John Winkle

STATISTICAL PROGRAMS AVAILABLE FOR COMPUTER ANALYSIS OF ROCK ART

The feasibility of using computerized statistical analysis to study prehistoric rock art has been investigated. This approach has not been used extensively in previous rock art studies. Therefore, an effort has been made to define the significant aspects of rock art research that lend themselves to statistical investigations.

Rock art research has moved in recent years away from a mere cataloging of petroglyphs and rock paintings toward a scientific investigation of stylistic elements. At the same time, emphasis appears to be shifting away from earlier, more subjective interpretations of rock art towards interpretations of a quantitative nature. Thus, a present evaluation of rock art is likely to concentrate on identifiable stylistic elements whose basic size, shape, frequency of occurrence, geographical distribution and context could be statistically quantified. Such an approach promises to reveal a wealth of information about the movements of prehistoric peoples and ideas without a need for subjective, and perhaps erroneous, interpretations of rock art elements.

An extensive use of a statistically based evaluation of rock art types is likely to require the use of computerized statistical packages and large amounts of computer memory.

The principal advantage of a computerized statistical approach lies in the ability of computers to process vast amounts of information. Other advantages include increased capability to cross-correlate data from several sites, and ease of data retrieval.

The use of computers also introduces certain disadvantages. The ability to process vast amounts of information may tempt the user to overuse the computer; for instance, a full computerized approach should not be necessary when analyzing a limited number of sites, however, it would be highly desirable when analyzing all statewide occurrences of a particular figure type. The capability to easily cross-correlate data might similarly lead to the generation of vast amounts of useless information. Finally, the ease of data retrieval entails highly structured data input. This implies forethought and meticulous attention to formatting the data.

Before computerized statistical packages are applied to rock art research, an effort must be made to define the desired features and scope of a statistical approach. A preliminary definition might include the following desired features:

- The approach should begin with a manageable method of handling input. Computerized approaches necessarily involve presorting of the data into elements and figure types by the researcher. Consequently, figure types should be clearly defined and held to a reasonable number.
- Enough cases to provide a statistically relevant analysis should exist for each of the basic figure types that are defined. Unique rock art cases that fall outside the basic element, or figure type categories, should be noted by the researcher, so they are not lost in a statistical fog.

- The output from a statistical analysis of rock art elements and figure types should be easy to comprehend and interpret. Perhaps the clearest output format would be a series of frequency plots, such as those generated by Schaafsma in *The Rock Art of Utah*, and appropriate correlation data.

Similarly, the scope of the desired work must be defined before applying a computerized statistical approach. Important in this regard are such variables as: the number of rock art elements, figure types, the number of attributes associated with each basic figure type, size, location, and perhaps association with known habitation and/or hunting sites. These, and any other desired variables, will effect the scope of the statistical analysis. In turn, the scope of the analysis will effect the required input format. Ideally a standardized system with a broad enough scope of entered data would be agreed upon by a majority of rock art researchers so that it would be useful to future research and so that results can be easily compared.

Warner and others are proposing a standardized classification system so any figure type can be easily described by a computer code consisting of approximately ten characters. To illustrate how the number of input variables could effect the scope of a statistical analysis, consider some small statistics programs that appear to be limited to approximately 500 total input values. If each petroglyph must be described by ten characteristic variables, a total of 50 petroglyphs could be statistically compared with each other at a time. In addition to the limitations this would place on the scope of the analysis, it would probably lead to highly specific figure types. This would effect the quality of the analysis. The ten character classification system could, however, be altered by a preprocessing routine such that the descriptive information would go along for the ride while the actual analysis considered a more basic figure type or element; for instance: "abstract", "abstract circle", "anthropomorph", etc.

While the number of cases and variables does not appear to be a limitation in larger statistical programs, the possibility of statistical inaccuracy due to over specification of cases remains a valid concern. The number of variables used to describe a particular petroglyph will also have a significant bearing on analysis costs.

A set of criteria used for selecting a statistical program involves the following:

- **Availability**: The program must be locally available and affordable.
- **Accessibility:** The program must be accessible from an input/output standpoint and should be accessible on a large number of different machines.
- **Capability**: The program must have sufficient subroutines and options to address the desired problem.

Given this set of selection criteria, both small- and large-scale statistical programs have been evaluated.

Nearly all programmable, hand-held calculators (such as the T159, HP67 or HP41CV) contain fundamental statistical packages. These packages include mean, standard deviation, coefficient of variation, and linear regression subroutines that might be useful in dealing with quantitative data such as petroglyph size or distance from a central location. In addition, frequency of occurrence could be easily calculated for a small set of data. These features are not as automatic as they would be in a large-scale statistical package. However, hand-held calculators are readily available, easily affordable, and would be accessible to most researchers. Therefore, the use of programmable calculators should not be ruled

out when analyzing a relatively small set of data.

When dealing with a large amount of data, such as state-wide occurrences of several figure types, large-scale statistical packages will be necessary. *A Statistical Package for the Social Sciences* (SPSS) would be an example of this type of program. SPSS is limited to 500 independent variables while the number of cases is determined only by the amount of memory available in the computer being used. The program is available and is accessible on a number of different computers, including IBM, CDC, XEROX, and Univac Machines. SPSS contains a full library of statistical functions. The functions listed in the section above are also available in SPSS.

The costs of Using SPSS or a similar large-scale program for rock art research analysis have not yet been determined. Several other large-scale statistical programs need to be evaluated and reported upon. A future update of this topic and its implications to rock art research will be necessary.