CONTEMPORARY STATISTICAL MAPS
EVIDENCE OF SPATIAL AND GRAPHIC IGNORANCE

George F. Jenks
University of Kansas

INTRODUCTION

Maps are windows into the minds of their creators and, if one peeks into these inner recesses, evidences of spatial and graphic ignorance become readily apparent. An inspection of a jumble of statistical maps taken from magazines, newspapers, professional journals and governmental reports gives rise to the following semantic range of contemporary map qualities: 1/

communicative/secretive	simplistic/complex
clear/indistinct	pleasing/ugly
directed/haphazard	accurate/erroneous
decorative/useful	

What do these polar adjectives imply about modern statistical maps? 2/ First, there is a confusion in the minds of many cartographers as to the function of a statistical map. Second, the graphic language, or the conventional symbology of thematic mapping, which has been developed through centuries of cartographic experience, is not part of the vocabulary of many map makers. Third, some statistical map makers do not seem to sense the relationship between data processing and the fidelity with which their map portrays the information they wish to convey. Last, too many map makers are either not trained in, or appreciative of, the subtleties of graphic communication. In this paper these deficiencies of contemporary statistical map making are examined, corrective procedures are suggested, and some opportunities for improving statistical maps are proposed.

WHY MAKE STATISTICAL MAPS?

If one raises the question, Why make statistical maps?, he poses the query because he assumes that there is a logical need behind each cartographic display. Additionally, since statistical maps are often included in a larger work, one can also assume that there is something about a map which makes it more useful than words, tables or graphs in the transfer of certain types of information. The enormous utility of a statistical map can easily be demonstrated if the reader will refer to Table 1, a typical censual map of enumerated data. 3/ Attempt to describe the spatial information contained in these tables. Most readers, even if provided with a base map (Figure 1) are unable to integrate areally the twenty-nine cropland acreage values given in the table. Furthermore, if the reader is provided with an "areal table" (Figure 2), he will still find it difficult to perceive "in his mind's eye" the distributional characteristics of cropland acreage in Utah.

Table 1

Cropland Acreage, by Counties, Utah, 1969

Beaver	29,917	Piute	15,302
Box Elder	360,571	Rich	66,550
Cache	176,926	Salt Lake	69,415
Carbon	14,692	San Juan	91,299
Daggett	8,106	Sanpete	98,029
Davis	40,946	Sevier	52,320
Duchesne	96,035	Summit	38,218
Emery	48,344	Tooele	39,643
Garfield	23,714	Uintah	93,023
Grand	3,132	Utah	139,987
Iron	65,973	Wasatch	20,116
Juab	77,275	Washington	33,650
Kane	11,215	Wayne	17,642
Millard	151,319	Weber	44,690
Morgan	16,527		

What then is the function of a map and particularly a statistical map? Simply, it is to provide the reader with a graphic display of information in such format that it promotes a conceptual relationship with spatial arrangements on the surface of the earth. The information communicated by a map may be the distance or directional relationships among things, the pattern of objects in space, or the shape and location of a region such as the Corn Belt. Very few readers of the information presented in Table 1 are able to perceive the geographical patterns and spa-

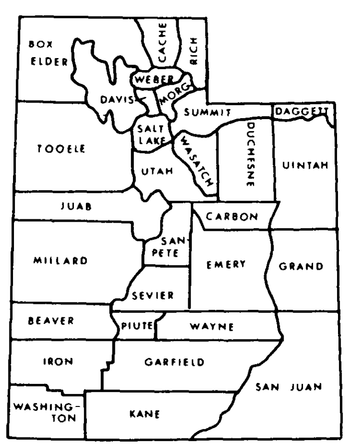

Figure 1: This base map of Utah provides the geographic network upon which a variety of statistical maps of the State can be formulated.

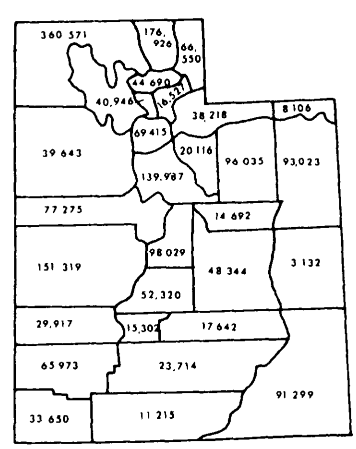

Figure 2: An areal table differs from a map in that it does not provide the reader with a graphic visualization of a distribution. Most readers are unable to integrate the numerical values on this table into a meaningful spatial pattern.

tial relationships of cropland acreages which are so clearly set forth on the map in Figure 3.

Contemporary map makers often confuse the primary purpose of a statistical map with that of a table. As a result they attempt to provide the reader with specific facts about specific places. Usually these attempts are no more successful than those achieved by inventors who have tried to create a hybrid between the automobile and the aircraft. There are two dominant reasons why authors ought not to confuse cartographic and tabular utility. The first, and most obvious, is that statistical maps are symbolized generalizations of the information contained in a table. Statistical maps tend, therefore, to be inefficient and inaccurate sources of data. For example, it is not easy to count the dots on Figure 3. Furthermore, each dot on Figure 3 represents five thousand acres, and in processing the data into symbols the map maker was obliged to round off the specific values given in the table. This procedure may create errors of as much as two thousand and forty-nine acres in some parts of the map.

A second, and geographically more sinister error is involved in attempting to provide the dual function in a statistical map and is illustrated in Figure 4. Most readers are thought to be unfamiliar with the geographic base information of an area and thus the cartographer feels that he must provide boundaries and names to assist his reader in obtaining data from the map. Thus, a portion of the map surface is occupied with lines and words prior to symbolization. This generally results in spatial patterns of symbols which are not properly related to earth position. What you really see on the map in Figure 4 is a combination of the patterns created by past historical-political decisions and a spatially erroneous display of cropland acreage.

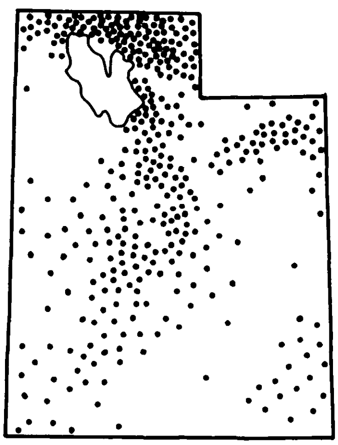

Figure 3: A dot map of the enumerated data presented in Figures 1 and 2. Each dot represents 5,000 acres of cropland and is placed in the approximate position of where that amount of cropland exists in the State of Utah.

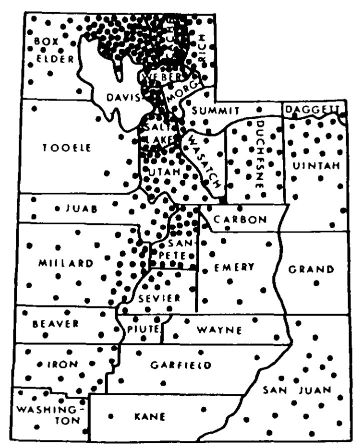

Figure 4: A dual purpose dot map of cropland acreage in Utah. This map provides the reader with rough acreage figures if he elects to count the dots but the spatial pattern of the distribution is confusing because the county names and boundaries interfere with the more accurate areal patterns shown on Figure 3.

Statistical maps have also been used for purposes other than those just described. If one reads the textual material associated with statistical maps, it becomes apparent that some authors use maps only for decorative purposes since no reference is made to the significance of the map message either in the caption or the exposition. Other authors seem to use maps as badges of membership in the fraternity of scientists and these individuals also ignore the map as a communicative device. If an author or a map maker is not primarily concerned with those unique spatial relationships which can best be illustrated in map format, some device other than a map ought to be utilized to communicate his concerns to his reader.4/

THE SYMBOLIC LANGUAGE OF MAPPING

The statistical maps found in contemporary literature are the products of a heterogeneous lot of individuals who are grouped together because of the function that they perform rather than because of the training that they have received. In one way, this diversity of academic experience and methodological outlook is advantageous because new informational concepts and types of statistical maps are produced. In another way it is disadvantageous, however, since many of these cartographers do not know the symbolic language of their trade. A review of the accepted practices in the use of map symbols and a common error in presentation are discussed in the following paragraphs.

In general, there are five types of symbols which are used to present information on statistical maps and each symbolic form has been assigned a particular function in thematic convention. Point symbols, e.g. dots, are used to represent phenomena which occur singly at points on the surface of the earth. Typically one finds population, either human or animal, and crop acreages represented by point symbols. Proportional symbols, e.g. circles, are used to present clusters of things which occur in very limited areas. The population of cities or the value added by manufacturing may be symbolized in this way. Area symbols are used to represent spatial phenomena which may be considered to be evenly distributed over the area. Due to this assumption, it is common practice to represent ratios with shading patterns. Lastly, we find certain phenomenon which occur ubiquituously on, above, or below the surface of the earth. Air pressure or elevation are distributions of this type and their configurations are represented by isarithms of which the contour is most common. Each of these symbolic types can be varied in color, size, shape and texture so that a very diversified array of marks is available to the cartographer. 5/

While there are no absolute rules for symbolic conventions, the statistical cartographer is well advised to practice recommended usage. Each mappable phenomenon has distinctive geographical characteristics and an understanding of these is the basis upon which symbolic conventions were established. It is also clear that many map users expect certain types of phenomenon to be depicted in a certain manner and if the symbolic language of the map coincides with the expectations of the reader the impact of the map message occurs more rapidly and with fewer aberrations than would be the case otherwise.

The two contrasting maps shown in Figure 5 illustrate the significance of failure to follow symbolic convention. The two maps communicate very different spatial messages but both were derived from the same cropland data used for the dot maps of

Figures 3 and 4. At least part of what is seen on map A results from the simple fact that large enumeration districts tend to have large numbers of things and small districts tend to have small numbers of things. If one flew over the State of Utah and observed cropland patterns he would perceive regions with a large proportion of cropland as distinct from regions with widely dispersed field patterns. His impression of the distribution of cropland would, therefore, be more like that shown on map B than on map A of Figure 5. It then follows that acreage and similar types of data should be areally standardized by the cartographer in order to present a truthful visual statement on his choropleth map about those distributions, the enumerations of which are directly related to the size of the enumeration units. 7/

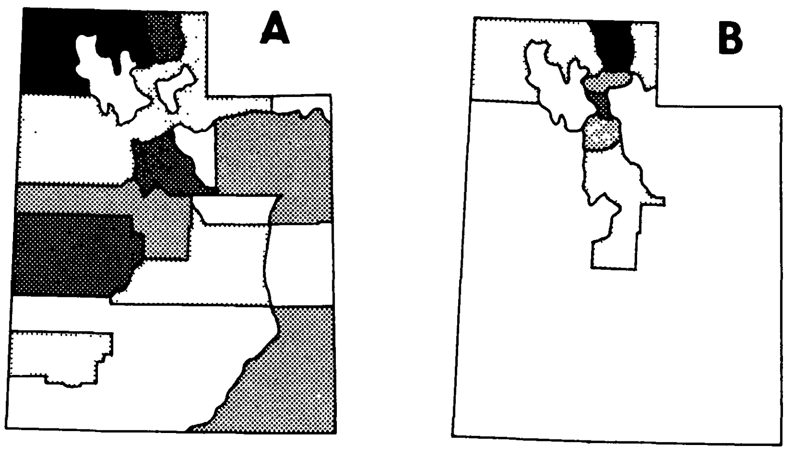

Figure 5: These two very different maps were created from the same data set, Table 1. Map A is a five-class representation of the raw enumerated data. The dark gray county in the middle of the western side appears to be more important, because of its size, than the dark gray county along the northern border. In Map B data were areally standardized before they were symbolized, and the visual importance of the two counties mentioned above has been reversed. The patterns on Map B more closely represent reality than those on Map A.

MAP ACCURACY AND DATA PROCESSING

Cartographers in survey mapping establishments take great pride in the high level of accuracy that they incorporate into their maps. Anyone who has visited a topographic mapping organization has no doubt been made aware of the care with which the contour and cultural data were measured, symbolized and plotted on the topographic sheets. This is as it should be, but this concern for positional and elevational accuracy has caused some thematic map makers (as well as the general public) to believe that their maps are equally accurate. This is an unfortunate circumstance because some statistical map makers, having taken great care in the arithmetic manipulation of their data, fail to realize that other types of significant error may creep into their map products.

In 1971 Joel Morrison introduced the term, "method produced error". 8/ This term was used to distinguish decision-making errors from arithmetic errors in the cartographic process. The two maps presented in Figure 6, although both arithmetically accurate and symbolically correct are quite different in appearance. This difference in map message is not related to the number, orientation or values (tones) of the symbols on the map nor is it related to the scale at which the maps were prepared. Instead, the difference in message is the direct result of the method used to subdivide the twenty-nine values (percent of land in cropland) into four classes. Quartile classes were used on map A while the classes on map B were derived by an optimizing technique. How can two arithmetically and computationally accurate maps produce such different distributional patterns?

Cartographers often face a choice like that posed by the two maps in Figure 6, and of necessity, they must select one to place before their readers. Too often the selection is made intuitively, but if one turns to traditional statistics, the problem can be resolved more rationally. Ideally, the cartographer should select the representation which minimizes the differences within classes and maximizes the differences between classes on his map, and this can be measured by analysis of variance. Application of this measure to the maps in question reveals that the within-class variance of map A is 2,524 while that of map B is 940. These values, when compared with the total variance of the data set, indicate that map B is ninety-four percent accurate while map A is only eighty-three percent accurate. With this information there is no problem in deciding which map should be used.

Statistical maps are considered to be poor data retrieval sources but many census map makers believe that the average map reader will use a map as a data bank. If their assumption is correct, and it has not been proven otherwise, the choice of classification procedure becomes critical. The standard error of the estimates for the maps in Figure 6 indicate that on the average a value obtained from map A will be in error 9.3 percent while a value obtained from map B will be in error 5.7 percent.

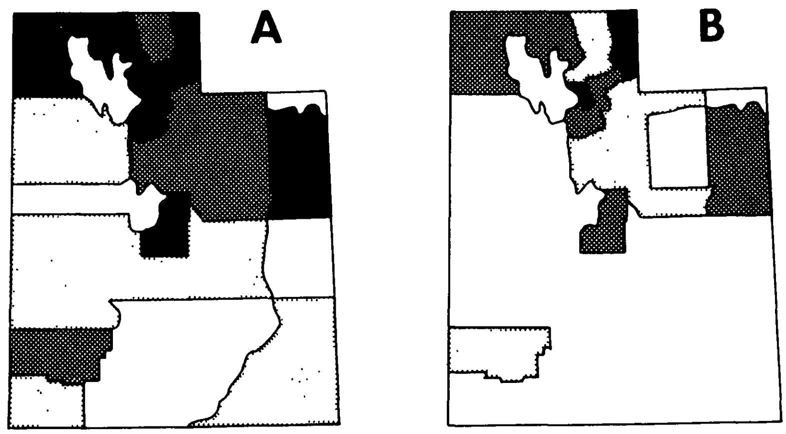

Figure 6: Two maps are different versions of one distribution, the percent of area in cropland in Utah. The classes for Map A are quartiles while those for Map B were created by an optimization technique. Eighty-three percent of the total variance is accounted for by the classes in Map A and ninety-four percent by those in Map B. Thus, Map B is the more accurate map.

The message of a map is communicated to the reader by a process of visual integration in which the characteristics of individual symbols are subjugated to the perceived patterns of sets of symbols. Recognition of this map reading procedure is essential in cartographic design because size, color, texture of symbols must be considered on two bases. In the first instance each symbol must be identifiable. In addition, each symbol must be compatible with all other symbols since they must work together to create a visual whole. Failure to recognize symbolic identifiability and compatibility results in map messages which are complex and "noisy" rather than simple and clear.

Several graphic principles are of concern to the cartographic designer. Figure and ground must be separated clearly so that the major symbols which present the distribution can be seen as distinctly different from those lesser symbols which may be necessary to the purpose of the map. Poor figure-ground relationships are demonstrated on the map presented in Figure 4 where, in addition to areal misplacement, the dots become visually confused with the boundary lines and the county names. A second aspect of figure-ground relationships is shown in Figure 7 where lack of contrast on map A subdues the map message which becomes clear and distinct on map B.

Many of the most confusing and displeasing maps to be found in contemporary literature are those which are symbolized with self-adhesive shading patterns. These patterns are available in a variety of styles, textures and values and they can be applied in all of the orientations of the compass. The patterns used on map A of Figure 8 are so coarsely textured that they are visually unstable. Compare this map with those shown in Figure 6 and there is little doubt as to which are easier to read and more pleasant to look at.

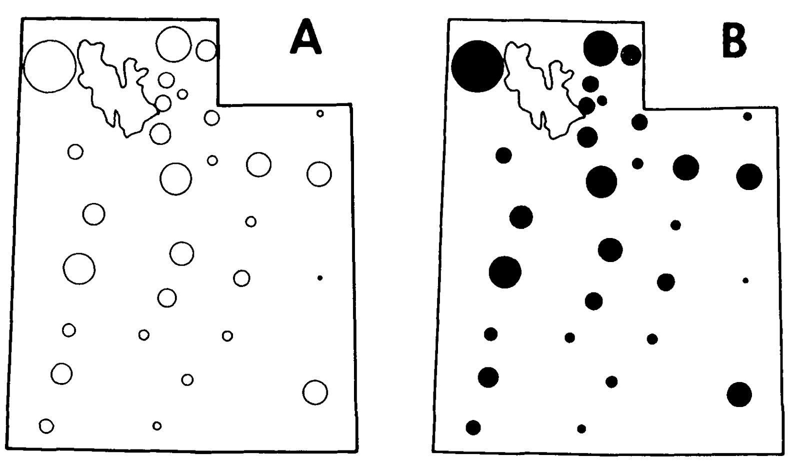

Figure 7: Minimal (Map A) and maximal (Map B) contrast between figure and ground. Map B is a clearer and more forceful visual statement than Map A.

At first glance, the reader of map B in Figure 8 might conclude that he is looking at a two-class map but in reality it is a four-class map. A single pattern is used for classes one and two and another for classes three and four and the only visual difference between these pairs of classes results from pattern orientation. This makes the task of the map reader more difficult than would be the case were the patterns also separated by visual differences in value. The intersection of the lines in these rather coarse patterns with the boundary lines of the regions causes a degree of noise on this map that could be reduced by decreasing texture.

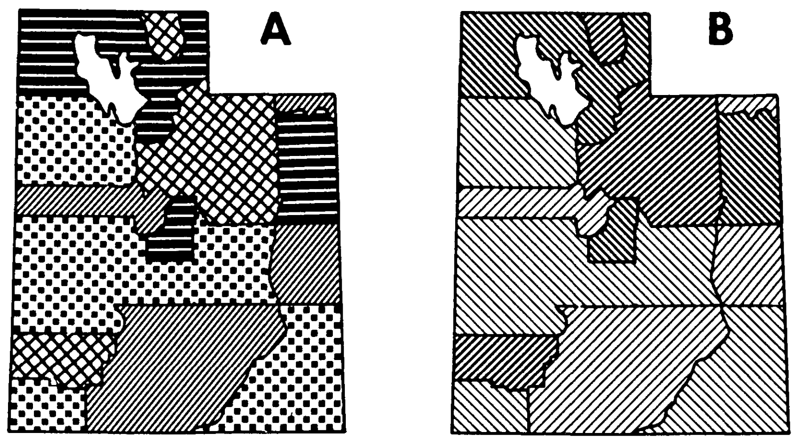

Figure 8: Two aesthetically unpleasant and visually unstable choropleth maps. The pattern is so coarse that they "dance" before the reader's eyes.

To paraphrase a current television commercial, statistical cartographers should prepare maps which are in "good taste." Great works of art like the Mona Lisa or Venus de Milo are appreciated because their creators presented their subject matter with simple and clear graphic statements. Statistical map makers would do well to remember and emulate these artists when they design their maps.

A LOOK INTO THE FUTURE

Rudolph Arnheim in his book, Art and Visual Perception, states that gestalt psychologists believe that "any stimulus pattern tends to be seen in such a way that the resulting structure is as simple as the given conditions permit." 9/ If this concept is applied to the perception of maps, one can assume that complex patterns of symbols are perceived as simplified or generalized patterns. Carrying the idea further, the simplified patterns that are perceived must be those that are remembered and later utilized when the reader wishes to recall the informational content of the map. These notions raise a very important question about map communication. Do all (or a majority of) map readers visually simplify a pattern of symbols in the same way and thus obtain and remember the same map message?

Current research in map communication seems to indicate that messages obtained from maps may have an individualistic and highly personalized informational content.

As a case in point, study the three maps presented in Figure 9. 10/ The center map (B) was shown to a group of individuals who a short time later reconstructed what they had seen and remembered. The response maps on the left and right (A and C) indicate that the readers obtained both an aberrant and a different spatial message. Research cartographers are unable to explain why these response maps are different, nor are they sure whether these responses represent extreme cases of map reader diversity. This example makes one wonder whether research cartographers can depend upon map reader generalization. Perhaps the map designer should take the initiative and create the simplifications he wishes to communicate.

Traditional map makers may think that such manipulation of the map message is an abhorrent misuse of cartographic prerogatives and akin to preparing propaganda. Upon further contemplation, however, we realize that speakers and authors select words, create phrases, and structure paragraphs to suit the message that they are attempting to communicate. Should cartographers be denied the same privilege?

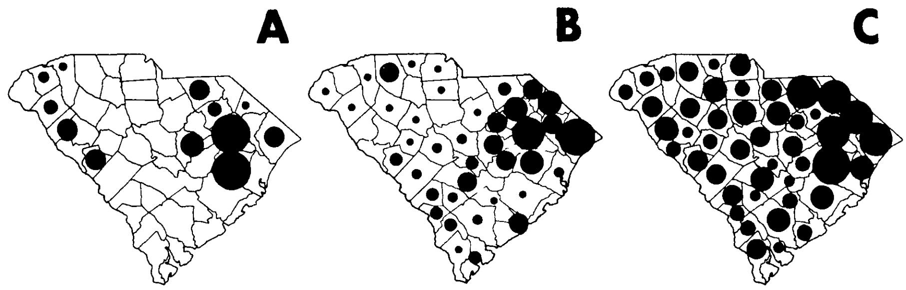

Figure 9: Map B was shown to a group of freshman college students and later these students were asked to reconstruct from memory what they had seen. Maps A and C are two of these memory reconstructions and they are both very different from the original (Map B) and from each other. This experiment raises serious questions about the efficiency of maps as tools for the communication of spatial information.

RESUME

A number of rather pernicious impressions about statistical maps and the cartographers who create them have crept into the minds of the public. It has been the objective of this paper to help dispel these notions and to provide alternative ways of looking at maps and their creators. The following statements summarize these attempts.

- Statistical maps are not good sources of specific information about specific places.
- Maps are created to provide information about spatial relationships. No other medium communicates distance, directional and areal pattern relationships as well.
- There is a symbolic language of mapping and good maps can be created only by those who understand it.
- Accuracy in cartography involves conceptual as well as arithmetic

precision in data handling.

- Maps are graphic communicators and clear and concise map messages result from easily read, clear rather than noisy maps.
- Communication via the map is not perfectly understood and there is need for continued research if we are to create more effective maps.

REFERENCES

1. Petchenik, Barbara B., "A Verbal Approach to Characterizing the Look of Maps," The American Cartographer, vol 1, No. 1 (1974), pp. 63-71.

2. Statistical maps are not necessarily products of statistical inference; rather many statistical maps are two-dimensional graphic representations of enumerated data obtained by censual procedures.

   This paper is an abbreviated version of an oral paper presented at the International Symposium on Computer-Assisted Cartography (Auto-Carto II), sponsored by the American Congress on Surveying and Mapping in cooperation with the Bureau of the Census.

3. The data presented in Table 1 and that used for other maps of Utah were taken from Utah Agricultural Statistics 1975, Salt Lake City:Utah State Department of Agriculture, pp. 86-7. Map makers frequently array these types of data in sequential order as a preliminary to processing the figures into categories for symbolization.

4. For a more sophisticated analysis of maps as communicative devices see: Robinson, Arthur H., and Barbara B. Petchenik, "The Map as a Communicative System," The Cartographic Journal, vol. 12, No. 1 (June 1975), pp. 7-15.

5. Also see Robinson, Arthur H., and Randall D. Sale, Elements of Cartography, 3rd edition, New York: John Wiley and Sons, 1969.

6. Dobson, Michael W., "Symbol-Subject Matter Relationships in Thematic Cartography," The Canadian Cartographer, vol. 12, No. 1 (June 1975), pp. 52-67.

7. Areal standardization of data for choropleth maps is achieved by converting simple enumerated data to ratios. In the case of map B in Figure 5, the cropland area of each county was divided by its total land area to derive the ratio, cropland as a percentage of total area.

8. Morrison, Joel, Method Produced Error in Isarithmic Mapping, Cartography Division Technical Monograph No. CA-5, Washington, D.C. : American Congress on Surveying and Mapping, 1971.

9. Arnheim, Rudolph, Art and Visual Perception, Berkeley: University of California Press, 1971, p. 44.

10. These maps are taken from the unpublished Ph.D. dissertation research of Theodore Steinke at the University of Kansas.