

THE SIZE-SHAPE APPROACH TO BIOLOGICAL DISTANCE:  
ITS APPLICATIONS TO ANTHROPOLOGY

by

Aaron Elkins

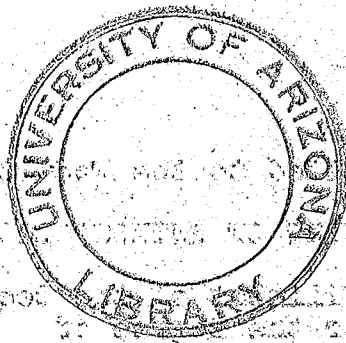
---

A Thesis Submitted to the Faculty of the  
DEPARTMENT OF ANTHROPOLOGY  
In Partial Fulfillment of the Requirements  
For the Degree of  
MASTER OF ARTS  
In the Graduate College  
UNIVERSITY OF ARIZONA

1959

UNIVERSITY OF ARIZONA LIBRARY

FOR INFORMATION OF THE UNIVERSITY



ALL INFORMATION CONTAINED  
HEREIN IS UNCLASSIFIED  
DATE 10/15/2010 BY 60322 UCBAW/STP

DATE OF REVIEW

DATE OF REVIEW

DATE OF REVIEW



## ACKNOWLEDGEMENTS

This thesis could not have been completed without the liberal aid of a number of people. I wish to express my deep appreciation for the thoughtful guidance which Dr. Frederick S. Hulse has freely extended in the trying capacity of adviser-by-mail. Drs. Emil W. Haury and Harry T. Getty, also of my thesis committee, made many suggestions which have improved this paper. Dr. William S. Laughlin was very helpful in the formative stages of this study, and in supplying the Greenland raw data. Drs. L. S. Penrose and Marshall T. Newman made several important suggestions for which I am grateful. Dr. Robert McGinnis and Mr. Robert Leik gave most generously of their time in discussing statistical questions with me. Drs. Marvin Kivett, John Champe, and Edward Fry kindly permitted and aided my measurement of Pawnee crania from the Nebraska State Historical Society collection, a series which has since been deleted from this paper. Finally, Toby Elkins has been invaluable in her capacity as secretary, consultant, and morale-booster.

## TABLE OF CONTENTS

	Page
Acknowledgements.....	iii
Part I: Size, Shape, and Distance	
1. Introduction.....	1
2. The Size-Shape Approach	
$C^2P$ , $C^2Q$ , and $C^2H$ .....	7
Correlation and $C^2R$ .....	17
3. Measurements .....	26
Part II: Applications of the Size-Shape Approach	
4. The Eskimo Occupation of Greenland ....	29
5. <u>Migration and Environment</u> .....	35
6. Discussion and Summary .....	47
Appendix (Examples of $C^2R$ calculations) .....	52
References Cited .....	54

## INTRODUCTION

The discrimination and classification of human populations has been a chief pursuit of physical anthropology since its earliest days, but it is only in the last half-century that an energetic attempt has been made to bring statistics into the picture. The union of the rigid, punctilious methodology of statistics with the unavoidably haphazard nature of anthropological data has been somewhat stormy. Methods of estimating "biological distances" between human groups have come thick and fast, and have been criticized out of existence almost as quickly as they have been formulated.

The earliest measure of "biological distance" between human groups to receive truly wide application was the coefficient of racial likeness (c.r.l.) of Karl Pearson, first proposed in a paper by Tildesly (1921), and more adequately presented some years later (Pearson, 1926.) Few innovations in modern physical anthropology have aroused as much comment, pro and (mostly) con as has the c.r.l., the deluge of criticism continuing into recent years (e.g., Fisher, 1936; Pearl and Miner, 1935; Rao, 1952; Rao, 1955; Seltzer, 1937). The c.r.l., in effect the difference between two means divided by the number of measurements, is no longer

used, its main faults being generally cited as the equation of the value of all characters used, the undesirable influence of population size, and, above all, the ignoring of the correlations between characters.

At the present time, there appears to be general agreement that the "generalized distance", or  $D^2$ , of Mahalanobis (1936) is, from the point of view of results, the best approach extant for discriminating between biological populations. Examples of its use may be found in Trevor, 1947; Mulhall and Jones, 1949; and Rao, 1955. The main drawback of generalized distance is the enormous amount of computation involved in this type of analysis (Roberts, 1954; Penrose, 1954; Spuhler, 1954).

Penrose (1954) and, in particular, Roberts (1954) have gone further in their objection to  $D^2$ , contending that the nature of anthropological data is hardly ever precise enough to warrant the use of extremely accurate statistical techniques. In his paper, Roberts also points out that:

Metrical differences between groups may be due to a) differences in genotype, arising from differential hybridization, or from drift in the case of non-adaptive "neutral" characters (if such exist) or from selection in adaptive characters; or b) to differential direct or indirect environmental modifications of the phenotypic characters during ontogeny (the groups being genetically identical); it is probable that most morphological features are the expression of interactions between both genetic and environmental components.... $D^2$  removes some...difficulties, but can

it be said to have touched on the essential problem?  
(p.148)

Many other workers have stressed the necessity for recognition and separation of genetic and non-genetic traits, among them Laughlin and Jorgensen, 1956; Sanghvi, 1953; Sanghvi and Khanolkar, 1949; and Spuhler, 1954. The problem of the inference to be made from a distance estimate based on phenotypic characters is more important than the statistical procedure involved in computing the estimate. "Biological distance" is, at present, a uselessly vague term in statistical analysis, and "genetic distance", "environmental distance", and painfully often, "morphological distance" are more appropriate for the type of information which can now be secured. The amelioration of this problem, it seems to me, is the obligation of physical anthropology, not statistics. Until a much more thorough knowledge of the mechanisms involved in the transmission of anthropometric characters is secured, the anthropologist has little choice but to follow tradition and assume that "morphological distance" can be roughly equated with "biological distance", recognizing that biology includes both genetics and environment.

Another criticism of  $D^2$  made by Roberts (1954) is that it can mask the large contribution of one trait, and the small contribution of another. Actually this is a



restatement of the criticism of the c.r.l. to the effect that all traits are given equal weight, an objection that can be made of any distance measure of which I know. Once again, I think, this is not a legitimate criticism of statistical inaccuracy, but a reflection of the lack of anthropological knowledge. Even if a statistical technique were introduced wherein the exact contribution of every character to total distance could be made, we would be only a small step closer to "biological distance", for that which is a single anthropometric character is not, of necessity, a single genetic unit, but is more than likely the end-product of quantitative inheritance. The use of discrete traits (e.g., Laughlin and Jorgensen, 1956) is a step towards overcoming this important problem.

Pitfalls in discrimination and distance studies are created not only by the nature of the characters used, and by statistical inadequacies, but by the relationships of the populations themselves. Le Gros Clark (1955) and Von Bonin and Morant (1938) have noted that statistical measures of distance become less reliable with increasing biological distance between the populations or individuals studied. What seems to me to be the prime reason for this effect may be explained with the aid of a simple, if irrelevant example. Suppose that we wish to calculate the relative morphological

distances between a cocker spaniel, a Saint Bernard, and a portable radio, and that we use size, weight, and color as our characters. It is very probable that the spaniel will appear closer to the radio than to the Saint Bernard. If, however, another type of dog is substituted for the radio, it may be that the relative distance measurements will approximate the biological situation. In more pertinent terms, we may say that no measure of morphological distance makes it possible to overcome poor judgement on the part of the scholar in the selection of the traits to be used. Nor can the degree of a particular distance, by itself, tell us whether it is primarily the result of divergent, convergent, or parallel evolution. Since it is the amount of divergence which has usually been sought in comparative studies in physical anthropology, it is this which is of most concern here. In order to be relatively certain that it is divergence which is being measured, and not parallelism or convergence, it appears necessary to limit studies of distance to groups which are positively known to have diverged in the not-too-distant past. A distinct line between "too distant" and "not-too-distant" is, of course, impossible to determine, for it will change with the traits measured or observed.

Recently, a number of scholars have applied relatively simple methods for the discrimination of human

populations. Most of these methods are more concerned with discrimination, per se, and with classification, than with the somewhat more subtle concept of distance between human populations. The work of Sanghvi (1953) and Sanghvi and Khanolkar (1949) are examples of the classification approach. Both of these papers are quite valuable, in that they separate and compare genetic and "morphological" characters. (The dichotomy is less than mutually exclusive, for the genetic components of "morphological" traits are largely unknown.) Since separation into "races" is a chief concern of both these papers, statistical analysis is by means of measures of significance.

Clark (1952) has suggested a measure of distance between populations, the coefficient of divergence. This measure represents the difference between the means of two populations, divided by the mean of the means, for each character considered. To my knowledge, this technique has been twice used in physical anthropology. Spuhler (1954) employed it as a tool in an analysis of variance study of Southwestern groups, the aim of which was grouping, rather than establishing relative distances between isolates; and Laughlin and Jorgensen (1956) made it their measure of distance in their study of the Eskimo population of Greenland.

## THE SIZE-SHAPE APPROACH

Size, Shape, and Mean Square Distance.

A most interesting approach to the study of distance has been put forth by L. S. Penrose, who has devised a simple method of approximating the results of  $D^2$  (1947, 1954). It is the purpose of this paper to present a discussion of Penrose's approach to the study of morphological distance between human populations - with particular emphasis on the elucidation of basic points not adequately (to my mind) covered by Penrose. Because of the possibility of misinterpretation or error on my part, I think it desirable to separate Penrose's thoughts clearly from my own. To this end, Penrose's ideas will not be paraphrased, but will be quoted. The responsibility is mine for statements neither enclosed in quotation marks, nor attributed to a worker other than Penrose.

Some of the basic aspects of Penrose's approach are summed up in his 1954 paper:

For purposes of all distance estimations common standard deviations for each character are agreed upon and the measurements are expressed in terms of standard deviation units. After this has been done, the

differences  $d_1, d_2, d_3, \dots$  to  $d_m$  between the mean values in any two populations can be tabulated. The c.r.l. in its original form or reduced (Morant, 1936) is equivalent from the point of view of discrimination to the coefficient used by Heincke (1898) which is found by summing the squares of the d-values from character no. 1 to character no.  $m$  and dividing by  $m$ : it may be called  $C^2H$ . Thus

$$\text{c.r.l.} = C^2H = (d_1^2 + d_2^2 + d_3^2 + \dots + d_m^2)/m.$$

Now  $C^2H$ , the mean square distance, can be broken down into two component parts which can be termed the 'size' and 'shape' distances respectively. 'Shape' distance will depend on the extent to which the d-values differ among themselves; it can be measured by their variance. The measurement of the component 'size',  $C^2Q$ , of a population in this sense will be obtained from the sum of the d-values of all the characters studied. The actual size distance between two populations will be the square of the difference in size between them, that is, the square of the mean of all the d-values. Thus

$$\text{Size distance } (C^2Q) \dots = \left[ \frac{\sum_{d=1}^m S(d)}{1} \right]^2 / m^2$$

If we subtract this  $C^2Q$  from  $C^2H$ , what remains is an estimate of the shape distance between the two populations; the magnitude of the remainder measures the amount of diversity among the d-values....

$$\text{Size } (C^2Q) = \left[ \frac{\sum_{d=1}^m S(d)}{1} \right]^2 / m^2$$

$$\text{Shape } (C^2P) = \frac{\sum_{d=1}^m S(d^2)}{1} / m - \left[ \frac{\sum_{d=1}^m S(d)}{1} \right]^2 / m^2$$

$$\text{Mean sq. distance } (C^2H) = C^2Q + C^2P = \frac{\sum_{d=1}^m S(d^2)}{1} / m$$

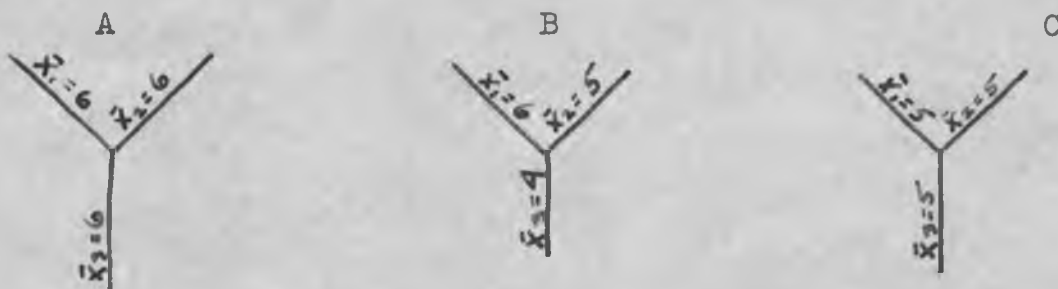
The meanings of size and shape in a single population are presented thus:

The sum of a set of equally weighted measurements, preferably expressed in terms of their standard deviations constitutes the compound estimate, which can be termed 'size'. If a set of measurements, again preferably expressed in standard units, is weighted before summation in such a manner that the sum of the weights themselves is zero, the sum constitutes a type of compound measurement which can be called profile or 'shape'. The virtue of the use of two compound measurements, size and shape, for discriminative purposes, lies in the fact that if...the characters are all equally intercorrelated, size is uncorrelated with any shape. (Penrose, 1947:228)

Since the first of the above quotations presents a simple method of computing size and shape distance between two populations, it is unnecessary to discuss the calculations entailed in measuring the size or shape of a single population. (For information on these calculations, see Penrose, 1947.) It is, however, certainly of value to have a clear understanding of the meaning of these terms. This can, perhaps, best be supplied with an example. Suppose that there are three populations of Y-shaped objects (A, B, and C), on each of which the means of three measurements (1, 2, 3) have been computed, each measurement being the length of a specific arm of the Y. These are illustrated in Figure 1.

Employing the formulae presented on page 8, the following inter-population differences in size, shape, and

FIGURE 1  
Hypothetical Populations A, B, C  
with Measurements 1, 2, 3.



mean square distance are obtained (Table 1). For a detailed example of calculations, see the appendix.

If we assume that the 3 measurements are taken in the same units, and are of the same range of magnitude, it is unnecessary to standardize the units.

TABLE 1:  
 $C^2Q$ ,  $C^2P$ ,  $C^2H$ :<sup>1</sup> HYPOTHETICAL Y-SHAPED POPULATIONS

	<u>A-B</u>	<u>A-C</u>	<u>B-C</u>
1.	0	+1	+1
2.	+1	+1	0
3.	+2	+1	-1
	<u>+3</u>	<u>+3</u>	<u>0</u>
Size distance ( $C^2Q$ )	1.00	1.00	0.00
Shape distance ( $C^2P$ )	0.67	0.00	0.67
Mean square distance ( $C^2H$ )	1.67	1.00	0.67

It will be immediately noted that  $C^2Q$ ,  $C^2P$ , and  $C^2H$  present very different results, and, therefore, offer

---

<sup>1</sup> Following Penrose,  $C^2Q$  = size distance,  $C^2P$  = shape distance,  $C^2H$  = mean square distance.

different information.

Shape. The difference in shape between A and C is zero. Since none of these distance estimates can be negative, zero is the smallest possible value, and means that no difference in shape exists, i.e., A and C have the same shape. This is because each d-value (difference between means for the same measurement) is the same (plus - 1) for each measurement. The shape of a set of characters is its variance, and the shape distance between two populations is the variance of their difference in shape, i.e., the variance of the difference in variance. Thus, it is not necessary that the means within each population be the same, as they are in this case. If the d values for A and C were

	<u>A</u>	<u>C</u>
1.	50	10
2.	80	40
3.	41	1,

the shape distance between the two would still be zero, (although the size distance would be greatly changed), because the d-values would be equal (plus - 40) for each measurement.

The rationale for using the term "shape" may be seen in further analysis of these two hypothetical populations. In both populations, 1) measurement 1 is 30 less than measurement 2, 2) measurement 2 is 39 greater than measure-



ment 3, and 3) measurement 1 is 9 greater than measurement 3. It is not the measurement values, per se, to which "shape" refers, but to their relationships, i.e., to the differences between them.

The inter-trait differences are exactly the same in both cases; therefore, their "shapes" are the same. This is manifested when the differences are listed:

	<u>A-C</u>
1.	40
2.	40
3.	40

Since shape distance is equal to the variance, it is obvious that it will be zero.

In summation, the shape of a population is not directly a function of its characters, but represents the interrelationships of its characters, or, as Penrose says, its "profile". The shape distance between two populations is the profile of the distance in shape existing between them, or, in other words, the variance of the d-values.

Size. Size is a more easily visualized concept than shape, being, as Penrose points out, a direct function of the sum of the differences between two populations (or, in other words, the difference between the sums of the characters of two populations). In the illustration above, B and C are equal in size because the d-values between them are plus - 1, minus - 1, and zero, totaling zero. As in

shape, magnitude of individual d-values is of no import.

If the values of B and C had been

	<u>B</u>	<u>C</u>
1.	25	100
2.	150	76
3.	2	1,

size difference would have remained zero, because the d-values would have been minus - 75, plus - 74, and plus - 1, which total zero.

Mean square distance. By using the formula  $S(d^2)/m$  (where d = d-values, m = number of characters) for mean square distance, Penrose does away with sign, and eliminates the possibility of the mutual cancellation of d-values seen in the example immediately above. Mean square distance may be thought of as a direct function of the magnitude of the individual d-values, differing from size in discounting signs, and thereby cumulating the differences, each as a completely independent element.

In practice, it will be remembered, size distance and mean square distance are calculated by separate formulae, and shape distance is found by subtracting the former from the latter. Why shape distance and size distance - at first glance, two unrelated measures - should add up to mean square distance - also seemingly independent - can be illustrated, I think, in the following manner. First, let us suppose that in a particular comparison, size distance equals zero.

For this to be possible the sum of the d-values must also equal zero, and, therefore, so must the mean of the d-values be zero. Consequently, each d-value will be equal to its own deviation from the mean. Because shape distance, being equal to the variance of the d-values, is the mean of the squared deviations from the mean, and mean square distance is equal to the mean of the squared d-values, the two estimates will be equal. Since, in this example, size distance is zero, mean square distance equals shape distance plus size distance.

For the second example, let us say that shape distance equals zero. Equality of all d-values is requisite for this to occur. In such a population,  $(Sd)^2/m^2$  will be equal to  $S(d^2)/m$ , these formulae being those for size distance and mean square distance, respectively. Since it has been stated that the shape distance is zero, we may say that mean square distance equals size distance plus shape distance.

Based on the discussion of the previous pages, the most important mechanical aspects of these measures appear to be the following:

1. Maximal size distance occurs when all the d-values are large, and all of the same sign. The variability of the d-values will, of course, be related to their sum, but variability, per se, is of no direct importance.

Minimal size distance occurs when the sum of the d-values is zero, and therefore, when the sums of the negative and positive d-values are equal.

2. Maximal mean square distance will result when the d-values are very large. Minimal mean square distance occurs with very small d-values. Once again, variability of the d-values is not of direct consequence.

3. In contrast to size distance and mean square distance, shape distance is a reflection of the variability of the d-values, and is not directly concerned with their magnitude. Maximal shape distance ensues with large differences between d-values; i.e., when population "profiles" are greatly different. Minimal shape distance is the result of equality, or constancy, among the d-values.

These three points are of great importance to Penrose's approach, as will be shown below.

Having discussed size, shape, and mean square distance as more or less abstract concepts, this paper now proceeds to a consideration of Penrose's method as it applies to the estimation of distance between human groups by means of the multivariate analysis of metrical characters.

The first complication to arise in the comparison of biological populations is the reduction of measurement data to some sort of uniform measure independent of the

units used. This is necessary lest a maximum cranial breadth measurement of 150 mm. have the same weight as a maximum length of 150 mm., or an upper facial height of 70 mm. be accorded one-half the weight of a maximum breadth of 140 mm.

Penrose (1954) recommends the employment of the standard deviation unit for this purpose. This unit, as described by Penrose (1958) is essentially similar to a familiar type of standardized unit called, among other things, the standard deviate, the normal deviate, the z-score, the standard normal variable, the unit normal deviate, and the standardized score (see, e.g., Mode, 1951, p. 145; Wallis and Roberts, 1956, p. 365). To express in standard deviation units the mean of a particular character in a particular population, that mean is subtracted from the mean of the means for that character of all populations being compared, and divided by the mean of all the standard deviations for that character.

A formulation of this is

$$\frac{\frac{\sum_{i=1}^n (\bar{x}_i) - \bar{x}_{A1}}{n}}{\frac{\sum_{i=1}^n (S.D._i)}{n}}$$

where A is a particular population, the subnumber 1 is a specific trait and n is the number of populations. (A detailed description, with an example, of the actual steps involved may be found in the appendix).

An S.D. unit, then, is an estimate of how much a mean of one population differs from the interpopulation mean for that trait. Throughout the rest of this paper, d-values refer to differences between populations in terms of S.D. units. This conversion need not complicate the ideas presented thus far; for the purposes of this paper, expression in S.D. units may be regarded simply as a reduction of population means to a uniform scale to facilitate multivariate analysis.

Correlation and  $C^2R$ . The problem of correlation is a major one in multi-character comparisons of human populations. Pearson's c.r.l. has been criticized on a number of counts, but the most consistent and damning is that the c.r.l. assumed lack of correlation between characters (e.g., Fisher, 1936; Pearl and Miner, 1935; Penrose, 1954; Rao, 1952, 1955; Seltzer, 1937), a deficit acknowledged by Pearson (1926). A keystone in Penrose's method is the fact that mean square distance (which, it will be remembered, is equal to the c.r.l. from the point of view of discrimination) can be thought of as the sum of two components, size and shape distance, and that it is

possible to weight these two components in a manner dependent on inter-character correlation, producing a mean square distance in the computation of which correlation is considered. Penrose calls this weighted mean square distance " $C^2R$ ". The premise upon which this weighting is based is that,

If the characters are uncorrelated with one another,  $R$  (average correlation coefficient) equals zero, and equal weightings of size and shape are appropriate, giving rise directly to  $C^2H$  or its equivalent, the c.r.l. If the characters are positively correlated, shape becomes more important than size, and if they are negatively correlated, size becomes more important than shape. (Penrose, 1954, p.340)

The reasons for this statement, although nowhere presented by Penrose, are of paramount importance in the appreciation of  $C^2R$ . It appears to me that the size-shape-correlation relationship put forth above is based on the inverse effects of positive and negative correlation on the shapes of populations. Positive correlation may be expected to promote the constancy of inter-trait differences from population to population, while negative correlation will result in the inconstancy of differences between traits, from population to population. For example, suppose that head length and breadth have a high positive correlation. If, then, the mean head length of one population is 170 mm., and the mean breadth is 140 mm., we might expect that a second population with a mean length of 175 mm. would have a mean breadth of about 145 mm., while a third population might have a 180-150 length-breadth ratio.

If, on the other hand, we suppose a high order of correlation - this time negative - for the two traits, our first population might again have a 170-140 length-breadth ratio, but a group with a length mean of 175 might have a mean breadth of 135, while a group with a mean head length of 180 mm. might have a mean breadth of 130 mm.

Now, it must be remembered that, while mean square distance is not directly concerned with intra-group relationships, but with differences between groups, this is not so for size distance and shape distance, which reflect differences between intra-group relationships. In terms of the above examples, the shape distance between any two populations depends on the difference between length and breadth in each population. If this difference is the same in both populations, it means that the population profiles are the same, and, thus, the shape distance between them is zero. The more disparity between the two differences, the greater will be the shape distance. Size distance, in the above examples, is dependent simply on the sum of length and breadth in each population, and on how different these sums are.

Keeping these conditions in mind, and remembering that given one character in a population, the other is completely predictable, we may safely state that any shape



or size difference between populations from one of the two illustrative groups above must be due to length-breadth correlations, and not to the shapes or sizes of the populations involved.  $C^2Q$  and  $C^2P$  results for these populations are presented in Table 2. Standardized units are not used, being unnecessary for this purpose.

It is clear from Table 2 that among the populations with positive correlation, the large size distances are functions of correlation, and not of differences inherent in the populations, while shape distance is not influenced by correlation. With negative correlation, the reverse is true. In a case where there is no correlation between characters, it would appear logical to reckon each difference as an independent unit, which can be accomplished by cumulating them without regard for sign. In effect, this is, of course,  $C^2H$ , the summation of unweighted size and shape distance. Thus, Penrose's suggestion of weighting size and shape in accord with correlation appears justified.

The first step in the actual weighting of size and shape distance is accomplished by estimating the "hypothetical constant correlation between characters", which, as may be imagined, presents a problem.

Penrose (1954) suggests using a mean correlation coefficient as  $\underline{R}$ , the hypothetical constant correlation, when observational data is available. If the inter-

TABLE 2  
HYPOTHETICALLY CORRELATED POPULATIONS

With Positive Correlation:

	Populations		Populations		Populations	
	<u>A</u>	<u>B</u>	<u>A</u>	<u>C</u>	<u>B</u>	<u>C</u>
Length	170	175	170	180	175	180
Breadth	<u>140</u>	<u>145</u>	<u>140</u>	<u>150</u>	<u>145</u>	<u>150</u>
	$c^2_Q = 25$		$c^2_Q = 100$		$c^2_Q = 25$	
	$c^2_P = 0$		$c^2_P = 0$		$c^2_P = 0$	

With Negative Correlation:

	Populations		Populations		Populations	
	<u>A'</u>	<u>B'</u>	<u>A'</u>	<u>C'</u>	<u>B'</u>	<u>C'</u>
Length	170	175	170	180	175	180
Breadth	<u>140</u>	<u>135</u>	<u>140</u>	<u>130</u>	<u>135</u>	<u>130</u>
	$c^2_Q = 0$		$c^2_Q = 0$		$c^2_Q = 0$	
	$c^2_P = 25$		$c^2_P = 100$		$c^2_P = 25$	

correlations are unknown, he offers a formula for estimating  $\underline{R}$ , based on the assumption that, "The two populations, or individuals, to be discriminated are drawn from the same parent population; more precisely, the correlation between the means of the characters in the two sub-populations is the same as the correlation,  $\underline{R}$ , within each sub-population." (1954, p.340) Elsewhere, in considering "two different species of the same genus, . . . two different races of mankind or two psychological types in the same population", he writes, "The constant  $\underline{r}$  between characters in the groups can . . . be supposed to be merely a repetition of the same property in the parent population." (1947, p.234)

It is my belief that Penrose underestimates the variability of human populations. Correlation coefficients calculated for the cranial measurements of various human groups appear to be characterized by diversity more than by constancy. (See, e.g., Pearson and Davin, 1924; Howells, 1957; Jones and Mulhall, 1949; Trevor, 1947.) For this reason, I have followed Penrose's first suggestion throughout this paper, in using a mean value of  $\underline{r}$  for  $\underline{R}$ .

Coefficients of correlation which I have calculated for the Greenland Eskimo isolates discussed later are presented in Table 3. The coefficients are for glabello-occipital length, maximum cranial breadth, and auricular

height. Population numbers are in parentheses.

TABLE 3:  
CORRELATION COEFFICIENTS, GREENLAND ESKIMO CRANIA

	<u>NW</u>	<u>SW</u>	<u>SE</u>	<u>NE</u>
Length-breadth	.40 (170)	.05 (52)	.02 (41)	.39 (24)
Length-height	.54 (164)	.54 (52)	.52 (41)	.21 (21)
Breadth-height	.38 (164)	.20 (52)	.29 (41)	.41 (21)

R was calculated by taking the mean for the three coefficients from each group, summing them, and taking a mean of the means. This method was adopted rather than simple pooling of all the values, to avoid the complications arising from unequal numbers. The NE population is probably too small to be reliable, but fortunately makes no difference in calculating an average r. R, with or without the inclusion of the NE values, is .33. Since computation of R's is not feasible in the populations of Japanese origin analyzed in this paper, this value is used as R in the two studies herein, which appears to me to be a more attractive alternative than calculating the various R's by the formula offered by Penrose.

Once R is determined, it is employed in the formula

$$c^2R = c^2P + \frac{(1-R)}{1-R+mR} c^2Q$$

where  $m$  is the number of measurements. This formula is extremely easy to apply, since the quantity  $\frac{(1-R)}{1-R+mR}$  is a constant, and need only be calculated once for each complete set of computations. With  $R$  equal to .33, this quantity is  $\frac{1-.33}{1-.33+3(.33)} = .4036$ . Thus, for all the distance estimates in this paper,  $C^2R$  was computed by using the formula

$$C^2R = C^2P + .4036(C^2Q).$$

The steps which led Penrose to arrive at this particular formula are beyond the scope of this paper, but are presented in his 1947 paper.<sup>2</sup> Several of its less esoteric aspects should, perhaps, be mentioned.

1. It is clear that positive correlation will, indeed, lower the weight of size, while negative correlation will increase its weight.

2. The construction of the formula allows raising or lowering the weight of size alone, so that shape distance is only indirectly weighted. Thus, no matter how high the negative correlation, shape distance can not be completely eliminated.

3. In an extreme case it is possible for the

---

<sup>2</sup> Mr. Robert Leik and Dr. Robert McGinnis of the Sociology Department of the University of Wisconsin have assured me that, from the statistical point of view, the derivation of the formula is valid.

denominator of  $\frac{(1-R)}{1-R+mR}$  to be zero, in which case the weight of size is undefinable.

The latter two points pertain to cases where R approaches +1 or -1, and not to the middle and lower ranges of correlation. Since the characters used in this study, like most anthropometric traits, are not highly correlated, these criticisms are irrelevant to our purposes, and may therefore be ignored.

## MEASUREMENTS

The traits used in this study have been selected neither for their discriminatory value nor for any knowledge we have of the mechanics of their inheritance, but because they were available in the material being studied. Three cranial measurements are used in each of the studies in this paper: maximum length (in the mid-sagittal plane), maximum breadth, and auricular height. The employment of this unusually small number of characters is a result of the fact that this paper had its beginning in the study of the Greenland crania presented below. Six measurements were recorded for these skulls: the three mentioned above, plus basion-bregma height, maximum horizontal circumference, and minimum breadth of the ascending ramus. At first, distance estimates were made using all six characters (with results much like, but not identical to those later computed from three measurements), but basion-bregma height and horizontal circumference were discarded because of the largely redundant nature of the information they provided. Since comparability to other studies was desired, minimum breadth of the ascending ramus was also deleted, as it is very often not recorded in the

literature, and is impossible to measure accurately in a living subject.



## APPLICATIONS OF THE SIZE-SHAPE APPROACH

In this paper,  $C^2R$  will be applied to two anthropological studies, both of which are unusual in their high degree of genetic control.

Examples of the computation involved in arriving at  $C^2R$  are presented in the appendix.

Both studies pertain only to measurements on adult males.

## THE ESKIMO POPULATION OF GREENLAND

Laughlin and Jorgensen (1956) have presented a study of the genetic distances between four Greenland Eskimo isolates, employing Clark's coefficient of divergence (Clark, 1952) as a measure of distance. These authors point out that archeological investigations have determined that the Cape York area was the sole port of entry for Eskimos migrating to Greenland, and that the immigrants, once there, moved in two directions; 1) going across or around Pearyland, and down the northeast coast, possibly as far as Scoresby Sound, and at least as far as Clavering Island; 2) leaving Cape York for the south, moving along the western coast, around the southern extremity, and then north along the eastern coast to Scoresby Sound. Since the Greenland ice cap prevented any non-coastal migrations, "Greenland presents what may be the only classical example in human population spread of a ring distribution where no gene exchange has taken place across the center of the ring but has been confined to gene flow between contiguous isolates on the periphery" (Laughlin and Jorgensen, 1956; p. 6). There is questionable evi-



FIGURE 2

ESKIMO MIGRATION ROUTES

dence for some contact between the two terminal isolates across Scoresby Sound, but, due to geologic barriers, it was probably sporadic, if existent. Figure 2 is a diagram of the direction of the migrations, and the locales of the isolates.

On the basis of the conditions outlined above, it appears possible to predict that the smallest genetic distances will be between contiguous isolates, and that the greatest of all the distances will be between NE and SE. Because the geographical distance between NW and SW is considerably less than that between SW and SE, and because harsher physiographic conditions are met in traveling from SW to SE than in moving from NW to SW, it appears reasonable to expect a greater genetic distance between SW and SE than between NW and SW. If it can be assumed, as seems reasonable, that differences between the northgoing and southgoing migrants increased with time, NE may be expected to be genetically closest to NW, next closest to SW, and, as noted above, furthest from SE.

In Laughlin and Jorgensen's 1956 paper, these expectations are clearly met by the coefficients of divergence for four cranial measurements and three indices; coefficients derived from seven discrete traits are somewhat ambiguous. In 1957, however, the original small series was enlarged to a size of greater statistical significance,

and a variant of Penrose's method, for use with discrete traits, was applied. The results, in this case employing eight traits, were completely in accord with the expected distances.

Dr. Laughlin has kindly supplied me with raw measurement data on the male crania of the four isolates, and Table 4 presents the resultant  $C^2R$  distances, in order of increasing distance.

TABLE 4

## DISTANCES OF GREENLAND ESKIMO CRANIA

NW-SW	.01
NW-SE	.05
SW-SE	.07
SW-NE	.18
NW-NE	.21
SE-NE	.30

## NUMBERS OF OBSERVATIONS

	<u>NW</u>	<u>SW</u>	<u>SE</u>	<u>NE</u>	<u>Total</u>
Length	170	52	41	24	287
Breadth	170	53	41	24	288
Auricular Height	165	52	41	21	279

Before interpreting Table 4, it is important to note that my own experiments with  $C^2R$  figures have convinced me that differences of up to .03 or .04 may result from rounding off, and, therefore, that distances of this degree should not be considered significant.

The  $C^2R$  results conform well to the predictions on the previous pages, and to Laughlin and Jorgensen's findings, and are nowhere contradictory to them. Interestingly, the clear difference in magnitude between the NW-SW-SE distances on the one hand, and the distances between these and NE on the other, does not occur in the figures of Laughlin and Jorgensen. This lack of agreement is not surprising, for it is not, of course, unexpected that different sets of traits change at different rates, and in different directions.

A detailed analysis of the  $C^2R$  distances, and of the differences between these results and those of Laughlin and Jorgensen is prevented by a lack of knowledge of the history of the Eskimo groups prior to their arrival in Greenland, a less than precise understanding of the nature of the traits used in the studies, and by the complicated picture of environmental relationships between the Greenland isolates.

Whatever information a more meticulous analysis would offer, it can be stated here that the application of the size-shape technique to the length, breadth, and height of the Greenland crania results in distance estimations which are in agreement with estimates derived from different traits and through different methods, and

that the common findings support independently made archeological deductions.

With the exception of the NW-SE contrast,  $C^2R$  distance is in agreement with geographical distance. That is, the further apart two populations are in miles, the further apart they are in head measurements. This fact would support the hypothesis that genetic drift has caused the differences in head shape. Without more detailed knowledge of variations in environmental conditions in Greenland, we cannot say whether it agrees with any hypothesis concerning environmental effects upon head shape.

MIGRATION AND ENVIRONMENT:  
FURTHER ANALYSIS OF CERTAIN DATA PRESENTED  
IN SHAPIRO'S 1939 STUDY

Shapiro's classic Migration and Environment is a rigidly controlled study of the physical characteristics of three large groups: non-migrating Japanese (sedentes), Japanese who migrated to Hawaii, and Hawaiian-born offspring of these migrants. To facilitate careful examination, each of these groups was divided into four subgroups, depending on the Japanese prefectural origin. These four prefectural divisions are Hiroshima, Yamaguchi, Kyushu (including Fukuoka and Kumamoto prefectures), and "Other Japan" (including Niigata and Fukushima prefectures). The number of Hawaiian-born persons of "Other Japan" descent was too small for analysis, and their measurements are not used. Altogether, then, there is a total of eleven groups.

Migration and Environment is not along the lines of a family study, but adequate genetic control is ensured by keeping strict account of the prefectural origins of members of all groups, and by gathering, for the sedent category, individuals who are blood relatives of those in



the migrant class. A large majority of the mothers of the Hawaiian-born were picture brides, and these mothers came, almost without exception, from the same prefectures as their husbands.

The unequivocality of genetic and environmental relationships in this study allows a fineness of interpretation which is impossible in an analysis of skeletal remains.

By comparing each of the eleven groups to every other group, fifty-five distance estimates result. Clarity and conciseness will be served if this study centers about six comprehensive comparisons, viz.; inter-sedent, inter-migrant, inter-Hawaiian-born, sedent-migrant, sedent-Hawaiian-born, and migrant-Hawaiian-born. Some discussion of these groups is called for before the presentation of the relevant  $C^2R$  results.

In a broad sense, these six comparisons may be divided into two subsets, one of which (inter-sedent, inter-Hawaiian-born, and inter-migrant) consists of comparisons of groups which are related in the sharing of environments, but are genetically rather distant; the other (sedent-migrant, sedent-Hawaiian-born, and migrant-Hawaiian-born) consisting of groups whose chief connections are genetic. It follows from this that differences between

groups in the former subset must be largely the results of genetic differences, while differences within the latter subset are likely to be environmentally induced.

To go further than this in analyzing these interrelationships is dangerous, at the present time, due to the unknown nature of the migrant-sedent differences. If we accept Shapiro's theory that the migrants are different from the sedentes on account of selection, we still do not know whether this selection is genetic, or based on non-heritable factors. Moreover, we do not know whether the migrants underwent any further bodily changes after leaving Japan. It does, however, appear relatively safe to assume the following: 1) of the six comparisons, the one with the strongest environmental tie is the inter-Hawaiian-born (with the possible exception of the migrant-sedent comparison).<sup>3</sup> 2) the strongest genetic relationship exists between migrants and Hawaiian-born (once again, with the possible exception of the one between migrants and sedentes).<sup>3</sup> 3) In terms of genetic relationship, the

---

<sup>3</sup> If the migrant-sedent difference is genetic, then their (the migrant-sedent) closest relationship is really environmental; if the difference is only phenotypic, they will have a high degree of genetic relationship. Statistically, the coefficient of relationship between the migrants and sedentes, and between the migrants and Hawaiian-born is equal (approximately .5), since migrants and sedentes are often brothers, while migrants and Hawaiian-born are often fathers and sons.

inter-Hawaiian-born distance is probably greater than the inter-sedent and inter-migrant distances. The Hawaiian-born individuals are of the generation following the migrants, and it is to be expected that genetic distance increases between groups with the passage of generations.

4) in terms of environmental relationship, the order of increasing distance of the following three groups is probably migrant-sedent, migrant-Hawaiian-born, and sedent-Hawaiian-born. This ordering is based simply on the amount of time each group shared a geographical environment with the others.

All possible  $C^2R$  distances are presented in Table 5. This formidable material is simplified in Table 6, which presents six key means derived from Table 5. Observation of Table 5 will show that these means are representative of their populations, since variability between the distances from which each mean is computed is not extensive.

As mentioned in the chapter on Greenland, differences less than .03 or .04 may easily be the results of arithmetical procedures, and we dare not assume that they signify real differences between populations.

Table 6 suggests several interesting conclusions. First, it appears that the environment is the primary agency in the determination of the precise details of head

TABLE 5:  
C<sup>2</sup>R RESULTS

1.	Yamaguchi Imm. - Yamaguchi H.B.	.01	29.	Yamaguchi Sed. - Kyushu Imm.	.20
2.	Yamaguchi H.B. - Kyushu H.B.	.01	30.	Kyushu Sed. - Kyushu H.B.	.20
3.	Yamaguchi Imm. - Kyushu H.B.	.01	31.	O.J. Imm. - Yamaguchi H.B.	.21
4.	Hiroshima Sed. - Kyushu Sed.	.03	32.	Kyushu Sed. - Yamaguchi Imm.	.22
5.	Kyushu Imm. - O.J. Imm.	.04	33.	Hiroshima Imm. - Hiroshima H.B.	.25
6.	Kyushu Sed. - Hiroshima Imm.	.04	34.	Kyushu Sed. - Yamaguchi H.B.	.27
7.	Hiroshima Sed. - Yamaguchi Sed.	.05	35.	Yamaguchi Sed. - O.J. Imm.	.28
8.	Hiroshima Imm. - O.J. Imm.	.06	36.	Kyushu Imm. - Kyushu H.B.	.29
9.	Hiroshima Imm. - Kyushu H.B.	.07	37.	Yamaguchi Sed. - Kyushu H.B.	.31
10.	Hiroshima H.B. - Yamaguchi H.B.	.08	38.	Hiroshima Sed. - Kyushu H.B.	.32
11.	Hiroshima Imm. - Yamaguchi Imm.	.08	39.	Yamaguchi Imm. - Kyushu Imm.	.32
12.	Yamaguchi Imm. - Hiroshima H.B.	.09	40.	Kyushu Sed. - O.J. Sed.	.33
13.	Hiroshima H.B. - Kyushu H.B.	.09	41.	Kyushu Imm. - Yamaguchi H.B.	.36
14.	Kyushu Sed. - O.J. Imm.	.09	42.	O.J. Sed. - Kyushu H.B.	.36
15.	Kyushu Sed. - Kyushu Imm.	.10	43.	Hiroshima Sed. - Yamaguchi Imm.	.37
16.	Hiroshima Sed. - Kyushu Imm.	.10	44.	Yamaguchi Sed. - Yamaguchi Imm.	.40
17.	Hiroshima Sed. - Hiroshima Imm.	.10	45.	Kyushu Sed. - Hiroshima H.B.	.42
18.	Hiroshima Imm. - Kyushu Imm.	.10	46.	Hiroshima Sed. - Yamaguchi H.B.	.42
19.	Hiroshima Imm. - Yamaguchi H.B.	.11	47.	Yamaguchi Sed. - Yamaguchi H.B.	.42
20.	Yamaguchi Sed. - Kyushu Sed.	.12	48.	O.J. Sed. - Yamaguchi Imm.	.44
21.	Hiroshima Sed. - O.J. Sed.	.12	49.	O.J. Sed. - Yamaguchi H.B.	.45
22.	Yamaguchi Sed. - O.J. Sed.	.12	50.	O.J. Imm. - Hiroshima H.B.	.48
23.	O.J. Sed. - O.J. Imm.	.15	51.	Yamaguchi Sed. - Hiroshima H.B.	.52
24.	Yamaguchi Sed. - Hiroshima Imm.	.15	52.	O.J. Sed. - Kyushu Imm.	.57
25.	Hiroshima Sed. - O.J. Imm.	.16	53.	Hiroshima Sed. - Hiroshima H.B.	.58
26.	O.J. Imm. - Kyushu H.B.	.17	54.	Kyushu Imm. - Hiroshima H.B.	.66
27.	O.J. Sed. - Hiroshima Imm.	.17	55.	O.J. Sed. - Hiroshima H.B.	.74
28.	Yamaguchi Imm. - O.J. Imm.	.18			

O.J. : Other Japan      Sed. : Sedentes  
Imm. : Immigrants      H.B. : Hawaiian Born

form among these groups. This conclusion is based on the fact that the three highest  $C^2R$ 's are between environmentally distant, but genetically close groups. Furthermore, the smallest  $C^2R$  exists between groups which are certainly the closest, environmentally, and the furthest apart, genetically; and the greatest  $C^2R$  is between groups which are possibly the furthest apart, environmentally.

TABLE 6:

MEAN  $C^2R$  DISTANCES

1.	Mean of all inter-Hawaiian-born distances.....	.06
2.	Mean of all inter-sedent distances.....	.13
3.	Mean of all inter-migrant distances.....	.13
4.	Mean of all distances between migrants and Hawaiian-born of the same prefectural origin.....	.18
5.	Mean of all distances between sedentes and migrants of the same prefectural origin.....	.19
6.	Mean of all distances between sedentes and Hawaiian-born of the same prefectural origin.....	.40

Shapiro's observation of a notable difference between migrants and sedentes is upheld by their relatively high  $C^2R$  distance of .19. Keeping this distance in mind, the equality of the inter-sedent and inter-migrant distances suggests that the difference between migrants and sedentes is of the same degree and direction in each prefecture. It does not appear unlikely that this uniformity of

difference is indicative of causal uniformity, i.e., that the same influences are responsible for the differences between migrants and sedentes in each prefecture. This difference in head form may be genetic in nature, it may be the result of the Hawaiian environment acting upon the migrants after their arrival in Hawaii, or it may be non-genetic, but due to factors other than the Hawaiian environment, e.g., selection for a "laborer type" or "migrant type". Drawing partly from the thoughts of Shapiro, we can make a number of predictions.

1. If the migrant-sedent difference is genetic, the migrants will be more like the Hawaiian-born than like the sedentes.

2. If the migrant-sedent difference is non-genetic, and due to the Hawaiian environment, the sedentes may be expected to be more like the migrants than like the Hawaiian-born, because the migrants shared the sedentes' environment during their early years, while the same environment was at no time common to sedentes and Hawaiian-born.

3. If the migrant-sedent difference is non-genetic, and due to factors other than the Hawaiian environment, then a) if the Hawaiian-born are not greatly affected by the Hawaiian environment, they will probably revert back to the sedentes in physical appearance, and be

more like them than like the migrants; b) if the Hawaiian-born are greatly affected by the Hawaiian environment, prediction is impossible, for the forces acting upon them will be different from the ones acting on either the sedentes or the migrants.

In Table 6, it will be noted that the migrants are not more like the Hawaiian-born than like the sedentes (thereby, the genetic hypothesis fails to be substantiated); the Hawaiian-born are not more like the sedentes than they are like the migrants (thereby failing to support the non-genetic, non-Hawaiian explanation); the sedentes are, however, much more like the migrants than they are like the Hawaiian-born, thus supporting the Hawaiian environment hypothesis. This interpretation is further borne out by observations of the means of the original measurements, as presented in Migration and Environment. In each of the three measurements, there is a cline from sedentes to migrants to Hawaiian-born. Such a cline is in complete accord with a situation in which the sedentes are totally unaffected by the Hawaiian environment, the migrants only slightly affected by it (having spent their most formative years in Japan), and the Hawaiian-born greatly affected, having been subject to no other environment.

In 1952, Lasker achieved similar results to these

in a study of Mexican migrants and sedentes. He found that differences between migrants and sedentes in a large number of morphological characters appeared to be attributable to the new environment to which the migrants moved, rather than to selection for migrants. As in the Japanese study, most of these migrants were past the generally accepted lower limits of maturity, but they were considerably younger, by and large, than the Japanese migrant group.

It should be added that, in the case of the Japanese, migrants tend to be younger brothers more often than do sedentes (Hulse, 1959), although the migrants are, on the average, older than the sedentes (Shapiro, 1939).

Much of the material thus far presented in support of the hypothesis that the Hawaiian environment is the major factor in head form differentiation of the migrants, sedentes, and Hawaiian-born could also be cited as evidence for an explanation of  $C^2R$  distances in terms of simple genetic relationships. The equal migrant-sedent and migrant-Hawaiian-born distances, for example, might be interpreted as following naturally from the equal genetic relationships of the two pairs (coefficient of relationship equals .5 in both cases). The sedent-migrant-Hawaiian-born cline, too, could reflect genetic differences. An



even stronger base from which to build a genetic hypothesis of Japanese head form differentiation may be found by further examination of Table 5. In this table, it may be seen that the sum of the  $C^2R$  distances of "Other Japan" sedentes from all the other sedentes (.57) is considerably greater than the sum of the distances between the other three sedent groups (.20). This is in agreement with the distance in miles, and with the probable genetic situation, thus paralleling the Greenland results.

The rejection, in this paper, of the genetic hypothesis in favor of the more complicated hypothesis of Hawaiian influence is based largely upon Table 6, in which, it will be remembered, distances between known environmentally and genetically related groups were presented. This table illustrated, rather convincingly, that the head length, breadth, and height of these groups are more influenced by environmental than by genetic relationships.

This does not imply that head form is solely the result of environmental circumstances. To make the assumption that these measurements are determined by one agency - either the environment, or heredity - would be most naive, in view of present anthropological and genetic knowledge. It is likely, if not certain, that, "Genes determine potentialities; the realization of these

potentialities depends on the environment in which the genes perform their functions" (Srb and Owen, 1958, p. 71). With this orientation, we may, perhaps, interpret the results noted in the previous paragraph, which set off "Other Japan" sedentes from the other sedentes, in concordance with genetic relationships. Each sedent group may be thought of as a separate group of unique genic structure, and therefore of unique potential in head form and size. Since the Hawaiian environment has had no opportunity to affect the sedentes, it is this genetic difference in head form which is reflected in the  $C^2R$  figures. It is possible to infer from this that head length, breadth, and height, notwithstanding their environmental susceptibility, may be efficacious in determining well-defined genetic relationships.

It may be asked: if we are using characters which can be molded powerfully by the environment, is it not likely that any distance estimates we derive from them will tend to underemphasize genetic relationships? The answer to this must be affirmative, and it is admitted that the traits used here are ineffective for properly evaluating genetic relationships. This circumstance does not, however, detract from the conclusions which can be made from observation of the  $C^2R$  figures, viz.: 1) the Japanese

sedentes, migrants, and Hawaiian-born males may be distinguished on the basis of differences in head length, breadth, and height. 2) these differences appear to be largely attributable to differences in the environments of these three groups; more specifically, to the amount of time each group has spent in Hawaii.

## DISCUSSION AND SUMMARY.

Penrose's approach to the problem of distance estimation has been discussed, and examples of its applications presented.

Two important advantages of this method are immediately apparent: 1) the ease of calculation. Fifteen or twenty distance estimates, utilizing six characters, can be calculated and checked in two evenings, with the aid of an ordinary desk computer; 2) the consideration of correlation. To my knowledge, no method of comparable simplicity takes into account this most important factor.

Aside from the defects to be found in any technique of biological distance estimation, one flaw unique to  $C^2R$  should be mentioned. Much of the effectiveness of  $C^2R$  is based on the calculation of  $\bar{R}$ , which does not seem completely satisfactory, since it is no more than the mean of the individual correlation coefficients. The validity of  $\bar{R}$ , then, is dependent upon the variability of the correlation coefficients, for, whereas a mean may be representative of a population with little variability, no measure of central tendency can adequately represent a highly

variable population.

The discussion thus far presented in this chapter pertains to elements inherent in Penrose's size-shape approach. It remains to consider the contributions which such a technique can make to anthropology. These contributions may be classified, I believe, into the three categories which are presented below, with examples of particular applications in this paper.

A. The Determination of the Nature of Physical Traits.

Of great importance in any study of distance between populations are the characters used. Ideally, different purposes should employ different traits, depending on the sensitivity of the traits to environment and heredity, and to various aspects of these all-inclusive domains, e.g., expressivity, penetrance, response to specific environmental stimuli, etc. With few exceptions (blood groups probably being the most notable), characters of anthropological interest do not yet lend themselves to analyses of this nature, and anthropological studies often make use of standard anthropometric characters which were devised with a view towards the acquiring of purely morphological information, and many of which originated before the widespread understanding of Mendelian genetics.

In this paper, the application of  $C^2R$  to head length, head height, and head breadth has added support to

the many other studies which suggest a strong environmental component in head form. (An extensive, annotated listing of such works may be found in Ewing, 1950).

Notwithstanding the environmental susceptibility of these three characters,  $C^2R$  distances based upon them are in agreement with probable genetic relationships in the cases of the Greenland isolates, and the Japanese sedent groups. It may be that where considerable genetic distance intervenes between groups, even such traits as these may be used with some confidence in determining genetic relationships. On the other hand, the environmental factor is difficult to take into account when considering the Greenland populations and the Japanese sedent groups, and it may be that differences between these populations are as much environmental, as they are genetic, or even more.

B. The Establishing and Testing of Broad Hypotheses and Theories.

General hypotheses of use to physical anthropology may be established and tested only by means of observations made on a large number of situations. An example of a broad hypothesis of anthropological significance (and one which is implicit in many studies, but not often tested or even stated) asserts that differences in mutations, and additions and deletions of genes are bound to accumu-

late between populations with time, and thus that separated populations will become more and more different with the passage of time. The  $C^2R$  distances between the Japanese sedentes and between the Greenland isolates appear to substantiate this hypothesis, and also support another which, in these studies, is closely related to the first, viz., unless complicating circumstances are present, geographical distance may often be used as a rough approximation of genetic distance between two groups which share a common ancestry.

C. The Determination of Local Interrelationships.

The analyses of particular anthropological situations are of great import, both for the unique knowledge which they afford, and because they are the blocks from which broader hypotheses and theories must be built. In the present study,  $C^2R$  has proven valuable in indicating the probable importance of the Hawaiian-environment in establishing the differences in head form between Japanese sedentes, migrants, and Hawaiian-born.

In summation,  $C^2R$  (and other multifactor distance estimation methods) appear to be of great potential use to anthropology: first, in establishing the environmental and genetic components of various sets of traits, and then using the knowledge so gained as an aid in tracing macro-

and micro-evolution, and in getting closer to the mechanisms and processes of population change.



APPENDIX

COMPLETE CALCULATIONS FOR  $C^2_R$ , GREENLAND

	<u>NW-SW</u>	<u>NW-SE</u>	<u>NW-NE</u>	<u>SW-SE</u>	<u>SW-NE</u>	<u>SE-NE</u>
Length	+0.098	+0.213	-0.136	+0.115	-0.234	-0.349
Breadth	-0.097	+0.469	-0.422	+0.566	-0.325	-0.891
Height	+0.044	+0.300	+0.666	+0.256	+0.622	+0.366
$s(d)$	+0.045	+0.982	+0.108	+0.937	+0.063	-0.874
$(sd)^2$	0.0020	0.9643	0.0117	0.8780	0.0040	0.7639
$s(d^2)$	0.02095	0.3553	0.6401	0.3991	0.5473	1.0496
$C^2_Q = \frac{(sd)^2}{m^2}$	0.0002	0.1071	0.0013	0.0975	0.0004	0.0849
$C^2_H = \frac{s(d^2)}{m}$	0.0070	0.1184	0.2134	0.1330	0.1824	0.3499
$C^2_P = C^2_H - C^2_Q$	0.0068	0.0113	0.2121	0.0355	0.1820	0.2650
$C^2_R$	0.0069	0.0545	0.2126	0.0748	0.1822	0.2993

## COMPUTATION OF S.D. UNITS

- METHOD: 1) Add means of each isolate for a particular measurement, and take the mean of the means.
- 2) Add S.D.'s of each isolate for this measurement, and take the mean of the S.D.'s.
- 3) Subtract each mean from the mean of the means, the difference being x.
- 4) Divide each x by the mean S.D., each result being an S.D. unit.

EXAMPLE:

		<u>Mean</u>	<u>S.D.</u>	<u>X</u>	<u>X/Mean S.D.</u> <u>(S.D. units)</u>
Glabello- opisthocranion length.	NW	188.49	6.00	+0.24	+0.043
	SW	187.94	5.18	-0.31	-0.055
	SE	187.30	5.87	-0.95	-0.170
	NE	<u>189.25</u>	<u>5.28</u>	+1.00	+0.179
	Mean	188.25	5.58		

d-values

NW-SW	+0.098
NW-SE	+0.213
NW-NE	+0.136
SW-SE	+0.115
SW-NE	-0.234
SE-NE	-0.349

## REFERENCES CITED

Clark, P. J.

- 1952 An extension of the coefficient of divergence for use with multiple characters. (Reprinted from *Copeia*, 1952, 2:61). *Yearbook of Physical Anthropology*, 209.

Ewing, J. F.

- 1950 Hyperbrachycephaly as influenced by cultural conditioning. *Papers of the Peabody Museum of American Archeology and Ethnology, Harvard University*, vol. 23, no. 2.

Fisher, R. A.

- 1936 The coefficient of racial likeness and the future of craniometry. *J. of the Royal Anthr. Inst. of Gt. Britain and Ireland*, 66:57.

Heincke, F.

- 1898 *Naturgeschichte des Herings*. Berlin.

Howells, W. W.

- 1957 The cranial vault: factors of size and shape. *Amer. J. of Phys. Anthr.*, n.s., 15:19.

Hulse, F. S.

- 1959 Personal communication

Lasker, G. W.

- 1952 Environmental growth factors and selective migration. *Human Biology*, 24:262.

Laughlin, W. S. and J. B. Jorgensen

- 1956 Isolate variation in Greenlandic Eskimo crania. *Acta Genetica et Statistica Medica*, 6:3.

Le Gros Clark, W. E.

- 1955 *The Fossil Evidence for Human Evolution*. University of Chicago Press.

Mahalanobis, P. C.

- 1936 On the generalized distance in statistics. *Proceedings National Inst. Sci. India*, 12:49.

Mode, E. B.

- 1951 *Elements of Statistics*. Prentice Hall, Inc, New York.

Morant, G. M.

- 1936 A contribution to the physical anthropology of the Swat and Hunza valleys based on records collected by Sir Aurel Stein. *J. of the Royal Anthr. Inst. of Gt. Britain and Ireland*, 66:37.

Mulhall, H. and G. I. Jones,

- 1949 An examination of the physical type of certain peoples of south-eastern Nigeria. *J. of the Royal Anthr. Inst. of Gt. Britain and Ireland*, 79:11.

Pearl, R. and J. R. Miner

- 1935 On the comparison of groups in respect of a number of measured characters. *Human Biology*, 7:95.

Pearson, K.

- 1926 On the coefficient of racial likeness. *Biometrika*, 18:105.

Pearson, K. and A. G. Davin

- 1924 On the biometric constants of the human skull. *Biometrika*, 16:328.

Penrose, L. S.

- 1947 Some notes on discrimination. *Ann. of Eugenics*, 13:228.
- 1954 Distance, size and shape. *Ann. of Eugenics*, 18:337.
- 1958 Personal communication.

Rao, C. R.

- 1952 *Advanced Statistical Methods in Biometric Research*. John Wiley and Sons, Inc., New York.
- 1955 (In the *Ancient Inhabitants of Jebel Moya (Sudan)*, by Mukherjee, R., J. C. Trevor and C. R. Rao). Cambridge Univ. Press.

Roberts, D. F.

- 1954 The Cretans. A geographical analysis of some

aspects of their physical anthropology. J. of the Royal Anthr. Inst. of Gt. Britain and Ireland. 84:145.

Sanghvi, L. D.

1953 Comparison of genetical and morphological methods for a study of biological differences. Amer. J. of Phys. Anthr., n.s., 11:385.

Sanghvi, L. D. and V. R. Khanolkar

1949 Data relating to seven genetical characters in six endogamous groups in Bombay. Ann. of Eugenics, 15:52.

Seltzer, C. C.

1937 A critique of the coefficient of racial likeness. Amer. J. of Phys. Anthr., 23:101.

Shapiro, H. S.

1939 Migration and Environment. Oxford Univ. Press, New York.

Spuhler, J. N.

1954 Some problems in the physical anthropology of the American Southwest. American Anthropologist, 56:604.

Srb, A. M. and R. D. Owen

1958 General Genetics. W. H. Freeman and Company, San Francisco.

Tildesly, M. L.

1921 A first study of the Burmese skull. *Biometrika*,  
13:176.

Trevor, J. C.

1947 The physical characters of the Sandawe. *J. of  
the Royal Anthr. Inst. of Gt. Britain and  
Ireland*, 77:61.

Von Bonin, G. and G. M. Morant

1938 Indian races in the United States. A survey  
of previously published cranial measurements.  
*Biometrika*, 30:94.

Wallis, W. A. and H. V. Roberts

1956 *Statistics*. The Free Press, Glencoe, Illinois