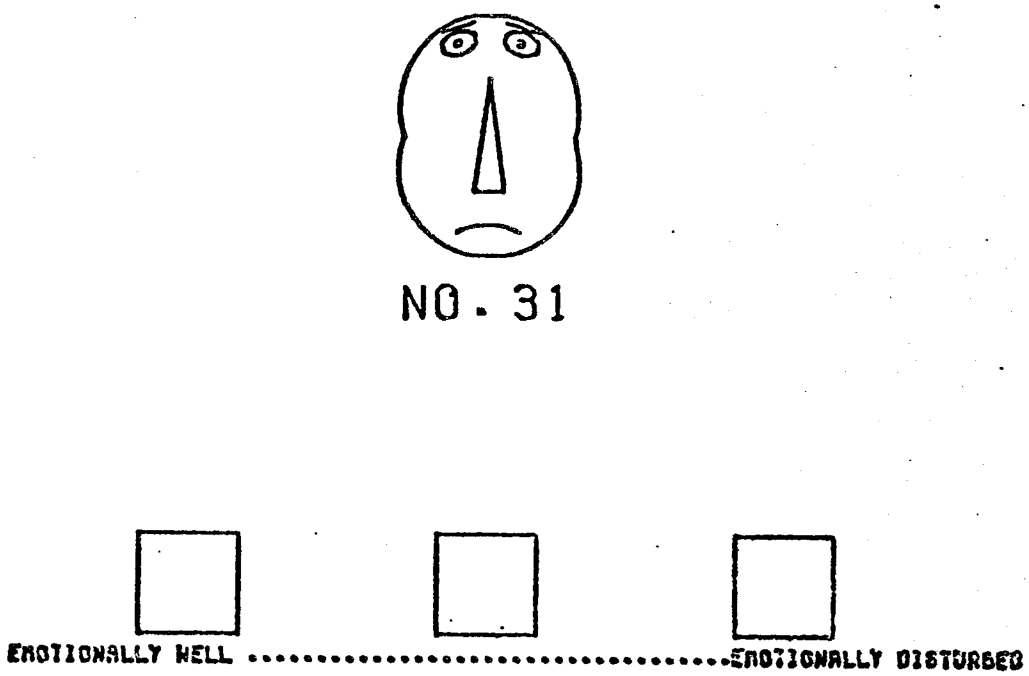

FIGURE 11. Example Face Stimulus from Experiment 6

```
-ABOVE AVERAGE NUMBER OF PHYSICAL COMPLAINTS, UNDUE CONCERN
 WITH BODILY HEALTH
-TOUCHY, UNDULY SENSITIVE, SUSPICIOUS, INCLINED TO BLAME OTHERS
 FOR OWN DIFFICULTIES

                          NO, 31




              I-----I        I-----I        I-----I
              I     I        I     I        I     I
              I     I        I     I        I     I
              I-----I        I-----I        I-----I

          EMOTIONALLY WELL......................EMOTIONALLY DISTURBED
```

FIGURE 12. Example Text Stimulus from Experiment 6

did college students.

The conclusions of this experiment are, then, unclear. While the text displays induced better performance, this turned out to be explainable by an irrelevant property they were found to possess. Nevertheless, the usefulness of faces for inducing good performance in processing Euclidean data was established by Experiment 2; and the ability of the transformation discovered in Experiment 3 to transmit data facially without training was established by Experiment 5. These continue to suggest that an improved version of Experiment 6 would indicate superiority for the facial representation.

The Orthogonal Subspace

A set of additional stimuli were appended to the facial portion of Experiment 6 in order to examine another idea. (Since these followed the regular stimuli, they did not affect subject's responses to them.) The facial representation of the MMPI disorders consisted of a five-dimensional subspace of the 18-dimensional space that represents all possible values of face parameters. There remains an orthogonal 13-dimensional subspace of facial variation. Under the linearity assumption of Experiment 3, any variation in this subspace should have no effect on the MMPI-related meaning of a facial expression. In particular, if variation in this 13-dimensional subspace were superimposed upon a face that lay at the origin of the five-dimensional d space (*i.e.*, one that was the facial depiction of a normal MMPI score), all of the resulting faces should also depict normal MMPI scores.

TABLE III. Results of Experiment 6 -- 32 Subjects

	Text	Faces
Mean correlation score	0.644	0.399
Standard deviation	0.095	0.111
Difference from chance--t(31)	38.485	20.416
Difference between means--t(62)	9.533	
Paired observations difference--t(31)	8.823	
Score using text algorithm	0.667	

Figure 13 shows the facial representation of five arbitrarily-selected, mutually-orthogonal axes lying in this 13-dimensional orthogonal subspace; they are plotted in the manner of Figures 8 and 9, with one row of faces representing movement along one axis and the origin at the center of each row. Examination of faces in this subspace suggested that the partition of the facial variation into two distinct orthogonal subspaces might be a more stable one than the further partition of the d space into its five MMPI-disorder axes, which was tested in the previous experiment. (Both of these partitions are a direct result of the regression analysis of Experiment 3.) It appeared that the two subspaces depicted an _emotional_ component and an _identification_ component of facial expressions. The latter represents variations that help distinguish the faces of particular individuals but transmit little emotional content.

To examine this notion, 10 additional faces representing normal MMPI scores plus random variation in the orthogonal space were generated. These faces all projected onto the origin of the d space. That is, while they varied along five of the axes of the orthogonal space, they contained no variation along any of the five MMPI disorder axes. If the d space indeed spanned most or all of the facial variation attributable to psychological disturbance, these faces should have been consistently rated "emotionally well," even though they contained considerable "non-psychological" variation. They can be viewed as representing psychologically-normal people of varying physical and ethnic characteristics. In order not to affect the subject's mental calibration with respect to the possible range of variation of the faces, five psychologically-abnormal faces were intermixed among these faces.

The results for this added portion of Experiment 6 were straightforward, in contrast to those above. If the three points on the rating scale are considered zero (normal), one, and two, the mean ratings for the two sets of faces are 1.575 for faces in the d space and 0.584 for faces in the orthogonal space. Chance responses would have given a mean of one, so the observed scores showed significantly ($p < 0.0005$) more psychological disturbance than chance for the d space and significantly less for the orthogonal space. Thus, the variation in the purportedly non-psychological, orthogonal subspace of facial variation was indeed perceived as having relatively little psychological meaning. This suggests that a reasonable partition of facial variation into two types may have been discovered; it could be used to construct displays that communicate in a single face both psychological and non-psychological data, each in its own subspace, with relatively little mutual interference.

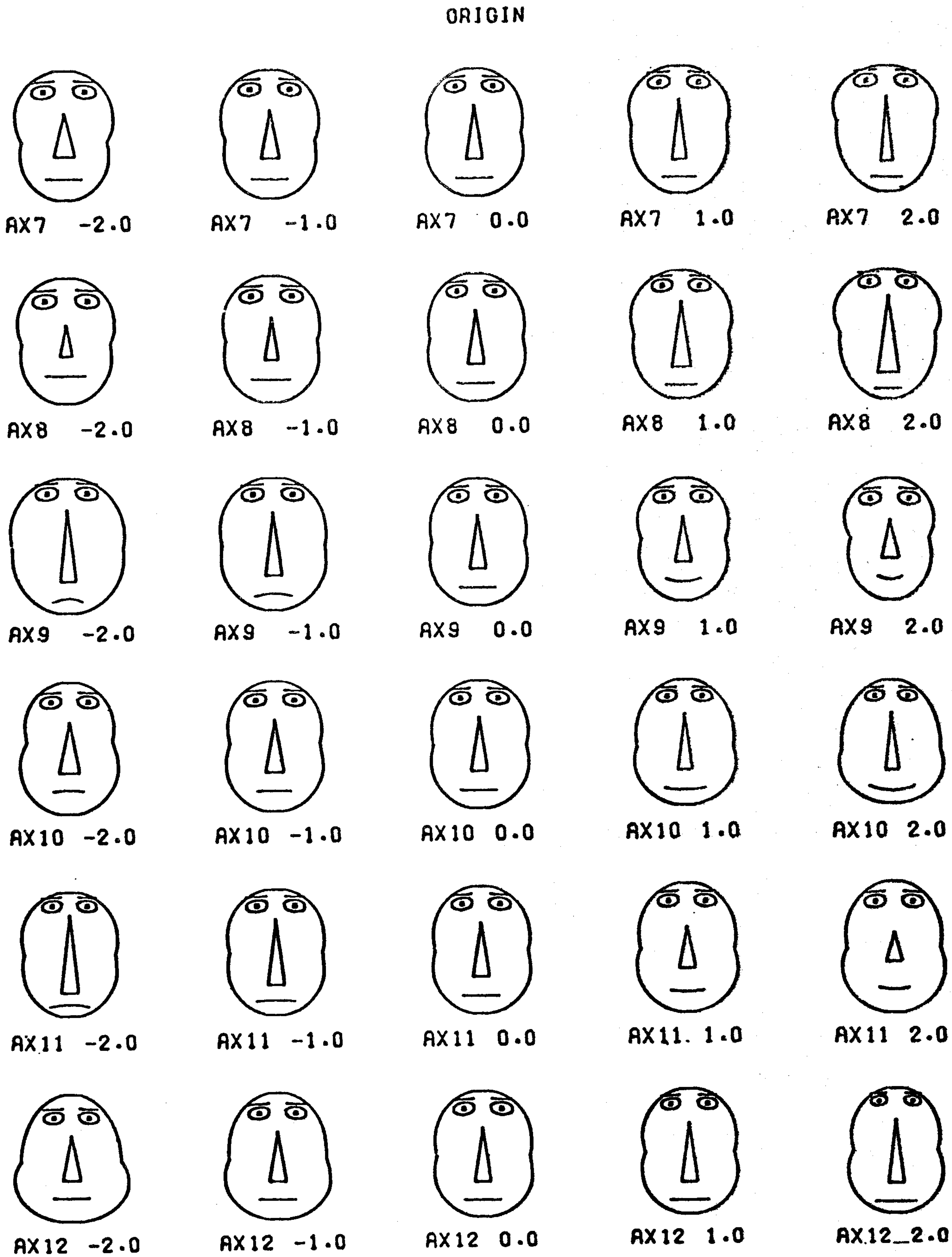

FIGURE 13. Facial Representation of the Axes of the Orthogonal Subspace

CONCLUSIONS

Two principal conclusions are drawn from this study. First, computer-produced faces are a particularly good representation for inducing superior performance of useful tasks on multivariate metrical data. Experiments with other iconic and symbolic displays indicate that it is the facial display itself, not merely the iconic mode, that accounts for this superiority. Second, the stereotype meaning already present in faces can be utilized in constructing a facial display. It was possible to measure and then exploit such meaning in order to create a demonstrably self-explanatory display for a particular set of data.

ACKNOWLEDGMENTS

Prof. William Huggins was the author's advisor while the author was a graduate student in Electrical Engineering at the Johns Hopkins University; he provided guidance and insight throughout this work. Profs. Howard Egeth and William Bevan at Johns Hopkins guided the work on the first set of experiments. Drs. Richard Hsieh and Faith Gilroy of the U. S. Public Health Service Hospital in Baltimore provided important assistance for the second set of experiments.

This research was supported by a contract between the Johns Hopkins University and the Engineering Psychology Programs, Office of Naval Research; and by the U. S. Public Health Service Hospital in Baltimore, Maryland.

REFERENCES

Anderson, E. (1960), "A Semigraphical Method for the Analysis of Complex Problems," Technometrics, 2, 387-391.

Arnheim, Rudolf (1969), Visual Thinking, Berkeley: University of California Press.

Chernoff, Herman (1971), "The Use of Faces to Represent Points in n-Dimensional Space Graphically," Stanford University, Department of Statistics, Technical report no. 71.

-------- (1973), "The Use of Faces to Represent Points in n-Dimensional Space Graphically," Journal of the American Statistical Association, 68, 361-368.

--------, and Rizvi, M. Haseeb (1975), "Effect on Classification Error of Random Permutations of Features in Representing Multivariate Data by Faces," Journal of the American Statistical Association, 70, 548-554.

Egeth, Howard (1966), "Parallel versus Serial Processes in Multidimensional Stimulus Discrimination," Perception and Psychophysics, 1, 245-252.

Entwisle, Doris R. and Huggins, William H. (1973), "Iconic Memory in Children," Child Development, 44, 392-394.

Garner, Wendell R. (1974), The Processing of Information and Structure, Potomac, Md.: Lawrence Erlbaum.

Harrison, Randall P. (1964), "Pictic Analysis: Toward a Vocabulary and Syntax for the Pictorial Code with Research on Facial Communication," Unpublished Ph.D. thesis, Michigan State University.

Hathaway, S. R. and McKinley, J. C. (1942), The Minnesota Multiphasic Personality Schedule, Minneapolis: University of Minnesota Press.

Huggins, William H. and Entwisle, Doris R. (1974), Iconic Communication: An Annotated Bibliography, Baltimore: The Johns Hopkins Press.

Jacob, Robert J. K. (1976a), "Computer-produced Faces as an Iconic Display for Complex Data," Unpublished Ph.D. thesis, Johns Hopkins University.

-------- (1976b), "PLFACE Program," Available from the author upon request.

--------, Egeth, Howard, and Bevan, William (1976), "The Face as a Data Display," Human Factors, 18, 189-199.

McKelvie, Stuart J. (1973), "The meaningfulness and meaning of schematic faces," Perception and Psychophysics, 14, 343-348.

Mezzich, Juan E. and Worthington, David R. L. (1978), "A Comparison of Graphical Representations of Multidimensional Psychiatric Diagnostic Data," in Graphical Representation of Multivariate Data, ed. Peter C. C. Wang, New York: Academic Press, 123-141.

Morrison, Donald F. (1967), Multivariate Statistical Methods, New York: McGraw-Hill.

Reed, Stephen K. (1972), "Pattern Recognition and Categorization," *Cognitive Psychology*, 3, 382-407.

Rome, H. P., Swenson, W. M., Mataya, P., McCarthy, C. E., Pearson, J. S., and Keating, R. F. (1962), "Symposium on Automation Techniques in Personality Assessment," *Proceedings of the Mayo Clinic*, 37, 61-82.

Secord, Paul F., Dukes, William F., and Bevan, William (1954), "Personalities in Faces: I. An Experiment in Social Perceiving," *Genetic Psychology Monographs*, 49, 231-279.

Siegel, J. H., Goldwyn, R. M., and Friedman, H. P. (1971), "Pattern and Process of the Evolution of Human Septic Shock," *Surgery*, 70, 232-245.

Smith, E. E. and Nielsen, G. D. (1970), "Representations and Retrieval Processes in Short-term Memory: Recognition and Recall of Faces," *Journal of Experimental Psychology*, 85, 397-405.

Tatsuoka, Maurice (1971), *Multivariate Analysis: Techniques for Educational and Psychological Research*, New York: John Wiley.

Yin, Robert K. (1969), "Looking at Upside-down Faces," *Journal of Experimental Psychology*, 81, 141-145.

--------- (1970), "Face Recognition: A Special Process?" Unpublished Ph.D. thesis, M.I.T.