Sex Determination by Discriminant Function Analysis of Crania

EUGENE GILES AND ORVILLE ELLIOT
Department of Anthropology, Harvard University, Cambridge, Massachusetts

Almost a century ago Dureau (1873, p. 487) reviewed the literature on the sexing of human crania for the first part of the nineteenth century and concluded that ". . . les sexes peuvent être aisément distingués dans la très-grande majorité des cas à l'aide d'un ensemble ou d'un groupe de caractères plus ou moins accentués; mais on rencontre toujours . . . un certain nombre de cas douteux, qu'on peut évaluer à environ un dixième." Since that time the "doubtful 10%" has prompted many more studies utilizing diverse techniques aimed at a more nearly perfect method of sexing skulls, but none has achieved significantly greater accuracy. Ability in cranial sexing claimed by different authorities shows considerable variation. While it is pointless to review all of these estimates, a few will give some indication of what reliability may be expected from traditional visual methods. Borovanský ('36) examined 247 skulls of known sex and opined that from observational criteria only 10% are impossible to sex. Hrdlička (Stewart, '52) believed approximately 80% of skulls without mandibles could be accurately sexed. Hooton ('43) suggested that even with the whole skeleton, sexing was possible in only 90% of cases. Keen ('50) has given a valuable survey of the statistical significance of many cranial features indicative of sex. By use of a combination of four measurements and three observations, he was able to sex correctly about 85% of his material. Ceballos and Rentschler ('58), using four measurements on cranial x-rays of the living, were able to sex 88% of 200.

Two exceptionally well-qualified physical anthropologists have given estimates of how well they did in blind tests on crania of known sex. Krogman ('49, '62) correctly identified 82–87% of 750 cranial specimens from the Todd collection of known-sex skeletal material. Stewart ('48) sexed accurately 77% of a smaller series using crania alone. Add to these results the statement of Howells ('41, p. 113) that ". . . where two observers examine the same series there is generally a disagreement between them as to the sex of ten to fifteen per cent of the individuals" and a situation emerges wherein the authors sense the potential value of a purely metrical method of sexing crania which provides an accuracy of 82–89% while requiring an investigator armed only with calipers and a knowledge of the more common cranial landmarks.

The approach chosen is that of the multivariate linear discriminant function. This concept was first introduced by R. A. Fisher in an anthropological context (Barnard, '35) and since then has found many applications. Discriminatory analysis attacks the problem of assigning an individual to a sample classified into two or more groups on the basis of some number p variables characteristic of the individuals comprising the sample. It does not attempt to sort out heterogeneous material, nor does it usually reserve judgment but will assign an individual to one group or other on the basis of the available information. In problems of sexing dealt with in this paper, the number of groups is limited to two, the variables are anthropometric measurements, p equals four to nine of these measurements, and the discriminatory function derived is that linear function most efficaceous in distinguishing the two groups, i.e., the two sexes. Efficaceous here means essentially that the number of individuals misclassified is minimized. Among others, Thieme and Schull ('57), Thieme ('57), Pons ('55), Hanihara ('58, '58), and Giles and Elliot ('62a,b) have utilized discriminant functions in the determination of sex and race from the skeleton.

One way of describing multivariate linear discriminatory analysis is by first considering regression analyses, which can be defined as the prediction of the value of one variable from the values of other given variables (Williams, '59). Ordinary linear regression involves the straight-line relation between two variables, one treated as causal, or independent, and the other in some degree resultant, or dependent. For example, it may be of interest to predict stature from femur length. In this case, stature is the dependent variable and femur length the independent variable upon which stature depends. The linear relation between the two can be expressed in the form of the simple regression equation: $y = a + bx$. In this example, femur length $= x$, stature $=$ y, and the coefficients a and b provide the best prediction of y from x. Multiple regression can be thought of as an extension of the idea of simple linear regression to the case of the regression of one dependent variable on several independent variables, e.g., stature on femur length and tibia length. The need for multiple regression is often found in the lack of a close regression relationship between the dependent variable and any single independent variable.

The multivariate linear discriminant function can be looked upon as the solution of a multiple regression problem (Hodges, '55). In the investigation at hand, each sex was given an arbitrary value (0 for males, 1 for females) and sex was treated as an artificial variable. This artificial variable is taken as the dependent variable in a multiple regression where the anthropometric measurements form the independent variables. It is interesting to note that a variable which by itself has little discriminatory value may heighten the power of another. The final result is that the p measurements on an individual are replaced by a single measurement called the discriminant function score, which is arrived at by summing the p measurements, each weighted by the appropriate coefficient. These multiple regression coefficients, which form the discriminant function, were calculated on electronic computers according to a design expressed in Kendall ('57). We see no purpose in detailing computation methods, since they are handled at length in the references cited above. Our purpose is to present results that will be of practical use for anthropologists.

If the discriminant function has done its job, the single measurements or discriminant function scores, each of which replaces several measurements for an individual, will be distributed for the whole series into two groups with minimum overlap. We can determine the mean value of the discriminant function scores for males by taking the mean male value for each measurement and entering them into the discriminant function. If this is likewise done for the females, the arithmetic mean of the two scores provides a sectioning point to use when we have no a priori reason to believe that a specimen is more likely to be male, say, than female (Kendall, '57). Following this procedure, we will say that any specimen falling on the side of this line towards the male mean will be called male, and any specimen falling on the other side will be called female. So doing should minimize the probability of misclassification.

To determine how effective our discriminant functions have been we have calculated the discriminant function scores for each of our specimens. If the scores for each sex are normally distributed (as they should be if we are dealing as assumed with a p-variate normal population), we can make two statements concerning the probability of misclassification based on properties of the normal distribution. These statements are of course separate from the empirical determination of misclassification based on the results from our sample. On the basis of tests described later, we feel the assumption of normality to be justified. Following this assumption, we have calculated the probability of misclassification as that area in the tail of a normal distribution which represents those specimens of one sex wrongly classified when allocated by the criteria given above (Kendall, '57).

A glance at the results of the discriminatory analysis will show that the possibility of non-significance of discrimination is out of the question and hence was not formally tested. Rather we have given an estimate of the degree of certitude that a specimen found on the male side, say, of the sectioning point is in fact a male. A probability of 19:1 that a specimen on the male side of the sectioning point is a male has been estimated by determining the point along the abscissa at which the ordinates of the two normal distributions are in a ratio of 19:1 in favor of the male. This particular position on either side of the sectioning point has been located for all the discriminant functions presented.

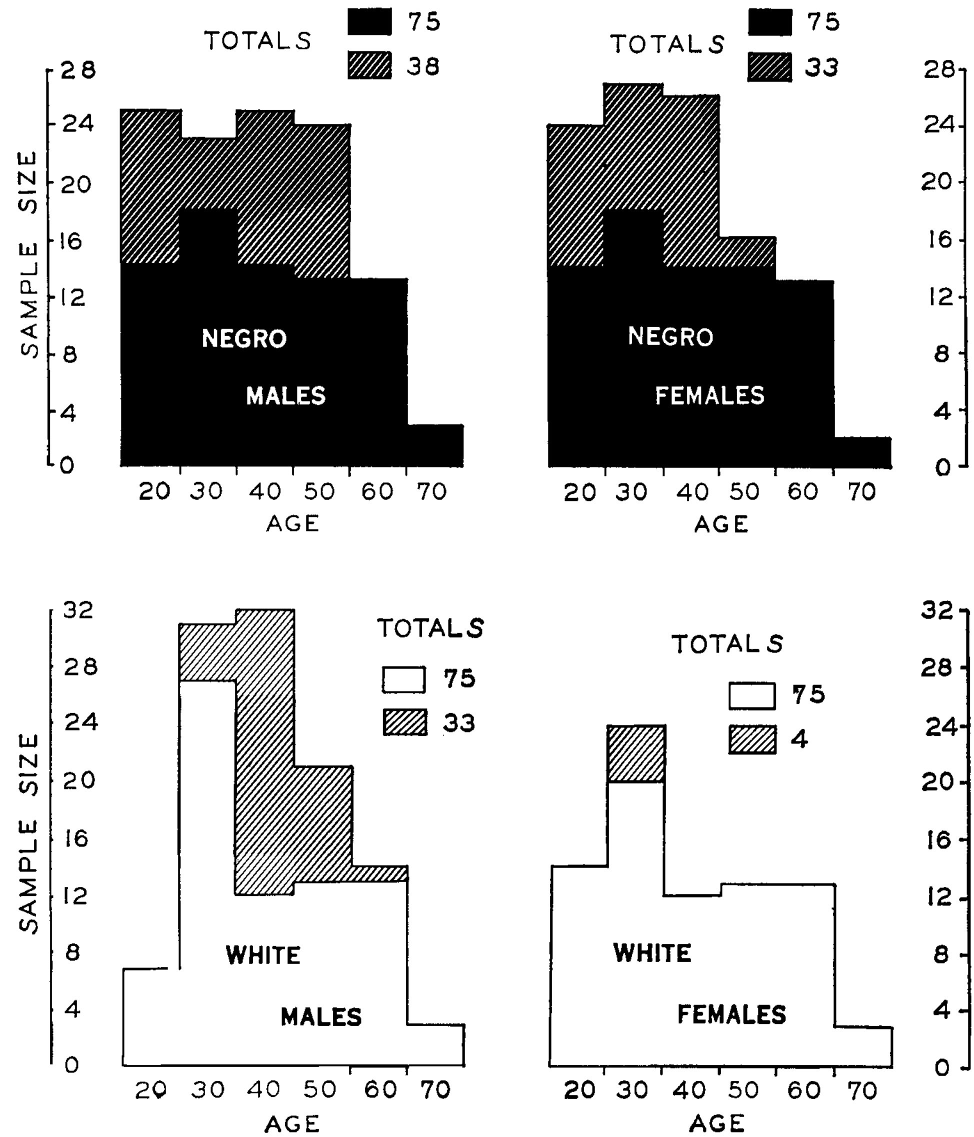

Fig. 1 Age distribution by decade for Negro and white sample. Shaded areas represent check specimens.

SAMPLE

A total of 408 skulls from the Terry collection, Washington University School of Medicine, St. Louis and the Todd collection, Western Reserve University School of Medicine, Cleveland were examined for this study. The problems presented by a medical school cadaver sample such as ours are many. Aside from the possible effects of the inherent socio-economic bias on skeletal morphology, the sample is too old and too male to be representative. The age and race data may not be exact. We had difficulty in securing white female skulls that permitted palatal measurements. Nevertheless, the Terry and the Todd collections provide an abundance of material that helps minimize some of these difficulties. And without question, any serious study of sex differences in the skeleton can only be based on collections such as these, where the sex is positively known from written records.

The age assignment of each specimen was taken from the records. In a few cases this had been estimated at death or given as a span of a few years. We have taken the age as given (or the mid-point of a span) with due appreciation of possible error. No other aging scheme, such as from the material itself, is likely to provide more specific information for the few cases recorded only as spans of five or ten years. All specimens were chosen to fall within the age range of 21 to 75 years. Figure 1 breaks age into six categories for each race and sex combination. The mean age for the total sample of 408 was 43.3 years with a standard deviation of 13.2 years; not, we feel, too old to vitiate the analytic results.

Both whites and Negroes were included in the sample. The sex and race breakdown is as follows: 108 white males, 79 white females, 113 Negro males, 108 Negro females. For all the statistical analyses, 75 specimens were chosen from each of these categories to enter into the computations; the rest were reserved as a check sample. Within each age decade, the specimens allotted the check sample were chosen randomly. The Negro specimens we so designated on the collection records, which reflected of course American cultural standards, not genetic ones. It seems reasonable to assume that any person exhibiting phenotypic evidence of Negroid admixture was considered a "Negro." Our Negro sample thus has an indefinite white American and possibly American Indian component, while our white sample represents American whites of European descent.

MEASUREMENTS

The measurements were chosen from the point of view of their ease in taking and their potential sex-discrimination power. In the latter consideration we were guided especially by the remarks of Stewart ('52), Hooton ('46), Keen ('50), and Parsons and Keene ('19). Eleven measurements were made on the 408 specimens by the senior author, but only nine of these were used in the discriminant function analyses. Table 1 gives the means and standard deviations of all the measurements and the age for the sample of 75 of each race and sex combination. In addition, table 1 presents the critical ratios (CR) computed for all measurements for the difference between means of the following pairings: white males and females, Negro males and females, male whites and Negroes, female whites and Negroes. The absolute size of the critical ratio of course varies with the sample size, but the relative differences among the measurements in the value of the critical ratio suggest how well these measurements distinguish race and sex. In every measurement for both races the male dimensions significantly exceed those of the females. If the female values for each measurement are adjusted by the same proportion necessary to eliminate the difference between the Negro and the white male means, then the Negro females exceed the white females in all measurements except maximum cranial width, prosthion-nasion height, and mastoid length. In racial differences, for both sexes, the whites exceed the Negroes only in maximum cranial width, basion-bregma height, and opisthion-forehead length. The whites have a higher and broader cranium than do the Negroes, but in terms of upper facial height, cranial base length, and facial projection, the Negroes have the greater dimensions.

TABLE 1

Means and standard deviations of four race and sex samples of 75 each and critical ratios (CR) of samples paired by race and sex for 11 measurements and age

	White males			White females			Negro males			Negro females		
	M	σ	CR with Negro ♂♂	M	σ	CR with White ♂♂	M	σ	CR with Negro ♀♀	M	σ	CR with White ♀♀
Glabello-occipital length	181.333	6.844	4.20	171.453	6.634	8.98	185.893	6.439	7.76	177.840	6.272	6.06
Opisthion-forehead length	156.600	5.307	3.06**	148.400	5.009	9.73	153.947	5.304	8.31	147.027	4.891	1.70†
Maximum width	143.013	6.178	4.06	138.707	5.675	4.45	139.347	4.788	6.04	134.053	5.892	4.93
Basion-bregma height	134.320	5.463	2.53*	127.453	4.981	8.04	132.093	5.332	5.80	126.680	6.078	0.85†
Basion-nasion	100.600	4.128	1.19†	95.067	4.325	8.02	101.467	4.751	6.99	96.307	4.274	1.77†
Maximum diameter bi-zygomatic	131.920	5.273	1.54†	122.707	5.117	10.86	133.253	5.347	10.95	124.400	4.522	2.15*
Basion-prosthion	95.400	6.703	7.13	90.480	5.589	4.88	102.960	6.275	4.44	98.720	5.393	9.19
Prosthion-nasion height	70.760	4.436	3.54	66.333	4.132	6.32	73.347	4.515	7.28	68.067	4.368	2.50*
Nasal breadth	24.267	2.042	8.39	23.120	2.292	3.24**	27.227	2.272	3.11**	26.147	1.978	8.66
Palate — external breadth	59.867	3.862	9.06	57.027	4.043	4.40	65.387	3.592	4.64	62.800	3.233	9.66
Mastoid length	28.067	2.675	5.07	25.213	2.748	6.44	30.320	2.762	8.99	26.347	2.651	2.57*
Age	44.6	13.7		44.6	14.1		44.3	14.1		43.9	14.3	

$CR = (M_1 - M_2)/\sqrt{\sigma_1^2/N_1 + \sigma_2^2/N_2}$.
** and * = significant at 0.01 and 0.05 level, respectively.
† = not significant.

Except for opisthion-forehead length and mastoid length, the measurements follow those given in Hooton ('46). The measurements require only the spreading and the sliding calipers, and are taken to the nearest millimeter. The designations given below are Hooton's; his text should be consulted if amplification is desired in the definition.

Glabello-occipital length. Maximum length of the skull, from the most anterior point of the frontal in the midline to the most distant point on the occiput in the midline.

Maximum width. The greatest breadth of the cranium perpendicular to the median sagittal plane, avoiding the supramastoid crest.

Basion-bregma height. Cranial height measured from basion (midpoint on the anterior border of the foramen magnum) to bregma (intersection of the coronal and sagittal sutures).

Maximum diameter bi-zygomatic. Maximum width between the lateral surfaces of the zygomatic arches measured perpendicular to the median sagittal plane.

Prosthion-nasion height. Lowest point on the alveolar border between the central incisors to nasion (midpoint of the naso-frontal suture).

Basion-nasion. From basion to nasion.

Basion-prosthion. From basion to the most anterior point on the maxilla in the median sagittal plane.

Nasal breadth. Maximum breadth of the nasal aperture perpendicular to nasal height.

Palate — external breadth. The maximum breadth of the palate taken on the outside of the alveolar borders.

Opisthion-forehead length. The maximum distance from opisthion (the midpoint on the posterior border of the foramen magnum) to the forehead in the midline.

Mastoid length. (Based on Keen, '50.) The length of the mastoid measured perpendicular to the plane determined by the lower borders of the orbits and the upper

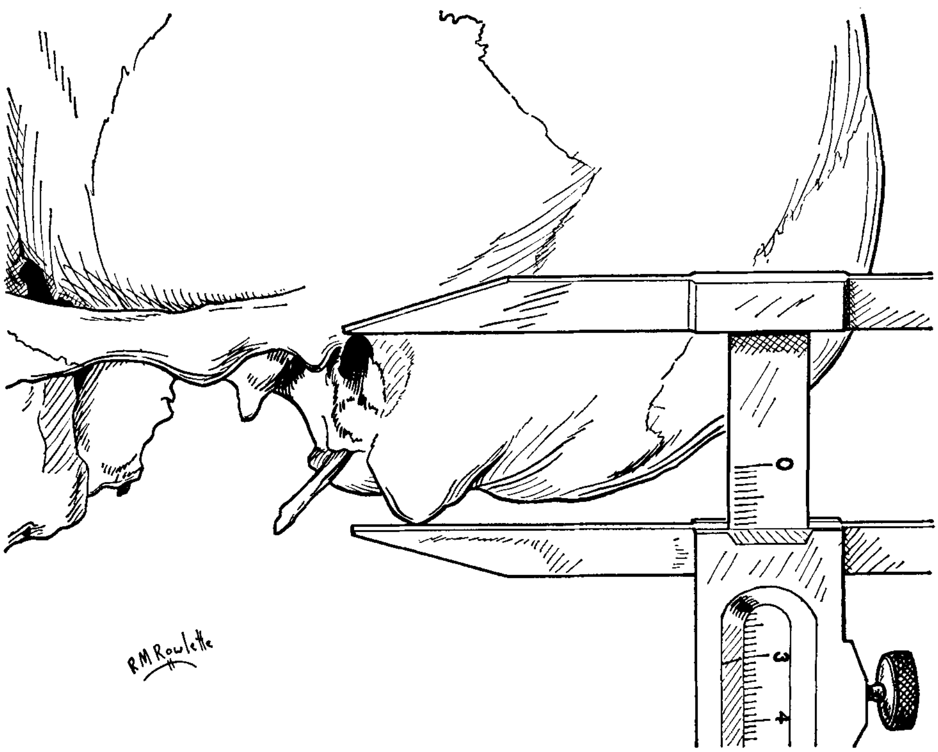

Fig. 2 Mastoid length measurement technique.

borders of the auditory meatuses (Frankfort plane). The upper arm of the sliding calipers is aligned with the upper border of the auditory meatus and the distance (perpendicular to the Frankfort plane) to the tip of the mastoid is measured (see fig. 2).

DISCRIMINANT FUNCTIONS

The sample, measurements and technique necessary for the calculation of the discriminant functions have already been described. From the Negro and white males and females samples of 75 each were chosen with an age distribution as indicated in figure 1. Then from combinations of nine cranial measurements a total of 21 discriminant functions were derived to indicate sex in whites, Negroes, and whites plus Negroes. The variations in the combinations of measurements employed should aid an investigator in finding a discriminant function that will fit the measurements possible on his specimen. Table 2 presents the basic data for the 21 discriminant functions we have calculated. Each is given a number in the first row of the table. The second row indicates the race or races (white, Negro, or white and Negro) upon which the discriminant function is based. The third row shows the number of measurements used. Below these are nine rows representing the nine measurements. If a discriminant function uses a particular measurement, the coefficient by which the measurement must be multiplied will be found opposite the measurement designation. These coefficients are then placed in a formula or equation form similar to that used in the example given later. To apply the discriminant function, each measurement on the skull in question is multiplied by its appropriate coefficient and the products added or substracted as indicated.

In table 2 the first row below those of the measurements, called "Male .05 level," gives that value of the discriminant function score (the sum of the coefficients times the measurements) at which the probability is 19 to 1 (i.e., .95) that the specimen is a male. Any score greater than this is even more likely male. The next row gives the mean score for males. The sectioning point is halfway between the mean score for males and the mean score for females. A score greater than the sectioning point is more probably male, one less than the sectioning point is more likely female. The final row gives that value which represents a probability of .95 that the score represents a female. Less than this is, of course, even more likely to be female. Figures 3–5 present graphically these points for representative discriminant functions based on whites (discriminant function no. 4), Negroes (no. 5), and Negroes and whites (no. 6) superimposed on the actual scores made by the sample of 187 whites, 221 Negroes, and 408 total.

Tepexpan man, a probable Paleo-Indian from Mexico, has been described and measured by De Terra, Romero and Stewart ('49), who provide sufficient data for this specimen to serve as an example of the discriminant function technique in use. Tepexpan man is also of interest since recently Genovés ('60) has suggested that the true sex is female. The following measurements are taken from De Terra, Romero and Stewart ('49, p. 107) except mastoid length, which was measured on a cast in the Peabody Museum: glabello-occipital length, 179 mm; maximum width, 143 mm; basion-bregma height, 136 mm; maximum diameter bi-zygomatic, 140 mm; basion-nasion, 94 mm; and mastoid length, 28 mm. These values are then entered into discriminant function no. 9, which does not require the presence of prosthion:

$$5.538(179) + 2.308(143) + 10.308(136) + 1.000(94) + 21.538(140) + 22.154(28) = 991.302 + 330.044 + 1401.888 + 94.000 + 3015.320 + 620.312 = 6452.866.$$

When the value of 6452.866 is compared with the figures given in table 2 for no. 9, Tepexpan man is seen to be quite definitely male, so much so that he exceeds the .95 level of probability. In other words, the chances are greater than 19 to 1 that the individual is male. This result must be tempered by what will be said later concerning the application of these functions to individual American Indian specimens, but even so, the above result in view of the general configuration and size of the specimen's measurements would lead one to suspect Genovés's assessment.

TABLE 2

Coefficients, sectioning points, means and 0.05 probability levels for 21 discriminant functions

Discriminant function number	1	2	3	4	5	6	7	8	9	10
Race	W	N	W & N	W	N	W & N	W	N	W & N	W & N
Number of measurements	8	8	8	8	8	8	6	6	6	6
Glabello-occipital length	3.107	9.222	6.083	3.400	3.895	4.692	1.800	3.533	5.538	5.550
Maximum width	−4.643	7.000	−1.000	−3.833	3.632	1.000	−1.783	1.667	2.308	−1.000
Basion-bregma height	5.786	1.000	9.500	5.433	1.000	8.769	2.767	0.867	10.308	
Basion-nasion				−0.167	−2.053	4.615	−0.100	0.100	1.000	4.100
Maximum diameter bi-zygomatic	14.821	31.111	28.250	12.200	12.947	21.308	6.300	8.700	21.538	19.800
Basion-prosthion	1.000	5.889	2.250	−0.100	1.368	−4.385				
Prosthion-nasion height	2.714	20.222	9.917	2.200	8.158	7.385				7.600
Palate — external breadth	−5.179	−30.556	−19.167							−10.400
Mastoid length	6.071	47.111	25.417	5.367	19.947	21.077	2.833	14.367	22.154	
Male 0.05 level	2799.21	8522.44	6519.83	2711.23	4246.26	6245.69	1357.03	2619.80	6399.85	3867.50
Male mean	2779.66	8487.56	6466.17	2687.53	4223.20	6174.82	1343.54	2608.60	6319.28	3813.61
Sectioning point	2676.39	8171.53	6237.95	2592.32	4079.12	5972.03	1296.20	2515.91	6119.50	3686.41
Female mean	2573.12	7855.50	6009.72	2497.10	3935.03	5769.25	1248.86	2423.22	5919.72	3559.20
Female 0.05 level	2550.29	7825.78	5961.92	2478.73	3897.84	5712.62	1241.15	2394.80	5846.77	3508.25

Discriminant function number	11	12	13	14	15	16	17	18	19	20	21
Race	W	N	W & N	W	N	W & N	W & N	W & N	W	N	W & N
Number of measurements	6	6	6	4	4	4	4	8	5	5	5
Glabello-occipital length				1.236	2.111	2.184	3.833	4.850	9.875	2.867	1.165
Maximum width				−1.000	1.000	1.000	1.000	−1.500			
Basion-bregma height								6.150			
Basion-nasion	10.714	1.000	2.417					−0.100	7.062	−0.100	1.659
Maximum diameter bi-zygomatic	16.381	19.389	5.867	3.291	4.963	6.224	11.267	19.350	19.062	12.367	3.976
Basion-prosthion	−1.000	2.778	−0.100					1.050	−1.000	−0.233	−1.000
Prosthion-nasion height	4.333	11.778	2.483				4.067	7.150	4.375	6.900	1.541
Palate — external breadth	−6.571	−14.333	−3.567					−9.900			
Mastoid length	14.810	23.667	5.867	1.528	8.037	6.122					
Male 0.05 level	3537.62	3630.78	1151.83	565.79	1448.76	1570.43	2679.00	4099.50	5308.00	2682.43	935.75
Male mean	3472.30	3615.33	1138.92	558.22	1436.80	1546.15	2631.86	4052.74	5230.05	2652.71	920.55
Sectioning point	3348.27	3461.46	1094.99	536.93	1387.72	1495.40	2551.52	3922.26	5066.69	2568.97	891.48
Female mean	3224.24	3307.59	1051.06	515.63	1338.64	1444.64	2471.17	3791.78	4903.33	2485.22	862.41
Female 0.05 level	3181.43	3288.56	1037.33	509.72	1316.72	1419.82	2429.00	3745.75	4854.25	2453.43	849.99

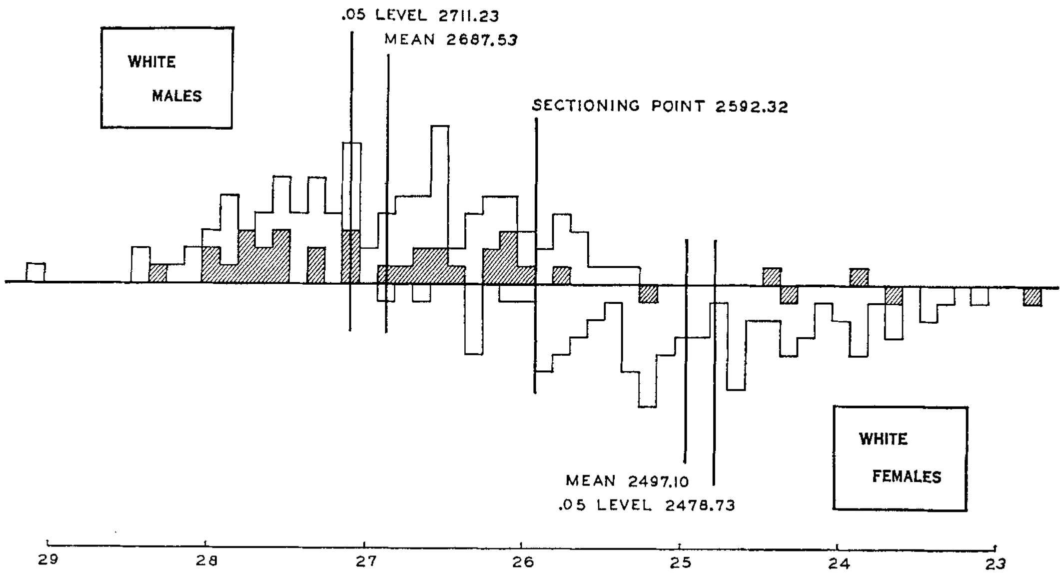

Fig. 3 Distribution of scores for 187 white specimens for discriminant function no. 4. Shaded areas represent check specimens.

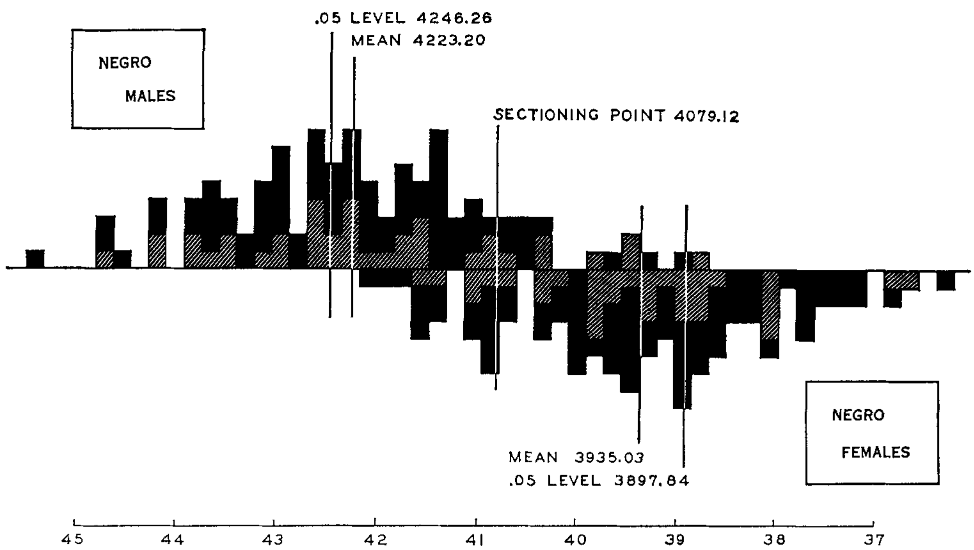

Fig. 4 Distribution of scores for 221 Negro specimens for discriminant function no. 5. Shaded areas represent check specimens.

A natural question to ask of a discriminant function approach is how accurate are the results. In the present situation, this question has two aspects. One concerns the accuracy on the material which formed the basis of the examination. Here we can empirically determine how well the material was sexed, since the specimens are of known sex, and we can also compute on the basis of properties of normal distributions what might be expected in a statistical sense. But it is also of importance to estimate just how reliable the discriminant functions may be when

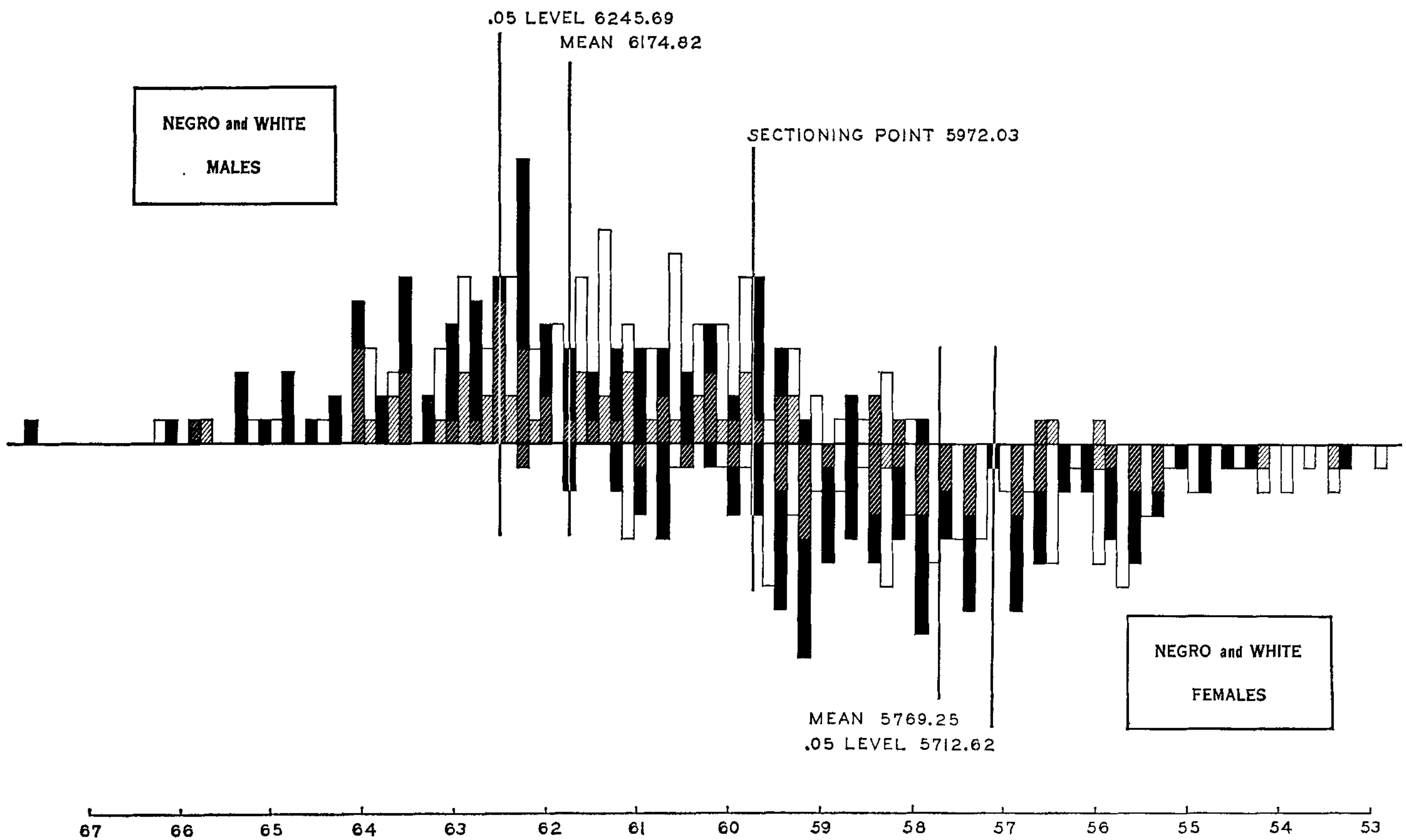

Fig. 5 Distribution of scores for 408 Negro and white specimens for discriminant function no. 6. Shaded areas represent check specimens.

used on individuals that are obviously not members of the populations upon which the statistics were based. We are asking to what extent do these discriminant functions reflect universal human morphological differences between the sexes, and to what extent do they merely represent morphological differences characteristic of St. Louis and Cleveland whites and Negroes likely to turn up in a dissecting room. Answering this latter question is made difficult by the paucity of known sex material for comparative purposes.

The calculations of the discriminant functions were made on samples of 75 of each race and sex combination. This left a little better than a fourth of the total sample as an independent check. Table 3 lists the percentages correctly classified by each of the 21 discriminant functions, along with the sample size upon which these percentages are based. These figures indicate that the level of accuracy is 82–89% based on actual tests with known sex material.

If the distributions of the discriminant function scores for males and for females can each be said to be normal, we may then determine statistically the probability of misclassification. It seemed rather too easy merely to assert that the distribution of the scores must be normal. To provide evidence for this assumption, we undertook a rigorous test of normality for those distributions pictured in figures 3–5. According to the procedure outlined by Pearson and Hartley ('58, pp. 61–63, 183), which involves calculating the third moment about the mean, the sex distributions show insufficient kurtosis and skewness to reject at the .05 level the null hypothesis that these distributions are normal. The only exception is that the white male distribution (fig. 3) shows a positive skewness exactly at the .05 level, i.e., only 1 in 20 times would a sample of this size from a normally distributed population show such skewness by chance. On the other hand, the combined Negro and white males sample (fig. 5) shows an extraordinarily close fit, the value of a being 0.7983 (normal = 0.7979, σ = 0.0095) and $\sqrt{b_1}$ being 0.00075 (normal = 0, σ = 0.1147).

We believe the test results justify our presenting probabilities of misclassification based on normal distributions. Table 3 shows then, in addition to the observed classification, the expected amount of misclassification based on the means and

TABLE 3

Percentage correctly sexed by 21 discriminant functions and estimated probability of misclassification

Discriminant function number	Whites			Negroes			White + Negro	Calculated probability of misclassification
	Basic sample N = 150	Test sample N = 37	Total N = 187	Basic sample N = 150	Test sample N = 71	Total N = 221	Total N = 408	
1	86.0	86.5	86.1					13.4
2				85.3	83.1	84.6		12.4
3	86.7	86.5	86.6	85.3	84.5	85.1	85.8	13.6
4	88.0	91.9	88.8					13.6
5				84.0	81.7	83.3		13.4
6	88.0	89.2	88.2	84.0	84.5	84.2	86.0	14.5
7	86.7	91.9	87.7					13.6
8				86.0	80.3	84.2		13.5
9	84.0	86.5	84.5	82.0	83.1	82.4	83.3	15.1
10	82.0	86.5	82.9	83.3	85.9	84.2	83.6	15.4
11	86.0	89.2	86.6					15.5
12				84.7	88.7	86.0		12.5
13	84.0	86.5	84.5	86.0	88.7	86.9	85.8	14.4
14	83.3	86.5	84.0					14.5
15				85.3	81.7	84.2		14.7
16	82.0	78.4	81.3	83.3	83.1	83.3	82.4	16.0
17	82.0	83.8	82.4	82.0	83.1	82.4	82.4	16.5
18	85.3	83.8	85.0	82.7	83.1	82.8	83.8	14.9
19	83.3	86.5	84.0					15.1
20				83.3	83.1	83.3		15.0
21	84.0	83.8	84.0	82.0	87.3	83.7	83.8	15.9

standard deviations of the discriminant function score distributions for each of the 21 variations. These probabilities are determined as that percentage of the total area under a normal curve (with the appropriate mean and standard deviation) to the right of the sectioning point in the case of males, to the left in the case of females. Thus the probability of misclassification for males, say, gives the frequency with which we would expect to misclassify males. If we have a population wherein the estimated sex ratio is 1:1, the probability of misclassifying specimen X from this population is the sum of the probability of X being a male and misclassifying him and X being a female and misclassifying her. This sum would be the arithmetic mean of the two misclassification probabilities, which is given in the final column in table 3 and represents the error expected from the discriminant functions. It can be seen that the predicted and the observed misclassification errors are quite similar.

The evidence presented in the preceding discussion demonstrates that the discriminant function technique is a valuable tool in the determination of sex from the cranium. This is the part of the skeleton most often in question in medicolegal work (Krogman, '49), besides requiring identification to supplement archeological interpretation and description. The discriminant functions provide an accuracy easily comparable with visual sexing, yet require no value judgments based on interpretations of qualitative characteristics. There is no problem of associating the mandible with the cranium, since it is not used, while the variation in measurements utilized should provide necessary flexibility with damaged skulls. Even though the work has been based on American Negroes and whites, it should not be overlooked that discriminant functions based on the two races combined do, for practical purposes, equally well for both races as do those based on and applied to a single race. This important result suggests that the sex discriminant function is employing basic differences and relationships in cranial morphology which are largely independent of racial variation. These functions should be applied to groups of other racial composition to test this proposal.

TESTS AND APPLICATIONS

If it is true that these sexing discriminant functions are based on criteria that are not specific to the samples from which they were derived, they would obviously have a much greater utility. As an experiment to push the supposition of universality to an extreme, we have applied a discriminant function based on whites and Negroes (no. 10) to a sample of 85 adult chimpanzees. These chimpanzees, 46 males and 39 females (according to collectors' records) from the Peabody Museum collection, had the following measurements taken on them by the senior author: glabello-occipital length, maximum width, basion-nasion, maximum diameter bi-zygomatic, prosthion-nasion height and palate — external breadth (Giles, '56). With a sectioning point (3509.39) midway between the mean male chimpanzee and the mean female chimpanzee scores (3637.56 and 3381.22), four females and five males were misclassified, leaving 89.4% correctly assigned. When the chimpanzees were scored on a discriminant function calculated from the same six measurements taken on the chimpanzees themselves, 11 were misassigned. On this evidence, the function based on human cranial measurements, with an adjusted sectioning point, does very slightly (but undoubtedly not significantly) better in sex discrimination than does a formula based on the specimens themselves. The unexpected accuracy of the human function on apes raises the possibility of using the discriminant function approach as an aid in determining the sex of hominoid fossils.

To calculate discriminant functions for sex differentiation with any confidence, a series of known sex must be available. Likewise, other known-sex series should be used for reliable testing. These are disappointingly rare. Consequently, in this paper we stress the application of those functions we have calculated rather than the statistical technique of calculation itself. Our functions have been tested on two human populations, early Irish whites and American Indians. Neither of the samples of these populations is strictly of known sex, but in both cases we have better than average data concerning the probable sex of the specimens. American

Indians are of special interest, since they are often disinterred by archeologists and their skeletons continue to accumulate in museums. Perhaps the first question to be asked about a human skeleton it its sex. Yet archeologists may not be particularly confident of their ability to sex visually skeletal remains, nor have ready access to a physical anthropologist, nor be able to preserve and store all the skeletal material they excavate. And occasionally collectors or other individuals who have found a cranium may be referred to the "local expert" — often an archeologist rather than a physical anthropologist. There is no question that in the hands of an expert the pelvis provides a surer identification of sex, yet in archeological material the pelvic portion of the skeleton is least likely to be well-preserved (Hooton, '30). Thieme and Schull ('57) have presented discriminant functions for sexing based on Negro skeletons, but in all of their functions measurements of both the pelvis and the femur are required. Because in some archeological contexts association of bones is not clear, the restriction of our measurements to the cranium without mandible should be helpful.

The first test sample can safely be considered white: some 200 skeletons uncovered, examined, and reburied in '35 at Gallen Priory in County Offaly, Ireland. The anthropological results of this study were published by Professor W. W. Howells ('41), and with his kind permission we have been able to apply one of the discriminant functions (no. 17) to the unpublished raw cranial measurements of this series. The evidence indicates that these skeletons date from the sixth through the sixteenth centuries, during which time the monastery functioned and presumably provided the greatest number, if not all, of the burials surrounding it. We should expect, then, that female skeletons would be very scarce, and in fact Howells ('41, p. 113) classed only 12 of 139 crania (most with associated skeletons) as female, and these with some reservations. Of the crania judged male, 42 provided the measurements necessary for discriminant function no. 17. Only two of the 42 would be classed as female by this function. Of the eight with requisite measurements that Howells considered female, only three would be so classified by the discriminant function technique. As Howells ('41) has commented, the lack of comparative material made the sexing of supposed females difficult at best, but this should not have affected the sexing of the males. Thus the very close agreement with Howells in the case of the males adds circumstantial evidence for the efficiency of the discriminant function method in situations of this sort.

The second test sample consists of three series of American Indian crania: from Indian Knoll, Pecos Pueblo, and Florida. The Indian Knoll site in west-central Kentucky, dated by radiocarbon approximately 3450 B.C., has produced a very large series of well-preserved skeletons that have been measured and described in full, and the individual measurements published by Snow ('48). Recently this material has been carefully re-evaluated with regard to age and sex by Johnston and Snow ('61). Their unpublished resexing data have very kindly been placed at our disposal. The Indian Knoll material, in our opinion, provides one of the best possible Indian series from the point of view of size, preservation, and exactness of visual sexing.

The Indians of Pecos Pueblo in north-central New Mexico, dating from approximately 1300 A.D. on, have been analyzed and described by Hooton ('30). His measurements and sex identifications have been available to us through the Peabody Museum files. Hooton states that he was very painstaking in the sexing, and his results were compared with an independent determination by Professor T. W. Todd. Again we feel justified in believing the sex estimations on these specimens to be of exceptional accuracy. Since these specimens were available in the Peabody Museum we have been able to measure mastoid length and add this to the other measurements, which are Hooton's. Because of the mastoid measurement, this is the only Indian series to which the majority of the discriminant function formulas can be applied. Unfortunately many of the Pecos specimens are deformed. For the purposes of this paper, we have divided the Pecos sample into two groups: those with slight or no deformation and those with greater deformation according to Hooton's standards. We have applied the functions to both groups, and the results speak for themselves. Giles and Bleibtreu ('61) demonstrated on the basis of their sample of the Pecos material that there was no significant difference between deformed and undeformed for basion-prosthion length, but that there was for cranial length, breadth and height. Discriminant functions nos. 11, 12 and 13 omit these latter three measurements.

The third series consists of crania from Florida measured by Hrdlička ('40). As with Indian Knoll, the lack of the mastoid length measurement prevented the application of the majority of the discriminant functions. In this series we have no particular reason to claim above average accuracy for the measurments or sexing (e.g., Stewart, '42), but Hrdlička presents the individual measurements for these specimens, and the sample is instructive in showing certain limitations in the techniques.

Table 4 summarizes the results of a number of the discriminant functions when applied to these Indian series. The table shows the percentage in agreement with the original sex estimate both when the sectioning point is based on the white and Negro sample of this study, and when it is based on mean scores of the series under consideration. No crania were used where the original investigator was unable to assign the sex. It must be emphasized that these results indicate merely the extent to which the discriminant function estimation agrees with the visual appreciation estimation of Snow and Johnston, Hooton, or Hrdlička, and not the percentage "correctly" identified in the sense used with the whites and Negroes.

Several observations are warranted from the data in table 4. On the whole the discriminant functions assign the proper sex, assuming for the moment that the original determinations were unerring, with the same order of magnitude as they do for the Negro and white sample. For the Florida Indians, however, this is true only when the sectioning point is based on mean values of these same Indians. It is safe to conclude, then, that these discriminant functions provide an adequate means

TABLE 4

Percentage of Indians correctly sexed, sample size (N), and sectioning points based on the particular series

Discriminant function number	Pecos Pueblo: Undeformed, Percentage correct	Undeformed, N	Undeformed, Sample sectioning point	Pecos Pueblo: Deformed, Percentage correct	Deformed, N	Deformed, Sample sectioning point	Indian Knoll: Percentage correct	Indian Knoll: N	Indian Knoll: Sample sectioning point	Florida: Percentage correct	Florida: N	Florida: Sample sectioning point
3	81.2 *78.1*	32	6289.81	85.1 *86.6*	67	6260.59						
6	76.9 *82.1*	39	6062.14	85.2 *81.5*	81	6052.78						
9	83.3 *83.3*	42	6190.88	86.0 *87.2*	86	6185.88						
10	79.4 *82.4*	34	3741.55	80.0 *82.9*	70	3693.23	91.9 *90.8*	271	3734.60	64.4 *90.7*	118	3917.07
13	87.9 *75.8*	33	1110.21	89.7 *85.3*	68	1110.21						
16	80.8 *80.8*	52	1494.65	85.5 *84.5*	110	1483.23						
17	82.0 *88.0*	50	2589.10	82.9 *81.9*	105	2564.94	90.7 *89.2*	344	2568.28	70.0 *85.3*	217	2702.21
18	79.4 *85.3*	34	4009.93	85.7 *82.9*	70	3979.32	90.3 *87.3*	267	4006.01	61.8 *92.7*	110	4190.63
21	86.0 *83.7*	43	909.48	85.1 *81.6*	87	900.48	91.0 *90.3*	267	903.28	65.4 *89.7*	107	944.82

Percentages in roman type based on Negro and white sectioning points.
Percentages in italic type based on sectioning points derived from that particular series and given in this table.

for sexing American Indians. What apparently may not always be relied upon is the sectioning point based on the whites and Negroes. The mean values of the Florida Indian series indicate that, using Hrdlička's sex estimations, they are in nearly all dimensions a large-headed group, and exceptionally long-headed, especially the females. These factors may help to account for the large proportion of females (by Hrdlička's determination) misclassified when the white and Negro sectioning point is used. It is interesting to note at this point that there is little difference in sexing accuracy when whites are sexed by Negro-based functions and vice versa.

On the other hand, the Indian Knoll and the Pecos material was classified just about as well by means of the white and Negro sectioning point as by a sectioning point determined from their own mean values. In many but not all cases, however, the number of males and females misclassified was somewhat more proportionate to the total males and females when the sectioning point was based on the sample in question. Apparently the amount of deformation present in the Pecos series does not affect the accuracy of the discriminant function.

In dealing with crania from the Pueblo period in the Southwest, or with material akin to Indian Knoll, it would be reasonable to use either the sectioning points based on the Negroes and whites, or those based on the samples used in this study (table 4). Single Indian specimens from other areas must be approached with caution, and this includes Tepexpan man used previously to illustrate the application of a discriminant function. If a discriminant function is applied to a number of skulls from a locality that includes both male and female specimens, the characteristic bimodality of the distribution of the scores will soon become apparent. It is then possible to place the sectioning point midway between the two modes of the series under investigation, and eliminate the need for any previously determined dividing line. The larger the sample size, the more accurately the sectioning point can be estimated. The Florida series, treated in this fashion, can be sexed (again with reference to Hrdlička's observations) as accurately as those whose scores fall close to the original white and Negro values.

SUMMARY

Nine measurements taken on a series of 408 known-sex American white and Negro crania have been used in 21 combinations to form discriminant functions for sex determination. Similar measurements from a skull of undetermined sex are multiplied by the proper coefficients, the products summed, and the resultant score compared with a male-female dividing point based on the known-sex material. An accuracy of 82–89% is attained with the Negro and white material. Estimates of the reliability of the technique have been calculated, and tests were made on chimpanzee, early Irish and American Indian crania to determine the possibility of extending the orignal results to other groups. It was found that the discriminatory power held up, but the male-female dividing point in some cases may need to be adjusted to the population at hand. If one of the discriminant functions presented in this paper is used to calculate scores for a number of specimens of both sexes from a population unlike those used in this study, the bimodality with which the scores will distribute themselves provides nevertheless a means for obtaining an appropriate male-female dividing point.

ACKNOWLEDGMENTS

We owe our thanks to a number of people who have aided us in the course of this study: Professor W. W. Howells for continued advice and financial support for computation through the Earnest Albert Hooton Research Fund; Dr. Mildred Trotter, Washington University School of Medicine, St. Louis and Dr. C. C. Francis, Western Reserve University School of Medicine, Cleveland for permission to examine the Terry and the Todd skeletal collections, respectively; Drs. C. E. Snow and F. E. Johnston for unpublished data on the Indian Knoll specimens; Drs. E. E. Hunt, Jr. and A. P. Dempster, H. K. Bleibtreu, C. P. Giles, P. deY. Kay, R. M. Rowlette and F. P. Saul for various benefits; the Harvard Statistical Laboratory and the Computation Center at M.I.T., Cambridge, Mass. for providing electronic computer services; and the National Science Foundation for granting the senior author a Graduate Fellowship during the period of the research.

LITERATURE CITED

Barnard, M. M. 1935 The secular variations of skull characters in four series of Egyptian skulls. Ann. Eugenics, *6:* 352–371.

Borovanský, L. 1936 Sexual differences in human skulls. Czech Acad. Sci. and Arts, class II. (In Czech, English summary.)

Ceballos, J. L., and E. H. Rentschler 1958 Roentgen diagnosis of sex based on adult skull characteristics. Radiology, *70:* 55–61.

DeTerra, H., J. Romero and T. D. Stewart 1949 Tepexpan man. Viking Fund Publ. in Anthrop., no. 11.

Dureau, A. 1873 Des caractères sexuels du crâne humain. Rev. d'Anthrop., *2:* 475–487.

Genovés, S. 1960 Revaluation of age, stature and sex of the Tepexpan remains, Mexico. Am. J. Phys. Anthrop., *18:* 205–217.

Giles, E. 1956 Cranial allometry in the great apes. Human Biol., *28:* 43–58.

Giles, E., and H. K. Bleibtreu 1961 Cranial evidence in archeological reconstruction: a trial of multivariate techniques for the Southwest. Am. Anthrop., *63:* 48–61.

Giles, E., and O. Elliot 1962a Negro-white identification from the skull. VI[e] Congrès international des Sciences anthropologiques et ethnologiques, Paris. Vol. 1, pp. 179–184.

——— 1962b Race identification from cranial measurements. J. Forensic Sci., *7:* 147–157.

Hanihara, K. 1958 Sexual diagnosis of Japanese long bones by means of discriminant function. J. Anthrop. Soc. Nippon, *66:* 187–196. (In Japanese, English summary.)

——— 1959 Sex diagnosis of Japanese skulls and scapulae by means of discriminant function. Ibid., *67:* 191–197. (In Japanese, English summary.)

Hodges, J. L. 1955 Discriminatory analysis. I. Survey of discriminatory analysis. Project no. 21–49–004, rept. no. 1. Air University School of Aviation Medicine, U. S. Air Force, Randolph Air Force Base, Texas.

Hooton, E. A. 1930 The Indians of Pecos Pueblo: a study of their skeletal remains. Papers of the Southwestern Expedition no. 4, Yale University Press, New Haven.

——— 1943 Medico-legal aspects of physical anthropology. Clinics, *1:* 1612–1624.

——— 1946 Up from the ape, rev. ed. Macmillan, New York.

Howells, W. W. 1941 The early Christian Irish: the skeletons at Gallen Priory. Proc. Royal Irish Acad., *46:* (sect. C): 103–219.

Hrdlička, A. 1940 Catalogue of human crania in the United States National Museum collections: Indians of the Gulf States. Proc. U. S. Nat. Mus., *87:* 315–464.

Johnston, F. E., and C. E. Snow 1961 The reassessment of the age and sex of the Indian Knoll skeletal population: demographic and methodological aspects. Am. J. Phys. Anthrop., *19:* 237–244.

Keen, J. A. 1950 A study of the differences between male and female skulls. Ibid., *8:* 65–79.

Kendall, M. G. 1957 A course in multivariate analysis. Hafner, New York.

Krogman, W. M. 1949 The human skeleton in legal medicine: medical aspects. In: Levinson, S. A., ed., Symposium on medicolegal problems, series two. Lippincott, Philadelphia.

——— 1962 The human skeleton in forensic medicine. Charles C Thomas, Springfield, Ill.

Parsons, F. G., and L. Keene 1919 Sexual differences in the skull. J. Anat., *54:* 58–65.

Pearson, E. S., and H. O. Hartley, eds. 1958 Biometrika tables for statisticians, vol. 1. Cambridge University Press, Cambridge.

Pons, J. 1955 The sexual diagnosis of isolated bones of the skelton. Human Biol., *27:* 12–21.

Snow, C. E. 1948 Indian Knoll skeletons of Site Oh 2, Ohio County, Kentucky. Univ. Ky. Rept. in Anthrop., vol. 4, no. 3, part 2.

Stewart, T. D. 1942 Variation in the technique of measuring skulls. Anthrop. Briefs, no. 2: 1–6.

——— ed. 1952 Hrdlička's practical anthropometry, 4th ed. Wistar Institute of Anatomy and Biology, Philadelphia.

Thieme, F. P. 1957 Sex in Negro skeletons. J. Forensic Med., *4:* 72–81.

Thieme, F. P., and W. J. Schull 1957 Sex determination from the skeleton. Human Biol., *29:* 242–273.

Williams, E. J. 1959 Regression analysis. Wiley, New York.