

Facial feature variation: Anthropometric data II

ALVIN G. GOLDSTEIN
University of Missouri, Columbia, Missouri 65211

Facial feature variability as a function of age and gender was explored by using published anthropometric data (adopted by Goldstein, 1979) collected from several samples of faces of three racial groups. In addition, feature variability was compared to feature importance in face identification. Gender differences in variability were small except for Japanese women's faces, which displayed more variability than other samples in survey. Infants' faces might be more variable than adults' faces, but measurement problems cloud the issue. Whether feature variability is related to feature saliency is still unknown because essential information is still missing.

Using anthropological data obtained from several published sources, Goldstein (1979) documented the extent of variation of facial features across three racial populations by reporting coefficients of variation (CV) for a variety of craniofacial features obtained from Japanese, black, and white individuals. Coefficients of variation are derived readily from original measures by dividing standard deviations by their own mean values. The resultant CV is used typically by physical anthropologists to, for example, compare the variability of two (or more) populations of head circumference measurements whose mean values differ in magnitude.

Goldstein's (1979) main point, that errors in other-race face recognition is not a function of race-related differences in feature variability, was documented by a comparison of CV values from several samples of each of the three racial populations. The tabulated data contained additional information pertinent to three problems of interest to research workers in face recognition memory. This article is intended to present a summary of analyses of CV values asking whether feature variability is related to gender, age, and saliency of features in face identification.

Common sense and space limitation preclude reprinting of Goldstein's (1979) extensive set of tabulations. Nevertheless, the findings discussed here should be understandable without reference to the original tables (although the interested reader will find them helpful). In the remainder of this article, reference to "Table 1" or "Table 2" will be employed only to identify the location of data under discussion, and this reference identifies the two tables in the original article (Goldstein, 1979).

The author expresses his appreciation to Michael Robbins, James Gavan, and A. T. Steegmann, Jr. for their help in the initial stages of this project, and to June Chance for her critical reading of an earlier draft of this paper.

FACIAL FEATURE VARIATION IN MEN AND WOMEN

The literature of face recognition research occasionally refers to sex differences in recognizability, but the problem is as elusive as the solution because stable differences are difficult to demonstrate. Although no count is implied, face recognition performance data seem to favor women's faces, but this conclusion should be viewed with more than the usual amount of caution. However, if women's faces are easier to remember, perhaps this advantage is gained because women's facial features are more variable than men's. Accordingly, feature variability in men and women was explored by comparing CV values of Japanese and white men and women. Comparisons with black men and women were omitted because corresponding CV values were so limited in number.

Twenty-seven craniofacial measures, obtained from six populations of Japanese men and women (native Japanese, emigrants to Hawaii, and offspring of these emigrants) made possible 81 male-female comparisons.¹ In 53 of these, or 65%, Japanese women's facial features were more heterogeneous than Japanese men's faces (Table 1, Columns 3-8). White women's faces barely differ from Japanese men's faces in variability, but they are less variable than Japanese women's faces. Comparison of CV values across all samples of white faces indicates that women's faces might be slightly more variable than men's, but the overall difference is quite small; women's CV values are larger in 21, equal in 7, and smaller in 16 comparisons with men's samples. In summary, although all differences are small, women's faces displayed slightly more variability than men's, and Japanese women's features were the most variable of all the samples included in this limited survey. Examination of these data and additional CV values obtained from a variety of sources (not reported in

Table 1 or 2) suggests the conclusion that facial feature variability is not correlated with gender as a general rule, but appears to vary from one racial subgroup to another. But in any one particular subgroup, a stable difference in variability favoring one sex or the other may be found.

SALIENCY OF FEATURE VARIABILITY TO FACE RECOGNITION

From a Darwinian point of view, it could be argued that feature variability should be closely related to the "usefulness" of the feature as a discriminant stimulus in everyday face recognition. Stated in another way, the degree of variance of facial features should be correlated with their ecological importance as discriminant stimuli for facial recognition in ordinary interpersonal interactions. Thus, face recognition should be more dependent on a feature with a wide rather than a narrow range of variation within a population, and this prediction should be valid even though nothing is known about the magnitude of change that would be just discriminable. In any case, the CV values reported for the three racial groups provide some empirical information, albeit limited in scope, to answer the question, "which features of the face show the largest and smallest coefficients of variation?" Although much of the interpretation of the utility of the extreme CVs is speculative, its inclusion is justified because it suggests several problems for future research.

Of all the measures shown in Table 1, nose salient (a measure of relative nose protrusion) and lip thickness are the most variable. On the average, the amount the nose protrudes from nearby facial areas varies by 15% in the population from which these measures were taken. Unfortunately, nose salient was reported for only Japanese faces, so there is some reason to wonder whether measures obtained from other samples would produce equally large coefficients. Since for any single measure across all samples, CV magnitude appears to be roughly similar, it is probably safe to assume that nose-salient measures taken on other groups would be quite as large as reported here. In support of this contention, it should be noted that many of the larger mean coefficients involve the nose; for example, sagittal nose protrusion is the second most variable feature, and this measure has been obtained from a white as well as a Japanese sample.² In contrast, the smaller CV values are, with one exception, associated with head shape and size. The exception, forehead height, is quite variable; there will be more about this later.

From the small amount of evidence available, it appears that degree of feature variability is only roughly correlated with feature saliency in facial identification tasks and in tasks requiring facial description. Even though the nose seems to be one of the most variable features of the face, it does not appear to be important

in facial identification (Ellis, Shepherd, & Davies, 1975; Goldstein & Mackenberg, 1966; Lin, 1964; see Davies, Ellis, & Shepherd, 1977, for short review of pertinent literature). On the other hand, Ellis, Deregowski, and Shepherd (1975) report that when describing faces, the nose is mentioned as frequently as lips and eyebrows, a finding that suggests that some features may be salient for one cognitive process but not for another. Most of the research just cited studied full-face stimuli, and in those studies the nose may have lost its potency as a cue for identification because it is a three-dimensional feature best viewed in profile. Recall that nose protrusion was the important measure in some of the anthropological data. Only Lin's (1964) investigation speaks directly to this issue, and she found that detection of a missing nose in incomplete pictorial representations of faces was no better in profile than in front-view drawings, but her subjects were 3.5 to 5.5 years of age, so the problem is still largely without solution.

According to the limited sample of data displayed in Table 1, lip thickness might be the most variable feature of the face. Dimensions within a population could vary as much as 21% of the mean values in blacks, and 15% in whites. (Additional data, not shown here or in the original report, corroborate the accuracy of both measures, but see Footnote 1). In spite of its high level of variability, the mouth appears to be relatively unimportant for identification according to some studies (Davies et al., 1977; Ellis, Deregowski, & Shepherd, 1975; Fisher & Cox, 1975; Goldstein & Mackenberg, 1966), but other investigators report contrary findings (Lin, 1964; Seamon, Stolz, Bass, & Chatinover, 1978).

Forehead height, alluded to earlier, is another feature that is quite variable. In two studies (Davies et al., 1977; Goldstein & Mackenberg, 1966), the forehead area was demonstrated to be a salient cue for recognition, but unfortunately, in both studies, foreheads were confounded with hair and hair lines so that even a tentative conclusion will have to wait on further data. In still another investigation, where faces were described, not recognized, the forehead was found to be a relatively unimportant feature, yielding frequencies of report as low as any feature in that study (Ellis, Deregowski, & Shepherd, 1975). Interestingly, in that study, hair length and hair position were mentioned by subjects when describing faces three to five times more frequently than the forehead, which might imply that it was the hair and not the forehead that was the important cue in the two studies just cited in which both features were confounded.

The foregoing remarks have drawn attention to several interesting areas of research, but the lack of pertinent data both precludes a meaningful summary statement and emphasizes our need for additional systematic research.

FEATURE VARIABILITY AND AGE

Goldstein and Chance (1964) speculated that children's faces might be less distinguishable than adult faces because the younger face is less heterogeneous than the older face. Implications of this speculation were tested with ambiguous success, but the evidence seemed to suggest that face age was unrelated to degree of discriminability. A direct measure of heterogeneity is available from eight age-related anthropological measurements of the face reported by Herskovits (1930), and from 11 craniofacial measures tabulated by Miklashevskaya (1966). (Space limitations precluded tabulation of these extensive developmental data in this and the original article.) Analysis of the derived CV values indicate that feature variability of the very young face (1, 2, and 3 years of age) is surprisingly similar to the variation found in both teenage and adult faces. In fact, if anything, there was a slight tendency for the features of very young faces to display more variation than the adult faces, but once again, this increased variability could be an artifact caused by errors of measurement (see Footnote 2). In conclusion, examination of CV values in a variety of samples suggest that either very young faces may be more variable than adult's faces or they may be equal to adult faces in physical variability. Overall, women's and men's faces appear to differ little in feature heterogeneity, except for the unusual variability found in Japanese women's faces. Finally, whether there is indeed a relationship between a features' variability and its importance or usefulness in face identification is still very much an open question.

REFERENCES

DAVIES, G., ELLIS, H., & SHEPHERD, J. Cue saliency in faces as assessed by the "Photofit" technique. *Perception*, 1977, 6, 263-269.

ELLIS, H. D., DEREGOWSKI, J. B., & SHEPHERD, J. W. Descriptions of white & black faces by white & black subjects. *International Journal of Psychology*, 1975, 10, 119-123.

ELLIS, H. D., SHEPHERD, J. W., & DAVIES, G. M. An investigation of the use of the Photofit technique for recalling faces. *British Journal of Psychology*, 1975, 66, 29-37.

FISHER, G. H., & COX, R. L. Recognizing human faces. *Applied Ergonomics*, 1975, 6, 104-109.

GOLDSTEIN, A. G. Race-related variation of facial features: Anthropometric data I. *Bulletin of the Psychonomic Society*, 1979, 13, 187-190.

GOLDSTEIN, A. G., & CHANCE, J. E. Recognition of children's faces. *Child Development*, 1964, 35, 129-136.

GOLDSTEIN, A. G., & MACKENBERG, E. J. Recognition of human faces from isolated facial features: A developmental study. *Psychonomic Science*, 1966, 6, 149-150.

HERSKOVITS, M. J. *The anthropometry of the American Negro*. New York: Columbia University Press, 1930.

LIN, Y. *The perception of human face: A developmental study of stimulus potency of different facial features*. Unpublished Doctoral dissertation, University of Rochester, 1964.

MIKLASHEVSKAYA, N. M. Growth of the head and face in boys of various ethnic groups in the USSR. *Human Biology*, 1966, 38, 231-250.

SEAMON, J. G., STOLZ, J. A., BASS, D. H., & CHATINOVER, A. I. Recognition of facial features in immediate memory. *Bulletin of the Psychonomic Society*, 1978, 12, 231-234.

SHAPIRO, H. L. *Migration and environment*. New York: Oxford University Press, 1939.

NOTES

1. These data were collected by Shapiro (1939). Only 22 measures appear in Table 1; the 5 not shown were omitted because corresponding data on white faces were not available.
2. Some caution is advised in the interpretation of the nasal and lip data. Random measurement errors may be inflating the nasal coefficients relatively more than those found in other features because anatomical landmarks on and around the nose are difficult to locate. Also, nasal and lip dimensions are always small, making them proportionately more influenced by small errors of measurement than would be the case for other, larger features. Lip thickness measures have the added problem of dimensional "drift" over time; lip thickness can change as a function of aging, a process that would increase variability of measurement for any sample composed of individuals widely differing in age.

(Received for publication January 15, 1979.)