PEBBLE SHAPE (AND SIZE!)[1]

WERNER K. ILLENBERGER
Geology Department, Institute for Coastal Research
University of Port Elizabeth, P.O. Box 1600
6000 Port Elizabeth, South Africa

ABSTRACT: The mathematical technique of R-mode factor analysis was applied to the long, intermediate and short axes of one synthetic random data set and three sets of natural pebbles to extract the three fundamental factors related to these parameters. One factor corresponds to size, which is best described by the mean of the axes. Two further factors correspond to shape: sphericity and disc-rodness. These factors are revealed whether synthetic data or natural pebble samples are analyzed.

Extending the factor analysis to include shape indices revealed which indices are equivalent, which are most useful, and which should be discarded. Sphericity is best described with the Corey shape index, $S/(IL)^{1/2}$, and disc-rodness is best described with the disc-rod index, (L − I)/(L − S). Accordingly, the most effective shape diagram is a triangular plot of these two shape indices; the three end-members of shape are spheres, discs and rods. Shape is best investigated using this shape diagram; a 2-D technique such as contouring point density on this diagram should be used to determine the mean and standard deviation of shape.

"While others ask themselves 'What is there to do?', I ask myself 'What is there that I can avoid doing?' "

Masanobu Fukuoka, *One Straw Revolution*

INTRODUCTION

Sedimentologists have developed a plethora of shape indices to describe pebble shape (see Appendix). In this study, I apply factor analysis to one synthetic random data set and three sets of natural pebbles to determine which of these indices are the most useful. As a spin-off from the analysis, some insight is gained into pebble size.

R-mode factor analysis (FA) can be used to determine the fundamental, orthogonal factors in a data set (Cooley and Lohnes 1971; Davis 1973, 1986; Jöreskog et al. 1976, and others). Applying FA to a 2-D random data set (equivalent to rectangles), Cooley and Lohnes (1971) found that a rectangle can be described either in terms of its length and width or in terms of one size factor and one shape factor. The shape factor is the ratio of length to width; this is the only possible shape index for a 2-D object.

Davis (1973, 1986) extended Cooley and Lohnes's (1971) study to a set of 25 randomly generated 3-D objects (boxes or ellipsoids) defined in terms of three independent (but ordered) dimensions: the long (L), intermediate (I) and short (S) axes. He also used some shape and size parameters derived from the axes. He found that the fundamental factors required to describe a 3-D object are one size factor and two shape factors. Davis (1986, p. 537) deduced that one shape factor, the ratio S to a combination of I and L, is a measure of thickness (or sphericity, Sneed and Folk 1958). He deduced further that the other shape factor, represented by the ratio I to L, separates discs from rods. Figure 1, based on an excellent technique developed by Davis (1986), illustrates these factors.

It is instructive to consider the nature of the random 3-D data set. All the dimensions are random, but are then ordered into large, intermediate and small. This constraint induces some spurious correlations between the axes, as illustrated by Davis (1986, fig. 2.17). Spurious correlations may also arise between shape indices whose numerators and denominators contain common parts (Pearson 1897), as well as between the axes and shape indices. Natural pebble populations will also exhibit these spurious correlations.

It is also instructive to consider all shape indices that are geometrically possible. "First-order" indices are S/I, I/L and S/L. "Second-order" indices are products of two first-order indices; the possibilities are S^2/IL, I^2/SL and L^2/SI. Most of the shape indices derived to date (see Appendix) can be related to these first- and second-order indices. Some indices (e.g., flatness index) consist of a combination of arithmetic and geometric ratios; as will be shown later, these are similar to the indices that are purely geometric. Further indices can be derived by multiplication of first-order and second-order indices; however, nothing useful is gained from this exercise. There is only one other index that is not related to the above indices: (L − I)/(L − S), which Sneed and Folk (1958) derived to distinguish discs from rods. Folk developed this enigmatic ratio from consideration of the constraint that the I axis must have a value between the L and S axes (R.L. Folk, personal communication 1990). Folk did not name this index; I have called it the disc-rod index (DRI).

Analysis Methods and Data

Statgraphics Version 2.1 (1987), by Statistical Graphics Corporation, was used for the (R-mode) factor analysis in this study. The analysis was applied to the correlation matrices of the variables, rather than the covariance matrices, in order to give each index equal value in the analysis. Factors were rotated using the Varimax method.

I used Davis's random data set (n = 25) in this study. I also used data from 3 samples of beach and river pebbles (river pebbles n = 156; small beach pebbles n = 229; storm berm pebbles n = 221) from Algoa Bay and Swartkops River, eastern Cape Province, South Africa (Illenberger, unpublished data).

Variables should have normal distributions for FA (Da-

[1] Manuscript received 1 November 1989; revised 1 October 1990.

doi:10.1306/d42677c6-2b26-11d7-8648000102c1865d

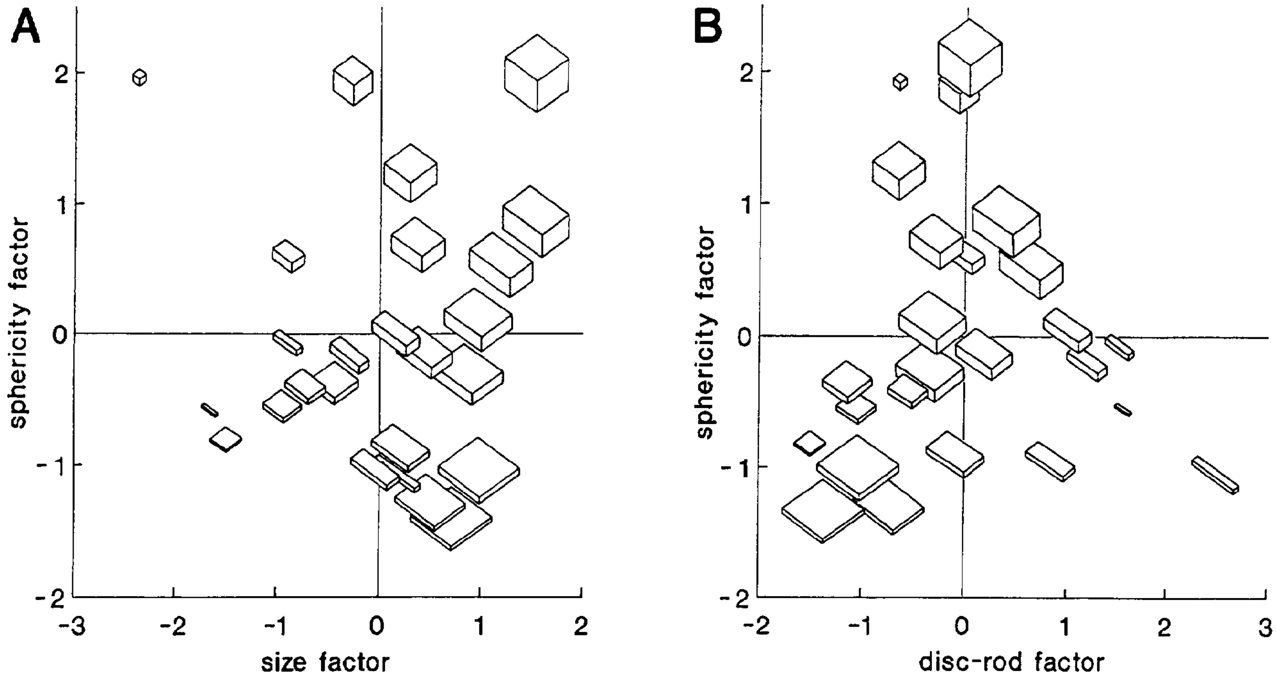


FIG. 1.—Random objects plotted on their respective factor scores, for the size factor, the sphericity factor and the disc-rod factor. Rods and discs are not separated by plotting the sphericity factor against the size factor (**A**). However, plotting the sphericity factor against the disc-rod factor (**B**) clearly separates all shapes and suggests that a triangular diagram is the most appropriate for shape. The spurious correlation between sphericity and the short axis can be seen on these diagrams; for example, objects with small short axes usually plot towards the lower left corner, and objects with large long axes plot in the upper right corner of the sphericity versus size diagram (A).

vis 1973, p. 502, and others); therefore, variables with lognormal distributions were transformed using $\log_{10}$. In particular, the size parameters of naturally occurring pebble populations usually have lognormal distributions (Blatt et al. 1972). The variables of the random objects have random distributions, so the requirement of normality had to be disregarded for this data set. To facilitate comparison, the variables of the random data set were transformed according to the distributions of the indices for the natural pebble populations.

ANALYSIS OF AXES ONLY

I first apply FA to the L, I and S axes only (Tables 1, 3, 5 and 7—see end of article). The eigenvalues of the three factors (the maximum number of factors that can be extracted from a data set with 3 variables) vary from 2.84 to 0.05; the factors must all be significant because 3 independent (orthogonal) parameters are necessary (and sufficient) to describe a 3-D object. I note in passing that factors with eigenvalues much smaller than one can be very significant, disproving Guttman's (1954) criterion that such factors should be discarded.

The first factor (F I) is related to size, with large loadings on (i.e., high correlations with) all three axes. Loadings on the second and third factors (F II and F III) are bipolar, with some loadings positive and some negative (a common feature of FA; see Cooley and Lohnes 1971; Jöreskog et al. 1976). The bipolar loadings indicate that F II represents the ratio S to a combination of I and L, while F III represents the ratio I to L (or perhaps I to a combination of S and L: Table 1); these are obviously shape factors, representing sphericity and disc-rodness, and are similar to the results of Davis (1973, 1986). I define these three factors as the (fundamental) factors inherent in 3-D data sets; there are two shape factors and one size factor. I use this terminology throughout the paper to distinguish these mathematically derived fundamental factors from shape indices and size parameters.

Factors can be rotated in FA to try to simplify loadings, retaining only significant factors and discarding smaller factors, with variables having either large or small loadings on the rotated factors (Jöreskog et al. 1976; Davis 1986; and others). Because the three factors in this analysis must all be significant, as noted above, all three factors must be retained for rotation, and none can be discarded. Rotation in this case is trivial, and simply returns the original variables (Tables 1, 3 and 5). However, this rotation yields the alternate description of a 3-D object in terms of its 3 axes, similar to Cooley and Lohnes's (1971) study of 2-D objects, as discussed above. The storm-berm pebbles are an exception. These pebbles are discoid; consequently, their I and L axes are very similar, and there are in effect only 2 independent variables in this data set (see Table 7: rotated factor loadings).

ANALYSIS OF AXES AND INDICES

A variety (some say a plethora) of shape indices has been derived from the axes for the study of pebble shape

(Appendix). FA including these shape indices will show which are equivalent and which are orthogonal, i.e., independent, and hence the most useful. It would be doubly useful if some of the indices corresponded closely to the (fundamental) shape factors.

The eigenvalues for the FA of the data sets that include the derived indices clearly show 3 significant factors (F I, F II and F III), separated by an order-of-magnitude step from further insignificant factors (only F IV is shown in Tables 2, 4, 6 and 8 because further eigenvalues are irrelevant). Hence, loadings for only 3 factors are shown in the Tables, and 3 factors are retained for the factor rotation. This fact reinforces the earlier observation that 3 factors are logically required to describe the data fully, even though Davis's (1986, p. 552) intuition tells him that 2 factors are sufficient, contradicting his earlier conclusion (1986, p. 537) that 3 factors are required.

The factor loadings of the indices calculated for the set of randomly generated 3-D objects, as well as for samples of beach and river pebbles, show similar trends in all cases (Tables 2, 4, 6 and 8). Varimax rotation of these analyses simplifies loadings on the indices and reveals the bipolar relationships of the L, I and S axes. Many of the indices are roughly equivalent, and some indices do indeed correspond to the shape factors. Furthermore, the average of the axes corresponds to the size factor. The indices are grouped accordingly in the Tables.

Plotting the shapes of the random objects on their factor scores vividly illustrates the interpretation of the factors (Fig. 1). Plotting the size factor against the sphericity factor shows the effect of the size factor and illustrates that rods and discs are not separated by the sphericity factor. However, a plot of the sphericity factor against disc-rod factor clearly separates all shapes and suggests that shape diagrams should be triangular.

As noted above, the first shape factor, the sphericity factor, corresponds to the ratio S to (I and L). Not surprisingly, indices with high loadings on this factor (Tables 2, 4, 6 and 8) consist of geometric or arithmetic ratios of S to (I and/or L) (Appendix). These indices are S/I, S/L, flatness index, Corey shape index (CSI: previously termed "Corey shape factor"; renamed here to avoid confusion with the fundamental shape factors), and maximum projection sphericity (MPS). The optimum index is different for each of the samples. However, MPS, CSI, the flatness index, and S/L always have high loadings on the (rotated) sphericity factor and are usually orthogonal to the other factors, as indicated by low loadings on the other factors. Therefore, these indices are the best indicators of sphericity. The flatness index (Wentworth 1922) is actually a sphericity index! The flatness index is the inverse of the other sphericity indices (Appendix) and hence correlates negatively with the sphericity factor. Similarly, I/L is the inverse of the other disc-rod indices, and correlates negatively with the disc-rod factor.

The second shape factor, the disc-rod factor, corresponds approximately to the ratio I to L. Again not surprisingly, indices which include geometric ratios of I to L (Appendix), such as I/L, $(L - I)/(L - S)$ (which I call the disc-rod index, abbreviated to DRI) the Aschenbrenner shape factor, and $(L + S)/I$, usually have high loadings on this rotated factor (Tables 2, 4, 6 and 8). The optimum indices, with low loadings on the other factors, are usually DRI and $(L + S)/I$.

A third group of indices (Krumbein sphericity and rod index, $(S + I)/L$) consists of geometric or arithmetic combinations of L to I and S. These indices usually have high loadings on both shape factors. This is for most purposes not a useful attribute, so these indices should be discarded.

The relative importance of the 3 factors is dependent on the relative variation in the variables characterizing each sample. In some cases, F I is sphericity (Tables 4 and 8) and in others disc-rodness (Tables 2 and 6); F II varies in antipathy. This difference is not relevant to the present study.

DISCUSSION

Shape

One could use the fundamental factors directly in a study of pebbles. However, different samples would not be directly comparable, since the precise meaning of the factors is difficult to determine and would vary between samples. To facilitate comparison, it would be advantageous to use shape indices whose precise meaning is known, instead of shape factors.

Many researchers have proposed shape indices, often not realizing that they should be looking for two independent (i.e., orthogonal) shape indices. Those who have been trying to find a single shape index are doomed to failure (Williams 1965), as are those who disregard the requirement of orthogonality (Stratten 1974, and others). Also, the sphericity factor often seems to be more significant sedimentologically (sphericity correlates with settling velocity and rolling velocity: Sneed and Folk 1958), obscuring the other shape factor and sowing confusion among hapless researchers.

Zingg (1935) proposed S/I and I/L as indices for a shape diagram. These indices are orthogonal, but they do not correspond with the shape factors. S/I discriminates spheres and rods from discs and blades (a flatness index), and I/L discriminates spheres and discs from rods and blades (an elongation index).

Folk (in Sneed and Folk 1958) reasoned that shape should be plotted on a triangular diagram, the end-members of shape being spheres, discs and rods. S/L is plotted against DRI in Folk's form triangle; these indices are fairly orthogonal and are similar to the shape factors. However, it seems that Folk was unaware that S/L is very similar to sphericity. Folk also developed the concept of maximum projection sphericity and pointed out that a disc and a rod can have the same sphericity, i.e., sphericity is not only dependent on flatness but is dependent on both flatness and rodness. This property of sphericity is at first a little confusing, which is probably why Wentworth (1922) misnamed his flatness index, as mentioned above.

Sneed and Folk (1958) deduced the necessity of two shape indices, without fully realizing the requirement of orthogonality. As a result, their final choice for a disc-rod index, the "form ratio", correlates increasingly with sphericity for more discoid or rodlike objects. Hence, this index is increasingly un-orthogonal for extreme discs or rods. Again, Dobkins and Folk (1970) derived the oblate-prolate index, which also correlates increasingly with sphericity for more discoid or rodlike objects (Dobkins and Folk 1970, fig. 19); plotting MPS against the oblate-prolate index as they suggest is therefore of limited value because of this rather peculiar "semi-correlation".

Other studies of pebble parameters using FA (Humbert 1968; King and Buckley 1968; Orford 1975; Isla 1984) have also revealed one size factor and two shape factors. However, these are not always recognized as such! The disc-rod factors in many cases seems to correlate with roundness (Humbert 1968; Orford 1975), which has led these researchers to conclude incorrectly that the fundamental factors are a size factor, a single shape factor, and a roundness factor.

One size factor and two shape factors are revealed in the analysis of the axes of a variety of artificial and natural sets of 3-D objects (Tables 1–8, and the studies referred to above), indicating an underlying similarity in the data sets, be they synthetic random numbers or natural pebble samples.

Algebraic Approach to Shape Factors and Indices.—Factor analysis yields approximate rather than exact solutions. An algebraic approach may help to reveal the shape factors and identify the optimum indices more exactly.

The sphericity factor is represented by the ratio S to (I and L), according to the bipolar factor loadings. MPS, CSI, and the flatness index (Appendix) correspond to this ratio. S/I and S/L can be discarded as possible contenders for the sphericity factor, because they each only represent part of this ratio. However, the close relationship of these two indices with sphericity is apparent, as is illustrated by their product:

$$S/I \times S/L = S^2/IL = (MPS)^3 = (CSI)^2$$

The second shape factor distinguishes discs from rods (Fig. 1). According to the bipolar factor loadings, the ratio I to L (or perhaps I to S and L) represents this factor. However, I/L is not an efficient separator of discs and rods, as is apparent from a perusal of shape indices for some "perfect" discs, rods and blades. I/L can have different values for "perfect rods" (8, 4, 4 and 8, 2, 2) and cannot distinguish a blade (8, 4, 2) from a rod (8, 4, 4) (Table 9). The only proper disc-rod index is DRI! None of the other indices is able to distinguish discs from rods unambiguously (Table 9). Interestingly, combinations of indices, as used in shape diagrams, are able to distinguish all shapes, illustrating the effectiveness of shape diagrams.

The Optimum Shape Indices.—Which are the best indices to use? Two orthogonal indices should be chosen, preferably similar to the (fundamental) shape factors. The indices should also have normal distributions for natural pebble populations to facilitate statistical calculations. The best choice for sphericity is either MPS or CSI, because they both comply with the above requirements and also have other sedimentologically desirable properties. MPS corresponds to a particle's settling velocity ($R = 0.97$) and rolling velocity ($R = 0.86$) (Sneed and Folk 1958). CSI ($=(MPS)^{3/2}$) also correlates well with settling velocity ($R = 0.98$ according to my calculations, using data from Sneed and Folk 1958, fig. 4). According to the factor loadings, CSI is slightly better than MPS, because CSI consistently has slightly higher loadings on the sphericity factor and lower loadings on the other factors (Tables 2, 4, 6 and 8). Thus CSI seems the better choice, but the differences are slight. Consideration of shape diagrams will help to make the final choice (see below).

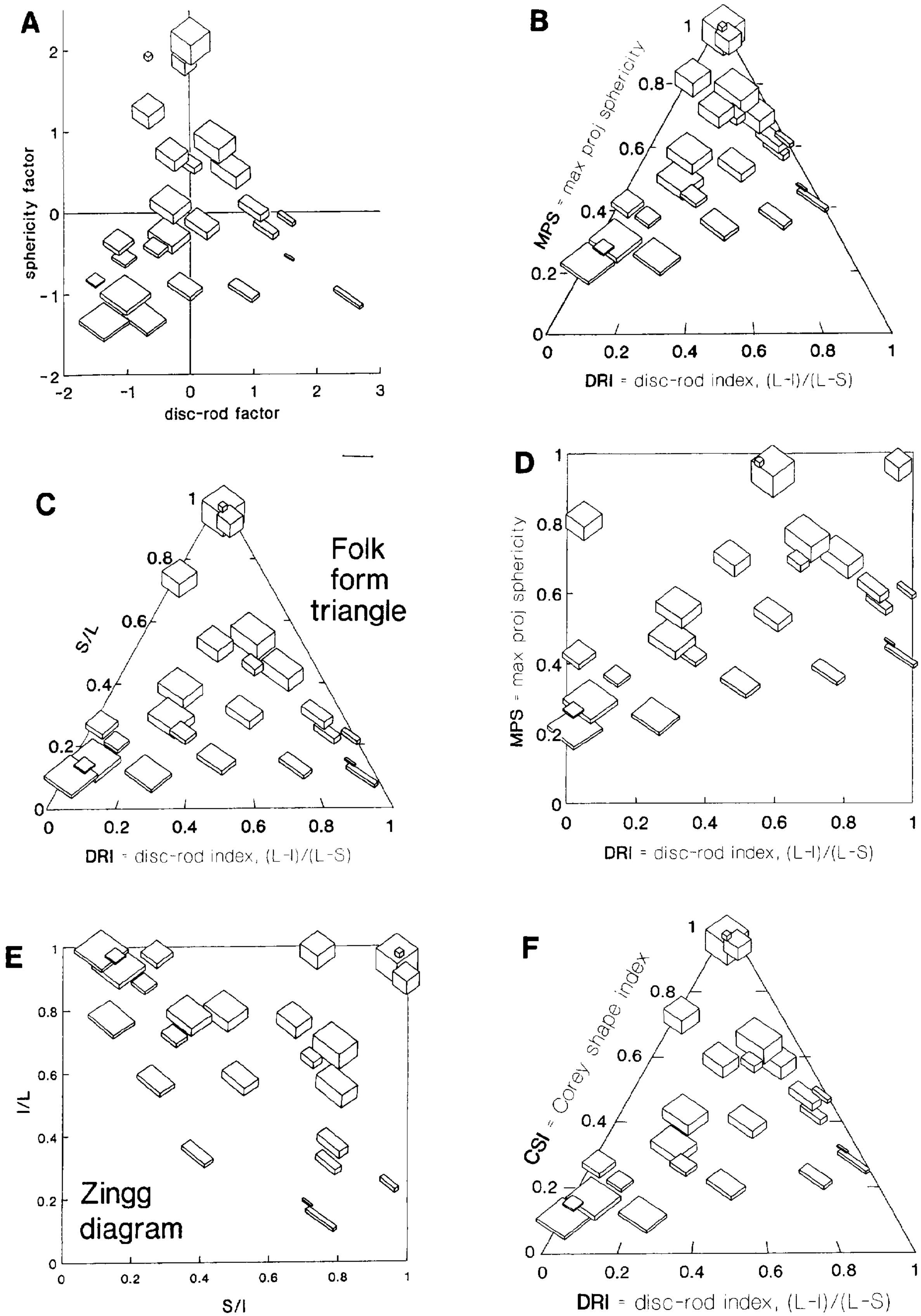
The only disc-rod index is DRI, so this is the only candidate for the disc-rod factor. However, this conclusion is based on the inference that one of the shape factors separates discs from rods (Fig. 1). This does seem to be a reliable inference, especially when shape diagrams are considered (below).

Shape Diagrams.—A close inspection of the shape factor diagram (Fig. 1) suggests that a triangular diagram is the most appropriate for the expression of shape. This is because disc-rodness becomes less important as sphericity increases, and confirms Folk's (in Sneed and Folk 1958) reasoning that a triangular diagram should be used for shape, the 3 end-members of shape being spheres, discs and rods.

The correspondence between the shape factor diagram and triangular plots of CSI versus DRI and MPS versus DRI is illustrated by plotting the random objects on shape diagrams (Fig. 2). Folk's form triangle (S/L vs. DRI) is also similar. The diagram which corresponds closest with the shape factor diagram is CSI versus DRI.

Comparing the triangular plot of MPS versus DRI with the rectangular plot of MPS versus DRI (Fig. 2) illustrates the distortion which results from using a rectangular diagram. This comparison yields another clue in the search for the optimum indices: because shape should ideally be displayed on a triangular diagram, a valid method to determine statistical parameters like mean and standard deviation would be to contour point density on a triangular shape diagram, analogous to the technique used for point contouring on an equal-area stereogram (Davis 1986, p. 339, and others). Hence, the shape diagram should be an equal-area projection. DRI is suitable for an equal-area projection, being a linear ratio; the index used for sphericity should also be linear. CSI is a linear ratio, unlike MPS which is a power ratio. Thus the ultimate choice for sphericity is CSI, and the optimum shape diagram is a triangular plot of CSI versus DRI. The shape classes proposed for this diagram (Fig. 3) correspond approximately to the shape classes on Folk's form triangle.

The Zingg diagram (Fig. 2E), although efficiently separating shapes, has some disadvantages. The shape factor axes are rotated through 45 degrees: the sphericity axis runs diagonally across the diagram from the lower left corner to the upper right corner, and the disc-rod axis lies at 90 degrees to this. This obscures the shape factors, as well as the triangular nature of the shape factor diagram.


A
sphericity factor
disc-rod factor
B
MPS = max proj sphericity
DRI = disc-rod index, (L-I)/(L-S)
C
Folk
form
triangle
S/L
DRI = disc-rod index, (L-I)/(L-S)
D
MPS = max proj sphericity
DRI = disc-rod index, (L-I)/(L-S)
E
Zingg
diagram
I/L
S/I
F
CSI = Corey shape index
DRI = disc-rod index, (L-I)/(L-S)

Interestingly, the indices used in a shape diagram do not have to be orthogonal. A comparison of the plots of CSI versus DRI and S/L versus DRI shows the slight distortion resulting from slightly different indices. Increasingly unorthogonal indices result in increasingly distorted shape diagrams, which are still effective, however, in discriminating shape. Even S/I versus S/L (not illustrated in this paper) is an effective shape diagram.

The triangular nature of the shape diagram probably results from the constraints in the 3-D axial data. As noted before, by definition the L axis must be greater than the I axis, which in turn must be greater than the S axis. It seems that some of the independence in the data is lost due to this constraint, and this loss is reflected in the triangular shape diagram being more appropriate than a rectangular one.

Any correlation between two orthogonal indices as used in shape diagrams is spurious. However, such a correlation may arise through geological processes (e.g., waves move discs of low sphericity onto beach berms, and rods and spheres to the subtidal beach zone: Illenberger, unpublished data).

Mean and Standard Deviation of Shape Indices.—As discussed above, the average and standard deviation of CSI and DRI should be determined using a 2-D technique such as contouring point density on the CSI versus DRI diagram. However, if a sample is well shape-sorted, the error resulting from calculating the mean and standard deviation of CSI directly will be minimal. For samples with poor shape-sorting, CSI should have a skewed distribution with the mode smaller than the mean, because the shape diagram is triangular.

If a sample is well shape-sorted, the error resulting from calculating DRI directly will also be small. In this case, standard deviation can also be calculated directly with only a small error; a correction has to be applied:

corrected std. dev.
= directly calculated std. dev. × (1 − CSI)

This correction compensates for the decreasing significance of DRI as objects become more spherical.

This phenomenon shows that the shape factors are inter-related and that the indices cannot be considered on their own when determining statistical parameters. Plotting data on the CSI versus DRI shape diagram will help to avoid misinterpretation of data.

←

FIG. 2.—Random objects plotted on various shape diagrams. A) Sphericity factor versus disc-rod factor (repeated from Fig. 1B for comparison). B) Maximum projection sphericity (MPS) versus disc-rod index (DRI)-triangular plot. C) Folk's form triangle. D) MPS versus DRI-rectangular plot. E) Zingg diagram. F) Corey shape index (CSI) versus DRI.

Comparison of triangular and rectangular plots of MPS versus DRI illustrates the distortion introduced by the rectangular form. Folk's diagram and CSI versus DRI are triangular; the latter is the most similar to the shape factor plot (Fig. 1B). The relationship between the Zingg diagram and the others is apparent: the shape factor axes are rotated through 45°—the sphericity axis runs diagonally across the diagram from the lower left corner to the upper right corner, and the disc-rod axis lies at 90° to this.

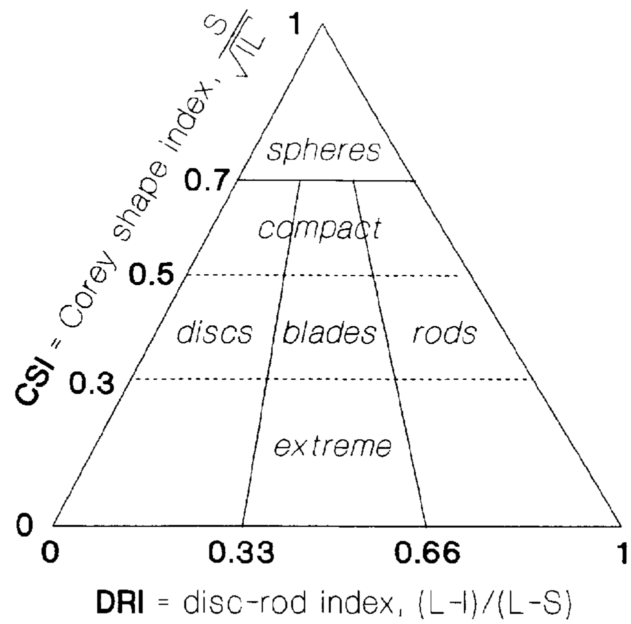


FIG. 3.—The shape diagram proposed in this paper: Corey shape index (CSI) versus disc-rod index (DRI) on a triangular diagram. The shape classes defined on the diagram correspond approximately to the shape classes on Folk's form triangle.

Size

The arithmetic mean of the axes usually has the highest loading on the rotated size factor and is more independent of the shape factors than the individual axes (Tables 2, 4, 6 and 8). Other interpretations of size, such as the geometric mean or volume, do not have such high loadings on the size factor (Table 2). (There is a high degree of inter-correlation between the axes of the pebble samples, because they have good size-sorting. This inter-correlation made it impossible to include the geometric mean and volume in the FA of these samples, because non-invertible matrices were generated in these cases.) Natural pebble populations usually have lognormal size distributions, so the logarithm of the mean size should be used as the optimum expression of size. The phi scale, a logarithmic scale, may of course be used.

As mentioned in the Introduction, there are some spurious correlations between the axes and shape indices. Thus, there will also be spurious correlations between axes and shape factors, particularly between the S axis and the sphericity factor, and to a lesser extent between the I axis and the disc-rod factor (Tables 2, 4, 6 and 8). The spurious correlation between the S axis and sphericity occurs because in any sample the less spherical objects will have shorter S axes; on average these objects will also be smaller. The plot of the size factor against the sphericity factor (Fig. 1A) illustrates this spurious correlation: objects with small S axes usually plot towards the lower left corner, and objects with large L axes plot in the upper right corner. Similarly, the I axis correlates with the disc-rod factor, because high values of I/L imply a large I axis.

Thus, using only one axis to represent size is incorrect, particularly the S axis because this axis correlates with size and sphericity. Numerous researchers have made this mistake (Carr et al. 1970; Williams and Caldwell 1988; and others); their conclusions that "size" (S axis) has the most significance in beach pebble dynamics actually means that a combination of size and sphericity has the most significance.

CONCLUSIONS

If three series of random numbers are ordered into triplets of large, intermediate and small, these triplets will exhibit properties akin to the long, intermediate and small axes of a pebble sample. In particular, FA will extract three factors, which are the three fundamental factors inherent in the dimensions of any population of 3-D objects, including pebbles. One factor corresponds to the size, and two further factors correspond to the shape (sphericity and disc-rodness) of the triplets.

Size and shape are related, although they are also orthogonal and independent (not too surprisingly).

Size is best described with the mean of the axes; for natural pebble populations the logarithm of the mean should be used to normalize the distribution. Sphericity is best described with CSI. Some aspects of the sedimentary behavior of a pebble (settling velocity and, to a lesser extent, rolling velocity) correlate with sphericity. The index that best represents the disc-rod factor is DRI. The most effective shape diagram is a triangular plot of these two shape indices; the three end-members of shape are spheres, discs and rods. These indices consistently have high loadings on the fundamental factors inherent in 3-D objects, whether using random shapes or pebble samples, indicating the robustness of these indices.

Shape should ideally be displayed on a triangular diagram of CSI versus DRI. Consequently, a valid method to determine the mean shape of a number of objects would be to contour point density on this diagram. However, if a sample population is well shape-sorted, the error in approximating the mean shape by averaging the indices will be small. The calculation of standard deviation of shape indices also requires special techniques. Thus the indices cannot be treated in isolation, indicating an underlying relationship between the shape indices and showing that shape is best investigated with the CSI versus DRI shape diagram.

Using a data set whose basic properties are known gives insight into the workings of FA. Eigenvalues much smaller than one may be significant in FA, if only primary variables are analyzed. FA extracts semi-fundamental factors, in the sense that the factor weights are dependent on the data set analyzed. FA needs an excess of derived variables "covering the factor field" to develop a clear distinction between significant and insignificant eigenvalues. Only in this case will Guttman's (1954) criterion (that factors with eigenvalues smaller than one be discarded) be valid. Davis (1986, p. 537) did not use any derived variables related to the disc-rod factor in his FA; consequently he "lost" this factor in his eigenvalues. The requirement that data should have normal distributions for FA needs to be re-examined, because analysis of the random data set gave results very similar to the analysis of the data sets with normal distributions.

ACKNOWLEDGMENTS

This paper originated from deliberations on pebble shape and size with Koos Reddering and Allan Williams, and from enlightening discussions on factor analysis with (in alphabetical order) Ted Donn, John (Speedy) Gonsalves, Gerhard Mels, Julian Orford, Peter Smith, Francis Thackeray and Oscar A.C. Williams. I am grateful to Jenny Burkinshaw, Johan Hattingh, various referees and editors, and my promoter, Izak Rust, for beneficial comments on the manuscript. The Council for Scientific and Industrial Research, South Africa, financed this study through the South African National Committee for Oceanographic Research.

REFERENCES

Aschenbrenner, B.C., 1956, A new method of expressing particle sphericity: Journal of Sedimentary Petrology, v. 26, p. 15–31.

Blatt, H., Middleton, G., and Murray, R., 1972, Origin of Sedimentary Rocks: Englewood Cliffs, NJ, Prentice-Hall, 634 p.

Carr, A.P., Gleason, R., and King, A., 1970, Significance of pebble size and shape in sorting by waves: Sedimentary Geology, v. 4, p. 89–101.

Cooley, W.W., and Lohnes, P.R., 1971, Multivariate Data Analysis: New York, Wiley, 364 p.

Corey, A.T., 1949, Influence of shape on fall velocity of sand grains [unpubl. M.S. thesis]: Colorado A&M College, 102 p. Quoted in Dobkins and Folk 1970.

Davis, J.C., 1973, Statistics and Data Analysis in Geology: New York, Wiley, 550 p.

Davis, J.C., 1986, Statistics and Data Analysis in Geology, 2nd Edition: New York, Wiley, 646 p.

Dobkins, J.E., and Folk, R.L., 1970, Shape development on Tahiti-Nui: Journal of Sedimentary Petrology, v. 40, p. 1167–1203.

Guttman, L., 1954, Some necessary conditions for factor analysis: Psychometrika, v. 19, p. 149–161. Quoted in Jöreskog et al. 1976.

Humbert, F.L., 1968, Selection and wear of pebbles on gravel beaches: University of Groningen, Geological Institute Publication No. 190, 144 p.

Isla, F.I., 1984, Caracteristicas texturales y comportamiento hidrodinamico de los rodados de playa: tecnicas multivariadas y experiencia de transporte en condiciones episodicas: Revista de la Asociacion Argentina da Mineralogia, Petrologia y Sedimentologia, v. 15, p. 33–45.

Jöreskog, K.G., Klovan, J.E., and Reyment, R.A., 1976, Geological Factor Analysis: Amsterdam, Elsevier, 178 p.

King, C.A.M., and Buckley, J.T., 1968, The analysis of stone size and shape in arctic environments: Journal of Sedimentary Petrology, v. 38, p. 200–214.

Krumbein, W.C., 1941, Measurement and geologic significance of shape and roundness of sedimentary particles: Journal of Sedimentary Petrology, v. 11, p. 64–72.

Orford, J.D., 1975, Discrimination of particle zonation on a pebble beach: Sedimentology, v. 22, p. 441–463.

Pearson, K., 1897, Mathematical contributions to the theory of evolution. On a form of spurious correlation which may arise when indices are used in the measurement of organs: London, Proceedings of the Royal Society, v. 60, p. 489–498.

Sneed, E.D., and Folk, R.L., 1958, Pebbles in the lower Colorado River, Texas: a study in particle morphogenesis: Journal of Geology, v. 66, p. 114–150.

Stratten, T., 1974, Notes on the application of shape parameters to differentiate between beach and river deposits in southern Africa: Transactions of the Geological Society of South Africa, v. 77, p. 59–64.

WENTWORTH, C.K., 1922, A field study of the shapes of river pebbles: U.S. Geological Survey Bulletin 730-C, p. 103–114. Quoted in Dobkins and Folk 1970.

WILLIAMS, E.M., 1965, A method of indicating pebble shape with one parameter: Journal of Sedimentary Petrology, v. 35, p. 993–996.

WILLIAMS, A.T., AND CALDWELL, N.E., 1988, Particle size and shape in pebble-beach sedimentation: Marine Geology, v. 82, p. 199–215.

ZINGG, T., 1935, Beitrag zur Schotteranalyse: Schweizerische Mineralogische und Petrologische Mitteilungen, v. 15, p. 39–140.

APPENDIX

Shape Indices and Size-related Expressions

Note that 10 (L + I)/2S, (L + I)/2S and (L + I)/S are exactly equivalent.

	Formula	Author
Long	L	
Intermediate	I	
Short	S	
Mean	(L + I + S)/3	
Geometric mean	$(LIS)^{1/3}$	
Volume	LIS	
S/I	S/I	Zingg 1935
S/L	S/L	Folk (in Sneed and Folk 1958)
Flatness index	(L + I)/S	Wentworth 1922
CSI = Corey shape index	$S/(IL)^{1/2}$	Corey 1949: renamed in this paper (originally named Corey shape factor)
MPS = max proj sphericity	$(S^2/IL)^{1/3}$	Folk (in Sneed and Folk 1958)
I/L	I/L	Zingg 1935
(L + S)/I	(L + S)/I	This paper
DRI = disc-rod index	(L − I)/(L − S)	Folk (in Sneed and Folk 1958): named in this paper
Oblate-prolate index	10((L − I)/(L − S) − 0.5)/(S/L)	Dobkins and Folk 1970
Aschenbrenner factor	LS/I^2	Aschenbrenner 1956
Rod index	(S + I)/L	This paper
Krumbein sphericity	$(SI/L^2)^{1/3}$	Krumbein 1941

TABLE 1.—*Eigenvalues and factor loadings: random data set, three axes only*

	FI	FII	FIII	Varimax Rotated FI	Varimax Rotated FII	Varimax Rotated FIII
Eigenvalue	1.79	0.82	0.39	1.01	1.00	0.97
Variance %	59	27	13	34	33	32
	Factor Loadings			Rotated Factor Loadings		
Long	0.80	−0.45	0.39	0.95	0.08	0.29
Intermediate	0.88	−0.13	−0.46	0.31	0.19	0.93
Short	0.61	0.77	0.15	0.08	0.98	0.16

TABLE 2.—*Eigenvalues and factor loadings: random data set, with size and shape indices. See Appendix for explanation of indices*

		FI	FII	FIII	FIV	Varimax Rotated FI	Varimax Rotated FII	Varimax Rotated FIII
	Eigenvalue	8.59	7.10	3.53	0.26	7.08	7.03	5.13
	Variance %	43	36	18	1.3	35	35	26
		Factor Loadings				Rotated Factor Loadings		
		FI	FII	FIII		FI	FII	FIII
Size	Long	0.27	−0.26	0.90		0.14	−0.38	0.89
	Intermediate	0.88	−0.31	0.30		−0.65	−0.07	0.73
	Short	0.59	0.73	0.30		0.00	0.78	0.60
	Mean	0.74	−0.03	0.66		−0.25	0.05	0.96
	log mean	0.70	−0.07	0.67		−0.23	−0.01	0.95
	Geometric mean	0.80	0.26	0.53		−0.23	0.36	0.90
	Volume	0.70	0.41	0.48		−0.12	0.48	0.80
Sphericity	S/I	−0.23	0.95	0.03		0.60	0.78	−0.07
	S/L	0.43	0.84	−0.28		−0.11	0.98	0.03
	log flatness	−0.24	−0.94	0.13		−0.14	−0.97	−0.05
	CSI	0.23	0.95	−0.17		0.13	0.98	0.01
	MPS	0.19	0.97	−0.14		0.18	0.98	0.01
Disc-rod	I/L	0.86	−0.11	−0.45		−0.93	0.30	0.11
	(L + S)/I	−0.86	0.24	0.34		0.92	−0.14	−0.19
	log(L + S)/I	−0.83	0.41	0.35		0.98	0.01	−0.17
	DRI	−0.65	0.61	0.30		0.89	0.26	−0.10
	Oblate-prolate	−0.74	0.56	0.31		0.95	0.19	−0.14
	log Aschenbr	−0.68	0.68	0.27		0.93	0.32	−0.13
	log rod	−0.87	−0.25	0.41		0.76	−0.62	−0.16
	Krumbein sph	0.69	0.62	−0.36		−0.45	0.88	0.10

TABLE 3.—*Eigenvalues and factor loadings: small beach pebbles, three axes only*

				Varimax Rotated		
	FI	FII	FIII	FI	FII	FIII
Eigenvalue	2.19	0.60	0.21	1.04	0.99	0.97
Variance %	73	20	7	35	33	32
	Factor Loadings			Rotated Factor Loadings		
Long	0.90	−0.29	0.32	0.23	0.88	0.41
Intermediate	0.91	−0.26	−0.33	0.25	0.42	0.87
Short	0.74	0.67	0.01	0.96	0.20	0.20

TABLE 4.—*Eigenvalues and factor loadings: small beach pebbles, with size and shape indices. See Appendix for explanation of indices*

						Varimax Rotated		
		FI	FII	FIII	FIV	FI	FII	FIII
	Eigenvalue	7.63	6.18	2.67	0.3	7.08	6.22	3.17
	Variance %	45	36	16	1.8	42	37	19
		Factor Loadings				Rotated Factor Loadings		
		FI	FII	FIII		FI	FII	FIII
Size	Long	0.17	−0.73	−0.65		−0.23	0.39	0.88
Size	Intermediate	−0.03	−0.24	−0.96		−0.08	−0.18	0.97
Size	Short	0.80	0.03	−0.57		0.71	0.24	0.63
Size	log mean	0.29	−0.49	−0.81		0.02	0.24	0.96
Sphericity	S/I	0.96	0.22	0.01		0.90	0.40	0.04
Sphericity	S/L	0.76	0.63	−0.08		0.99	−0.04	−0.04
Sphericity	log flatness	−0.84	−0.48	−0.07		−0.95	−0.18	0.10
Sphericity	CSI	0.88	0.46	−0.03		0.98	0.17	0.00
Sphericity	MPS	0.89	0.46	0.01		0.98	0.19	−0.04
Disc-rod	I/L	−0.33	0.91	−0.23		0.30	−0.93	−0.14
Disc-rod	(L + S)/I	0.75	−0.61	0.15		0.22	0.93	0.17
Disc-rod	log(L + S)/I	0.76	−0.61	0.19		0.22	0.95	0.13
Disc-rod	DRI	0.50	−0.73	0.27		−0.06	0.94	0.09
Disc-rod	Oblate-prolate	0.70	−0.62	0.29		0.15	0.97	0.05
Disc-rod	log Aschenbr	0.92	−0.29	0.21		0.54	0.83	0.04
	log rod	−0.23	−0.96	0.15		−0.77	0.60	0.15
	Krumbein sph	0.46	0.87	−0.09		0.90	−0.39	−0.14

TABLE 5.—*Eigenvalues and factor loadings: river pebbles, three axes only*

	FI	FII	FIII	Varimax Rotated FI	Varimax Rotated FII	Varimax Rotated FIII
Eigenvalue	2.84	0.11	0.05	1.18	1.16	0.66
Variance %	95	3.7	1.6	39	39	22
	Factor Loadings			Rotated Factor Loadings		
Long	0.97	−0.20	0.12	0.48	0.79	0.38
Intermediate	0.98	−0.06	−0.17	0.54	0.57	0.62
Short	0.96	0.26	0.05	0.81	0.46	0.36

TABLE 6.—*Eigenvalues and factor loadings: river pebbles, with size and shape indices. See Appendix for explanation of indices*

		FI	FII	FIII	FIV	Varimax Rotated FI	Varimax Rotated FII	Varimax Rotated FIII
	Eigenvalue	7.52	5.78	3.43	0.09	7.20	5.79	3.78
	Variance %	42	34	20	0.6	42	34	22
		Factor Loadings FI	Factor Loadings FII	Factor Loadings FIII		Rotated Factor Loadings FI	Rotated Factor Loadings FII	Rotated Factor Loadings FIII
Size	Long	0.17	0.52	0.83		0.07	−0.22	0.97
Size	Intermediate	0.36	0.32	0.87		−0.18	−0.11	0.97
Size	Short	0.07	0.17	0.96		0.01	0.16	0.97
Size	Log mean	0.20	0.47	0.82		0.02	−0.19	0.95
Sphericity	S/I	−0.91	−0.37	0.18		0.66	0.74	−0.08
Sphericity	S/L	−0.27	−0.92	0.28		−0.14	0.98	−0.07
Sphericity	log flatness	0.59	0.77	−0.24		−0.21	−0.97	0.11
Sphericity	CSI	−0.64	−0.72	0.26		0.27	0.96	−0.09
Sphericity	MPS	−0.64	−0.72	0.25		0.28	0.96	−0.10
Disc-rod	I/L	0.76	−0.64	0.09		−0.96	0.28	−0.01
Disc-rod	(L + S)/I	−0.93	0.33	0.02		0.98	0.10	−0.02
Disc-rod	log(L + S)/I	−0.95	0.31	0.03		0.99	0.12	−0.02
Disc-rod	DRI	−0.93	0.28	0.02		0.96	0.13	−0.03
Disc-rod	Oblate-prolate	−0.95	0.30	0.03		0.99	0.13	−0.02
Disc-rod	log Aschenbr	−0.99	0.10	0.07		0.94	0.33	−0.05
	log rod	−0.36	0.91	−0.19		0.71	−0.70	0.06
	Krumbein sph	0.22	−0.95	0.22		−0.60	0.80	−0.06

TABLE 7.—*Eigenvalues and factor loadings: storm berm pebbles, three axes only*

	FI	FII	FIII	Variamax Rotated FI	Variamax Rotated FII	Variamax Rotated FIII
Eigenvalue	2.26	0.67	0.07	1.86	1.06	0.08
Variance %	75	22	2.5	62	35	3
	Factor Loadings			Rotated Factor Loadings		
Long	0.95	−0.25	0.19	0.94	0.25	0.21
Intermediate	0.94	−0.28	−0.19	0.96	0.23	−0.18
Short	0.68	0.73	−0.01	0.24	0.97	0.01

TABLE 8.—*Eigenvalues and factor loadings: storm berm pebbles, with size and shape indices. See Appendix for explanation of indices*

		FI	FII	FIII	FIV	Varimax Rotated FI	Varimax Rotated FII	Varimax Rotated FIII
	Eigenvalue	8.61	5.01	2.96	0.21	6.76	6.49	3.33
	Variance %	51	29	17	1.2	40	38	20
		Factor Loading				Rotated Factor Loadings		
		FI	FII	FIII		FI	FII	FIII
Size	Long	0.46	−0.38	0.78		−0.36	0.03	0.92
Size	Intermediate	0.60	−0.05	0.77		−0.25	−0.31	0.89
Size	Short	−0.35	0.10	0.90		0.56	0.18	0.77
Size	log mean	0.38	−0.17	0.89		−0.14	−0.08	0.97
Sphericity	S/I	−0.98	0.08	0.15		0.80	0.57	−0.09
Sphericity	S/L	−0.79	0.57	0.18		0.98	0.08	−0.10
Sphericity	log flatness	0.92	−0.36	−0.07		−0.92	−0.32	0.20
Sphericity	CSI	−0.91	0.35	0.17		0.93	0.33	−0.10
Sphericity	MPS	−0.93	0.34	0.14		0.93	0.34	−0.13
Disc-rod	I/L	0.36	0.93	0.04		0.33	−0.94	−0.01
Disc-rod	(L + S)/I	−0.74	−0.65	0.05		0.15	0.98	−0.03
Disc-rod	log(L + S)/I	−0.78	−0.63	0.04		0.19	0.98	−0.05
Disc-rod	DRI	−0.67	−0.70	0.06		0.07	0.97	0.00
Disc-rod	Oblate-prolate	−0.83	−0.49	−0.02		0.29	0.91	−0.15
Disc-rod	log Aschenbr	−0.95	−0.30	0.03		0.51	0.84	−0.15
	log rod	0.24	−0.96	−0.12		−0.81	0.58	0.08
	Krumbein sph	−0.53	0.83	0.11		0.94	−0.29	−0.14

TABLE 9.—*Shape indices for some "perfect" shapes, illustrating that DRI = (L − I)/(L − S) is the only proper disc-rod index. I/L can have different values for "perfect rods" (8,4,4 and 8,2,2) and cannot distinguish a blade (8,4,2) from a rod (8,4,4). Similarly, none of the other indices is able to distinguish discs from rods. However, the combination of I/L and S/I will differentiate discs and rods unambiguously, as will other combinations. I^2/SL is the equivalent of the Aschenbrenner shape factor; SI/L^2 is the equivalent of Krumbein sphericity; S^2/IL is the equivalent of CSI and MPS*

	L	I	S	DRI	I/L	S/I	S/L	I^2/SL	SI/L^2	S^2/IL
disc	8	8	4	0	1	0.5	0.5	2	0.5	0.25
blade	8	4	2	0.67	0.5	0.5	0.25	1	0.125	0.125
rod	8	4	4	1	0.5	1	0.5	0.5	0.25	0.5
rod	8	2	2	1	0.25	1	0.25	0.25	0.0625	0.25