Beyond Stevens: A revised approach to measurement for geographic information

Nicholas R. Chrisman
CHRISMAN@u.washington.edu
Department of Geography DP 10, University of Washington
Seattle, Washington 98195 USA

ABSTRACT

Measurement is commonly divided into nominal, ordinal, interval and ratio 'scales' in both geography and cartography. These scales have been accepted unquestioned from research in psychology that had a particular scientific agenda. These four scales do not cover all the kinds of measurements common in a geographic information system. The idea of a simple list of measurement scales may not serve the purpose of prescribing appropriate techniques. Informed use of tools does not depend on the nature of the numbers, but of the whole 'measurement framework', the system of objects, relationships and axioms implied by a given system of representation.

Introduction

The approach to measurement in certain social sciences is still strongly influenced by Stevens' (1946) paper in *Science*. His 'scales of measurement' form the basis for geography (Unwin 1981) and for cartography (Muehrcke 1976; Chang 1978). While measurement has continued to develop in social science research (Churchman and Ratoosh 1959; Coombs 1964; Ellis 1966; Krantz et al. 1971; Narens 1985; Suppes et al. 1989; 1990), these continuing developments have not been followed in the cartography and GIS literature.

Development of theories of measurement
The 'classical' school of measurement developed in physics and other sciences by the end of the nineteenth century. In the classical view, measurement discovered the numerical relationship between a standard object and the one measured. The property was seen as inherent in the object. This viewpoint is deeply ingrained in our language and society.

Let us take the attribute 'length'. Every entity in space can be measured by comparing its length to some other length. If we adopt a 'standard' measuring rod, we can obtain numbers (the ratio between the length of the rod and the objects measured) by a physical procedure that mimics addition – laying the rod successively along the edge. Nineteenth century physics was able to build up a rather complex model of the world with remarkably few of these fundamental properties (length, mass, electrical charge, etc.). These properties were termed 'extensive' because they extended in some way as length does in space. Other properties (like density) were built up as ratios of the extensive properties and were thus 'derived'. The laws of physics prescribed the rules for derived measures.

Extensive properties are rather restrictive, and the idea of a universal standard measuring rod in Sèvres, France is not very practical for all the properties that must be measured. Physicists began to move beyond the classical concept that the meter was an intrinsic property of one particular rod. The method of measurement became just as important as the physical standard, thus separating the object and the measurement. A twentieth-century philosophy of measurement called 'representationalism' saw numbers, not as properties inherent in an object, but as the result of relationships between measurement operations and the object.

Exclusive focus on extensive measurement in physics left almost no room for the social sciences to develop a measurement theory. The physicists could not consider phenomena like perceived loudness of sounds as a measurement, since it did not involve extensive properties like addition. Stevens' system arises from this context, a part of the movement to create a quantitative social **science**.

This paper will begin with a quick review of Stevens' scheme, followed with some examples of measurements in one and more dimension which require another approach. At one end, the 'ratio' level is not the highest level, nor is it so unified. At the other, the nature of categories need to be reexamined. Stevens' hierarchy also fails to treat the circumstances of multidimensional measurement.

Stevens' Scales of Measurement

Table 1: verbatim copy of Stevens' (1946, 678) Table 1

Scale	Basic Empirical Operations	Mathematical Group Structure	Permissible statistics (invariantive)
NOMINAL	Determination of equality	*Permutation group* $x' = f(x)$ $f(x)$ means any one-to-one substitution	Number of cases Mode
ORDINAL	Determination of greater or less	*Isotonic group* $x' = f(x)$ $f(x)$ means any monotonic increasing function	Median Percentiles
INTERVAL	Determination of equality of intervals or differences	*General linear group* $x' = ax + b$	Mean Standard deviation Rank-order correlation Product-moment correlation
RATIO	Determination of equality of ratios	*Similarity group* $x' = ax$	Coefficient of variation

Stevens adopted the representationalist philosophy in a 'nominalist' form (Michell 1993), defining measurement as the 'assignment of numbers to objects according to a rule'. Table 1 reproduces Stevens' original table exactly so that his presentation is not clouded by the reinterpretations developed over the past fifty years.

The scales are defined by groups of mathematical operations that were increasingly restrictive. The focus was upon 'groups' of transformations under which the meaning of the scale remains invariant. A nominal scale could be replaced by any other scale that could be mapped one-to-one onto the original one. One subset of those operations are 'isotonic' [meaning monotonic]; a subset of these are linear, and a subset of these are simply multiplicative. From the start, the levels of measurement can be associated with an attempt to bring mathematical order into fields that do not seem to be as rigorous as physics, which had controlled the earlier developments of measurement theory (Campbell 1920; Bridgeman 1927; Michell 1993).

A key element of Table 1 is the connection between the 'scales' and 'permissible statistics'. Many textbooks on statistics for social sciences (beginning with Siegel's (1956) classic on non-parametric methods) adopted this connection between a variable and appropriate techniques. At the root, measurement is seen as a choice to represent an entity by a number, relationships were simplified to those inherent in the number system chosen. The association between numbers and methods may not be as simple as Stevens and Siegel conceived, particularly when dealing with geographic information.

Stevens tried to expunge the distinction that physicists had drawn between 'extensive' and 'derived' measurement. In Stevens' reductionist viewpoint, the properties applied to the number system, not the method by which it was generated [this had been the viewpoint of the operationists like Percy Bridgeman (1927)]. It is ironic that cartographers teach Stevens' system, with a unified 'ratio' level, then must make a distinction between those attributes permissible on choropleth maps (densities and other derived measures) and those permissible on proportional symbol maps ('extensive' measures where addition is the underlying mechanism). Using Stevens for cartography has been established for years (Muehrcke 1976; Chang 1978), and the inadequacies do not seem to be recognized.

Above Ratio

Stevens' four 'scales' are usually presented as a complete set, but they are far from exhaustive. Stevens (1959) himself proposed another scale at the same level as interval for logarithmically scaled measures. The invariance is the exponent, while the zero is fixed. This 'logarithmic-interval' scale is not cited in any of the geographic literature, though it is used for earthquake intensities and similar measurements. Following Stevens' invariance scheme to its conclusion, ratio is not the highest level of measurement. The ratio scale has one fixed point (zero) and the choice of the value of 'one' is essentially arbitrary. A higher level of measurement can be obtained if the value of one is fixed as well. Then the whole scale is predetermined or *'absolute'* (Ellis 1966) and no transformations can be made that preserve the meaning of the measurement. One example of an absolute scale is probability, where the axioms fix the meaning of zero and one simultaneously. Bayes' Law of conditional probability works because the scale is fixed between zero and one. Probability is just one example of a scale not recognized by Stevens.

Another class of geographic measurements consist of counts aggregated over some region in space. Counts are discrete, since there is no half person to count, but a count captures more mathematical structure than the other discrete levels (nominal and ordinal). Since the zero is a fixed value, counts may seem ratios, but, being tied to the discrete unit counted, it cannot be rescaled. Counts have different properties from the absolute scale, as well. Ellis (1966, p. 157) points out the difference between ratio scales and counting with the example that it is acceptable to posit a unit by saying "Let this object be 1 minch long", but it is not possible to say "Let this group contain one apple", since it either has one apple or some other number when you start. As I will demonstrate below, the process of counting depends upon the recognition of objects, a procedure tied to nominal measures.

Cyclical measures

While Stevens' levels deal with an unbounded number line, there are many measures which are bounded within a range and repeat in some cyclical manner. Angles seem to be ratio, in the sense that there is a zero and an arbitrary unit (degrees, grads or radians). However, angles repeat the cycle. The direction 359° is as far from 0° as 1° is. Any general measurement scheme needs to recognize the existence of non-linear systems. Some aspects of time, have repeating or cyclical elements. In environmental studies of all kinds, the seasons play an important role. Stevens' scheme does not allow for measurements that can be ordered spring–summer–fall–winter–spring or fall–winterspring–summer–fall. The seasonal relationships are invariant to the starting point in the cycle.

Spatial measurement raises questions about measurement scales. In the one-dimensional world of Stevens, the open-ended ratio scale seems to provide the most information content. A real number line contains the most promise for mathematical relationships. When representing a two-dimensional space, the normal scheme, attributed to Descartes, uses two orthogonal number lines. Analytical geometry can demonstrate the conversion between coordinates on two orthogonal axes and a radial system (Figure 1). These two representations are equivalent even though the units of measurement do not seem equivalent. The reason is that the two orthogonal distances create a triangle. The radial coordinates specify that same triangle using the hypotenuse and an angle. The theorems of geometry demonstrate that the two triangles are congruent, a finding that would not be apparent from their measurement scales.

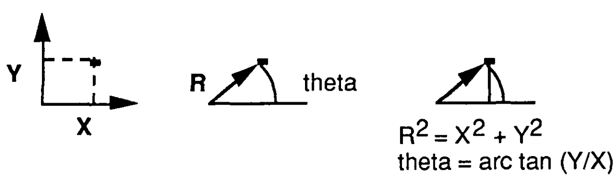

Figure 1: Cartesian axes convert to radial reference without loss

The conversion from two ratio scales to one ratio scale plus an angle is not unique to geometric constructions. Potentially infinite vectors can be simplified into lower dimensional renditions, adding great complexity to the intuitive structure propounded by Stevens. For example, all the gravitational forces from various directions can be resolved into a resultant force measured in three space. Similarly, radiant energy at various wavelengths coalesce into a particular color that can be represented in a simple conical object (Munsell's space or equivalent). The color cone is thus a "fact of nature, not a mathematical trick" (Suppes et al. 1989, p. 226). Multidimensional measurements create interactions not imagined in the simple linear world of Stevens. Since GIS is inherently multidimensional, the linear model limits our understanding concerning the interactions of measurements.

If there is any theory to GIS, it would have to start from the storage of attribute values in their spatial context. Tomlin (1983; 1990) has built a complex range of tools around the raster model of values stored for an array of point/areas. Goodchild (1987) contrasts the object view (isolated objects in a void) and the 'field' view (a z value for all pairs x,y). This commonality of thinking is strongly influenced by the storage systems that we have invented. We must remember that the slope of a surface is characterized by two numbers [gradient and aspect to use the terminology of Burrough (1986)]. We lose much understanding by the reductionism that treats these as arrays of numbers, not the vector space that the two numbers taken together portray. GIS is still stuck with scalar values as the basic conception, while vector fields and tensor fields are necessary to connect the representations to process. Higher numbers of dimensions require more complex spatial data structures (Pigot and Hazelton 1992; Worboys 1992, for example).

Rethinking nominal measurement

While Stevens' top end, the ratio scale, leads off in the direction of multidimensional measures, the bottom end is equally problematical. The nominal scale is not even considered to be a kind of measurement in many theoretical discussions (Ellis 1966; Krantz et al. 1971; Narens 1985). Social scientists had much discussion about identifying numbers, such a 'football numbers' (Lord 1953). A strictly arbitrary string assigned to each object is not really a category that groups together any individuals. Thus, it does not support Stevens' 'equality' operator. Furthermore, most numbering systems (like Lord's football team numbers) provide some kind of ordinal information about the sequence in which they are assigned or some other logic internal to the authority responsible. Identifying numbers are not really the categories that concern this discussion. Basically, a nominal category defies the logic expected of a 'scale'; order, systems of inequalities and some concatenation operations (Krantz, Luce et al. 1971, p. 4). These are the ingredients of ordinal measurement or higher.

Does this mean that nominal measurement must be abandoned? In a careful reading of measurement theory, the tide has changed from Stevens' simplification. For Stevens, the numbers determined the nature of the methods. Even some theorists who ignore nominal measures provide a basic definition that leads in another direction. Volume 1 of *Foundations of Measurement* (Krantz et al. 1971, p. 9) defines a scale as a construct of "homeomorphisms from empirical relational structures of interest into numerical relational structures that are useful". The key issue is not the invariance of some algebraic properties, but the invariance of the underlying *relationships*. Though restricted to numbers, this definition can be broadened to deal with categories. If the measurement preserves the empirical relationships and provides a useful structure for analysis, a nominal categorization fits the general requirements for a scale of measurement.

The trouble has been the oversimplification of nominal distinctions. In most treatment, Aristotle's rules are applied. Each member of a set must share common characteristics. Stevens adopts this rule by requiring that all members of a nominal group are equivalent. Certainly there is plenty of precedent for these kinds of rigid categories, but representations of the world do not always fit the simplicity of this logic. Many scientific categories, and even more of the categories of every-day life, do not live up to the purity of 'shared attribute' categories. Modern category theory (Johnson 1987; Lakoff 1987) describes at least two other alternatives; probabilistic and prototypes. While classical set theory assigns an object either as a full member or not in each category, a probability approach provides for a gradation of membership. The purest application of probability states a likelihood that an object will be discovered to belong in the classical sense. This is the approach taken to interpret soils classes by Fisher (1991). Goodchild (1992) suggests a gradation of membership that moves from the strict interpretation of probability towards a fuzzy set membership interpretation. Taken strictly, fuzzy memberships do not have to sum to one, though this normalization is often implied (Burrough 1989). This is a fracture zone for cartographers. Partial membership is often implied, and the specific model, whether probability or fuzzy sets, is rarely articulated.

On many occasions that cartographers refer to fuzzy sets or probability of membership, they really are using a 'prototype' approach to categories. The prototype refers to a 'central' example that represents the ideal form of the category. Objects are not matched attribute by attribute, but assigned to the prototype that fits most closely. There is some measure of 'distance' involved that may be mistaken for probability. The difference is that probability normalizes the separation so that every object sums to one. Distances from a prototype do not have to sum to any particular value. Some objects are just closer than others. Classification in remote sensing usually uses prototypes and distance based analysis internally before sending a sharp set out for final consumption. Supervised classification establishes the prototypes directly, then assigns each pixel to the 'closest' using the distance in spectral space. An unsupervised classification looks for the smallest set of clusters that will partition the spectral space, but pixels are also seen as more or less central to the cluster. The key trick, as always, is assigning the category names to the clusters; a process that often involves a complex interpretation of the spatial context. Lakoff and Johnson point out that the human mind tends to use prototype logic, rather than the rigid formalism of classical categories. The nature of human cognition is not the issue here, but the question of which relationships must be modeled to make the categories represent the scientific intentions.

Probability and prototype approaches to categories may dominate the real applications of geographic information, though classical categories pervade the explanation. There is a lot of literature talking about the inflexibility of categories, as if all geographic categories involve exact matches to a list of defining characteristics. Actual practice is far different. Categories are conceived in taxonomies, as a comprehensive system. All land is presumed to fit in a category, even if it is 'Not Elsewhere Classified' – a category that certainly does not share attributes amongst its members. The landscape is assigned to the closest fit category, or maybe to the most likely category. To return to measurement theory, geographers should remember that categories are not used to share formal properties along the Aristotelian scheme, but to partition a space into a nearest grouping. Geographic categories are developed to generalize.

Stevens' four scales of measurement are not the end of the story. The concept of a closed list of 'scales' arranged on a progression from simple to more complex does not cover the diversity of geographic measurement. Still, Stevens' terminology is so deeply entrenched that it may remain in use when it applies.

A larger framework for measurement

The largest difficulties with Stevens' scheme come not from the specific 'scales' of measurement, but with the overall model of the process. The levels of measurement presume a rather simple framework; the classical social science 'case' 'has' attributes. Such a model was proposed for most social sciences in the early quantitative period. The version proposed in geography was called the 'Geographical Matrix' (Berry 1964), simply a matrix with 'places' on one axis and attributes on the other. But all 'cases' or places do not have the same attributes. A more fruitful model sees measurement not in terms of properties, but in terms of relationships. Geographic information involves many more kinds of measurement. These distinctions have usually been discussed as 'data models', with an emphasis on representation. Viewed from the perspective of measurement, these old issues take on a new clarity.

This paper proposes a scheme of measurement frameworks developed from the simple taxonomy presented by Sinton (1978). Each model or framework for geographic measurement must account for each of these elements interacting in the roles of fixed, controlled and measured. In Sinton's scheme, in order to measure one component, one of the others had to be 'fixed' and one served as 'control'. At the most basic, Sinton's scheme distinguishes vector from raster because the first controls by object (attribute), while the later controls by space. This rough division provides a starting point, but it does not explain the divisions within these two approaches.

A Taxonomy for measurement: Object as control

When the attribute serves as the control, the spatial location is the subject of the measurement. While this is the common framework for a vector representation, there are large differences between the situation with an isolated category and a connected system of categories. Table 2 summarizes the distinctions.

Table 2: **Object Control Frameworks**

Isolated Objects	
Spatial Object	Single category distinguishes from void
Isoline	Regular slices of continuous variable
Connected Objects	
Network	Spatial objects connect to each other, form topology (one category possible)
Categorical Coverage	Network formed by exhaustive classification (multiple categories, forming an exhaustive set)

The simplest object control framework involves isolated objects, distinguished by a single category. While 'cartographic feature' might be apt, this framework will be termed 'spatial object'. Each point or area object is described as a geometric whole, since it will forcibly occur in isolation. The message of the object framework is: 'Here is an airport'; 'Here is another airport.' and so on. In the pure form of this framework, the only relationship is between the object and a position; there are no relationships between objects. Linear objects depart from this to some extent, creating the need for the network framework discussed below.

Isolines are formed by controlling for a specific value on a surface. Since isolines follow the contours and do not intersect, they have no topological relationships, beyond the ordering of nested contours.

In the creation of advanced GIS software, it was important to recognize that there were relationships between the objects in a database. When a coverage is formed with multiple categories, there will be topological relationships. Similar structure can be created by linear networks. The basic topology is required whether the categories form strict equivalence classes or some form of probabilistic or prototype categories. The distinction between the isolated coverages and connected coverages is not a matter of database design, but a recognition of the underlying measurement structure of the source material.

Spatial Control

Control can also come from a set of predefined spatial objects (Table 3).

Table 3: **Spatial Control Frameworks**

Point-based Control	
Center point	Systematic sampling in regular grid
Systematic unaligned	Random point chosen within cell
Area-based Control	
Extreme value	Maximum (or minimum) of values in cell
Total	Sum of quantities (eg. reflected light) in cell
Predominant type	Most common category in cell
Presence / absence	Binary result for single category
Percent cover	Amount of cell covered by single category
Precedence of types	Highest ranking category present in cell

Control by a set of points has different rules compared to control by areas. While both would be encoded in a raster representation, they must be understood differently. With a point-based control there are not too many rules. Center point provides a regular sampling of a landscape. Digital Elevation Matrices tend to use a point-based sample, though the photogrammetric equipment may actually work on a tiny area to match the photographs. Systematic unaligned is recognized in textbooks, but rarely performed.

Control by area is more common for remote sensing and other applications of grid sampling. In each cell there is some rule that has been applied to all the possible values. Some sensors add up all the reflectance in a certain band width; other gridding takes the highest or lowest value. A system that optimizes each cell by taking the most likely value for the cell may remove all traces of linear features and the minority elements. Unless these rules are known to the analyst, the information can be sorely misconstrued.

Other kinds of control
Control by object and control by space seem to be the only options, but they do not cover all the cases found in existing geographic information. The well-known *choropleth* map is an example of a **composite** framework, in that the base map is created using a categorical coverage for the set of collection units, then these objects serve as a secondary form of spatial control to tabulate the variable in question. Due to these two stages, the spatial measurements of the boundaries have little bearing on the precision of the measurement.

Triangular Irregular Networks (TIN) do not fit the scheme either. While the points may come from an isolated bunch of measurements, the TIN represents a set of relationships that cover space. The ideal TIN is constructed so that the triangles represent zones of uniform slope and aspect, within the resolution available. Thus, a TIN represents a novel class of measurement frameworks where relationships form the control, not the values of the attribute or the location.

Conclusions

The list of measurement scales developed by Stevens do not serve the purpose of providing a structure for geographic measurement. Any scheme to handle geographic measurement must deal with relationships between attribute and location, and eventually with time. A system of 'measurement frameworks' may provide a clearer focus on the design and implementation of geographic information systems. The frameworks proposed here place the measurement in the context of axioms and relationships to preserve.

Acknowledgements
The research reflected in this paper was supported by sabbatical leave from the University of Washington. Portions of this paper are extracts from a book manuscript currently under review by John Wiley.

References Cited

Berry, B. J. L. 1964: Approaches to Regional Analysis: A Synthesis. *Annals of the Association of American Geographers* **54**: 2-11.

Bridgeman, P. 1927: *The logic of modern physics.* Cambridge MA: Harvard Press.

Burrough, P. A. 1986: *Principles of Geographical Information Systems for Land Resource Assessment.* Oxford: Clarendon Press.

Burrough, P. A. 1989: Fuzzy mathematical methods for soil survey and land evaluation. *Journal of Soil Science* **40**: 477-492.

Campbell, N. R. 1920: *Physics: the elements.* Cambridge: Cambridge University Press.

Chang, K.-T. 1978: Measurement scales in cartography. *The American Cartographer* **5**: 57-64.

Churchman, C. W. and Ratoosh, P., eds. 1959: *Measurement: Definition and Theories.* New York: John Wiley.

Coombs, C. H. 1964: *A theory of data.* New York: John Wiley.

Ellis, B. 1966: *Basic Concepts of Measurement.* Cambridge: University Press.

Fisher, P. F. 1991: Modelling soil map-unit inclusions by Monte Carlo simulation. *International Journal of Geographical Information Systems* **5**(2): 193-208.

Goodchild, M. F. 1987: A spatial analytical perspective on geographical information systems. *International Journal of GIS* **1**(4): 327-334.

Goodchild, M. F., Guoqing, S. and Shiren, Y. 1992: Development and test of an error model for categorical data. *International Journal of Geographical Information Systems* **6**(2): 87-104.

Johnson, M. 1987: *The Body in the Mind: The Bodily Basis of Meaning, Imagination and Reasoning.* Chicago: University of Chicago Press.

Krantz, D. H., Luce, R. D., Suppes, P. and Tversky, A. 1971: *Foundations of Measurement: Volume 1: Additive and Polynomial Representations.* New York: Academic Press.

Lakoff, G. 1987: *Women, Fire and Dangerous Things,.* Chicago: University of Chicago Press.

Lord, F. M. 1953: On the statistical treatment of football numbers. *American Psychologist* **8**: 750-751.

Michell, J. 1993: The origins of the representational theory of measurement: Helmholtz, Hölder and Russell. *Studies in the History and Philosophy of Science* **24**: 185-206.

Muehrcke, P. C. 1976: Concepts of scaling from the map reader's point of view. *The American Cartographer* **3**: 123-141.

Narens, L. 1985: *Abstract measurement theory.* Cambridge MA: MIT Press.

Pigot, S. and Hazelton, B. 1992: The Fundamentals of a Topological Model for a Four-Dimensional GIS. *Proceedings of the 5th International Symposium on Spatial Data Handling* **2**: 580-591.

Siegel, S. 1956: *Nonparametric statistics.* New York: McGraw-Hill.

Sinton, D. F. 1978: The inherent structure of information as a constraint to analysis: Mapped thematic data as a case study. In *Harvard Papers on Geographic Information Systems*, ed. G. Dutton, p. 1-17. Reading MA: Addison Wesley.

Stevens, S. S. 1946: On the theory of scales of measurement. *Science* **103**: 677-680.

Stevens, S. S. 1959: Measurement, psychophysics and utility. In *Measurement: Definitions and Theories*, eds. C. W. Churchman and P. Ratoosh, p. 18-63. New York: Wiley.

Suppes, P., Krantz, D. H., Luce, R. D. and Tversky, A. 1989: *Foundations of Measurement: Volume 2.* New York: Academic Press.

Suppes, P., Krantz, D. H., Luce, R. D. and Tversky, A. 1990: *Foundations of Measurement: Volume 3.* New York: Academic Press.

Tomlin, C. D. 1983: Digital Cartographic Modeling Techniques in Environmental Planning. unpublished Ph.D., Yale University.

Tomlin, C. D. 1990: *Geographic Information Systems and Cartographic Modeling.* Englewood Cliffs NJ: Prentice Hall.

Unwin, D. 1981: *Introductory Spatial Analysis.* London: Methuen.

Worboys, M. F. 1992: A Model for Spatio-Temporal Information. *Proceedings of the 5th International Symposium on Spatial Data Handling* **2**: 602-611.