Toward a Perceptual Science of Multidimensional Data Visualization: Bertin and Beyond

Marc Green, Ph. D.

ERGO/GERO Human Factors Science
green@ergogero.com

Available from ERGO/GERO in Word 97 format.

ABSTRACT

A true science of data visualizations requires both a theory of perception and of computer graphics. However, visualization designers have paid relatively little attention to perceptual issues. In this article, I outline how knowledge of human visual perception and physiology can lead to more effective visualizations. Bertin's (1983) Image Theory, the only comprehensive perceptual theory in the visualization literature, will serve as the medium for the discussion. Experimental vision research grounds Image Theory in "first-principles" and suggests corrections, modifications and extensions. The resulting updated version of Image Theory can serve as a guide to visualization design.

1.0 Data Visualization Is A Joint Function of Computer Graphics and Perception

"The Purpose of computing is insight, not numbers."
--Richard Hamming (1962)

The power of computers to collect, store and manipulate numbers has increased dramatically since Hamming's pointed observation. Much of this increased power, however, is wasted because humans are poor at gaining insight from data presented in numerical form. As a result, visualization research takes on great significance, offering a promising technology for transforming an indigestible mass of numbers into a medium which humans can understand, interpret and explore.

The transformation from numbers to insight requires two stages. As shown in Figure 1, the first maps from numbers (data/processes) to images by means of some algorithmic technique. The second maps from images to insight by means of perception. A true science of visualization must incorporate both a formal theory of computer graphics and a theory of human perception. However, visualization research, like most areas of computer graphics, has focused largely on image formalisms and has ignored issues relating to visualization and perception (Greenberg, 1988; Blake, 1990; Pun and Blake, 1990). Papers published in proceedings of visualization conferences, computer graphics journals, etc. are typically written by computer scientists who describe new software for mapping data to images but who ignore the mapping from images to perception. They offer neither a coherent rationale for the creation of a particular visualization nor any empirical validation. They employ no explicit model of perception, relying instead on intuition and introspection. In sum, they are concerned only with questions of what could be done and ignore issues of what should be done. This is unfortunate, since empirical studies repeatedly find that design evaluations based on simple introspection and preference are poor predictors of user performance.

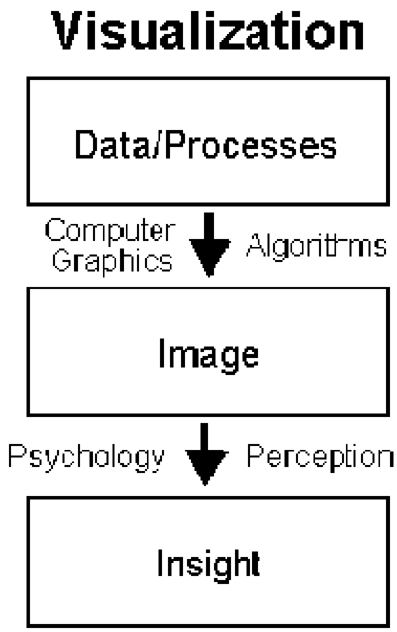

Figure 1

There is a small but steadily growing awareness (e. g., Levkowitz and Herrman, 1992; Merwin and Wickens, 1993; Rogowitz and Treinish, 1994) that perceptual issues are important for the design of good visualizations. Visualization conferences and journals contain increasing numbers of papers on topics such as specification of color scales, etc. In spite of this interest, however, there still seems little attempt to develop a general perceptual theory based on "first-principles" from human vision nor to use the vast body of existing perceptual data to guide design.

This article introduces visualization designers to relevant "first-principles" perceptual theory and research. An appreciation of basic perception, visual psychophysics and physiology can provide first-principles for a science of visualization, and suggest ways to develop new visualizations and to improve old ones. It can help move visualization from its current trial-and-error methodology to a true science.

To provide a bridge between the worlds of visualization and of vision research, I'll focus on Bertin's (1983) "Image Theory." It provides a particularly good interface, both because it is the only coherent perceptual theory in the visualization literature and because it closely parallels recent theorizing in human vision.

The first section briefly outlines the general problem of multidimensional data perception in order to introduce key conceptual issues which will reappear later. The subsequent section reviews relevant aspects of Bertin's theory and introduces the key concept of the *image*[1], the fundamental perceptual unit of a visualization. The next section describes current theories in visual perception, psychophysics and physiology, demonstrating the close parallels between Bertin's conceptual framework and current ideas in vision research. The discussion is divided into two parts, each covering a different area of vision research related to *Image* Theory. The first outlines relevant basic research in visual search/segmentation and demonstrates the great similarities between these tasks and data visualization. I show that the distinction between "preattentive" and "attentive" visual processing, so central to the search/segmentation research, is also a principle issue in data visualization design. The second outlines relevant research in psychophysical scaling, which can explain many aspects of *Image* Theory. The goal in both sections is to provide a "first-principle" underpinning for *Image* Theory. The following section uses these first-principles plus additional data to interpret, correct and extend Bertin's theory and to suggest some important directions for future research. The closing section provides a final perspective on the relationship between visualization and vision research.

Before beginning, I want to clarify the scope of the ensuing discussion and to offer some disclaimers. First, I focus on visualization of abstract data and processes. Other areas of visualization, such as volume rendering, etc., have a somewhat different set of perceptual issues. Second, I will only cover a subset of *Image* Theory. The full theory is large and often obscure because Bertin uses idiosyncratic concepts and terminology. I will generally use examples from what Bertin calls "diagrams," simple XY(Z) graphs, but the concepts discussed below generalize to other forms of visualizations, including networks, maps and even symbols. This would be far too much to cover, so the discussion is restricted to the most relevant aspects and detail is omitted. Lastly, the perceptual literature is vast and complex. For the sake of clarity, I'll ignore many details and skirt many contentious issues.

2.0 Data Visualization is Limited by the Number of Data Dimensions

A good starting point for examining visualization and perception is the question: why is perception of multidimensional data a difficult problem? The obvious answer is that we can readily perceive data values for two variables represented on spatial axes, but we have difficulty at higher dimensions.

There have been several strategies for circumventing this limitation. One is divide-and-conquer: provide a number of different views of the same data. For example, data could be decomposed into an array of relatively simple 2 or 3-D graphs (e. g., Cleveland, 1984) with different axes. A more recent technique is to present different views in different formats or granularities. In either case, the observer must integrate information which has been distributed over space and/or time. This method is constrained by human processing limitations in attention and memory.

A second technique is to represent multidimensional data by means of "emergent features" (e. g., Montgomery and Sorkin, 1993). For example, one method is the use of Chernoff faces, which code individual data dimensions as facial features. The single features presumably combine to produce an emergent feature, facial expression, which is a nonlinear combination of the individual features.

The last, and most common technique, is to maintain the high dimensionality of representation but to code some data dimensions into nonspatial features, such as combinations of color, shape, size, brightness, etc. As discussed below, however, this method is usually limited to a single nonspatial variable.

While the success of these strategies varies, it seems clear that humans perceive data coded in spatial dimensions far more easily than those coded in nonspatial ones. Since there are only two spatial dimensions directly represented in the retinal image, observers have problems visualizing data in more than 2 spatial dimensions. The third spatial dimension may be added by means of depth cues although its usefulness is limited (see below). Beyond this point, however, multidimensional data are very difficult to perceive. This is the basic problem which visualization techniques must overcome.

In attempting to better represent high dimensionality data, it might be worthwhile to consider why this constraint exists. After all, there is no *a priori* reason to expect a limitation in the number of dimensions which we can process simultaneously or to treat the dimensions of space any differently from nonspatial dimensions such as shape, hue, brightness, etc. Instead of mapping dimensions to spatial axes, for example, a visualization could be constructed by mapping one data dimension to brightness, another to color, third to orientation, etc. Data could then be depicted by different combinations of brightness, color and orientation of single lines. In practice, of course, this does not work very well.

Why are data best represented on spatial axes and why are we able to perceive only a limited number of data dimensions at single time? A satisfactory explanation must in turn answer several more questions: Are spatial dimensions somehow more fundamental than nonspatial ones? Does the visual system use different sets of rules for coding spatial and nonspatial image attributes? If so, what are the differences and can they be overcome?

Bertin (1983) attempts to answer some of these questions within the context of *Image* Theory. The perceptual literature also provides partial answers to these questions by revealing how the visual system itself deals with the complexity of high dimensional images. In doing so, it suggests methods to improve data visualization.

3.0 Image Theory Is an Informal Perceptual Theory of Data Visualization

Bertin's *Image* Theory is an ideal medium to demonstrate the close connection between visualization research and the basic literature in perception, psychophysics and physiology. Bertin's interest is the problem of creating good multidimensional data visualizations. To serve as a guide, he formulated a theory of graphic perception, which was not explicitly based on any systematic, empirical observations but which was still scientific in form: it analyzes visualizations into a set of primitive components and specifies procedures for combining the primitives to create good visualizations.

Bertin's key concept is the *image*, from which the theory derives its name. Roughly speaking, an *image* is the fundamental perceptual unit of a visualization. An ideal visualizations will contain only a single *image* in order to optimize "efficiency," the speed with which observer can extract the information. Most of the theory is an attempt to explain how to create visualizations with a single *image*. In order to understand the concept of the *image*, however, it is first necessary to explain some concepts which Bertin uses.

3.1 *Image* Theory Specifies Graphic Primitives and Procedures for Data Visualization

The fundamental primitives of his theory are "invariants," "components," "correspondences," and "marks." Every graphical data representation consists of two primary parts, variational concepts called "components" and an "invariant," which is used to relate the components. These relations are depicted by mapping the components to some visual graphic variable and placing a "mark" on the visualization to show "correspondences" between/among components. Marks come in three different types, or "implantations," point, line and area.

The viewer extracts information in three stages. The first is "external identification," where the viewer determines what components are being represented in the visualization. The next stage is "internal identification," the process of determining which components are mapped to which visual graphic variables. Lastly, the viewer perceives the correspondences between/among components. It is the correspondences which actually convey information.

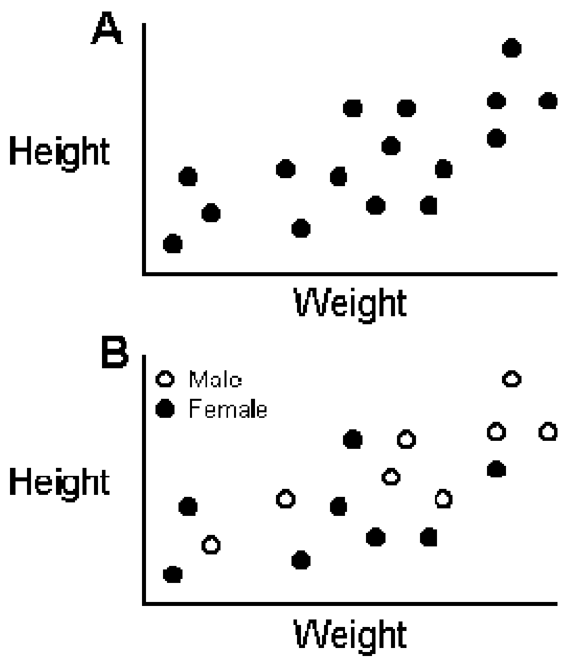

Figure 2

The simple example of a diagram shown in Figure 2a will clarify this schema. The components are height and weight, which have been mapped to the Y and X dimensions of space, and the invariant is the population of individuals. Each data point on the graph represents a "mark," showing the correspondence between a particular height and weight - a relation between the variational components. In terms of Bertin's three stages of information extraction, 1) external identification is the realization that the graph is showing something about height and weight, 2) internal identification is the perception that the height component is mapped to the vertical axis of the plane and weight to the horizontal axis, and 3) perception of the correspondences is the noting that the data point is located at a particular intersection of X and Y visual variables.

Image Theory, however, goes far beyond simply specifying a set of primitives and an explanation for the way observers read a data visualization. It also offers a model that describes the characteristics of optimal data visualizations and suggests methods for their construction. Both model and methods are stated in terms of perception.

3.2 Visual Variables Fall into Two Distinct Classes

According to Bertin, there are two functionally different classes of visual variable, planar and retinal. In Figure 2a, height and weight components are mapped to the two spatial dimensions of the "plane," the "planar" variables. In addition, components may be represented by six "retinal variables," size, color, shape, orientation, texture and brightness[2] (cf. Cleveland, 1984). Each retinal variable can be used with the three types of implantation.

In a typical visualization, the retinal variables allow representation of a third component. In Figure 2b, for example, the height/weight diagram might be extended to a third dimension, gender, by adding a brightness component and representing males and females with different brightness marks. A mark then would show a correspondence among three components, height, weight and gender, which have been mapped to three graphic variables, the two planar dimensions plus one retinal. Bertin believes that efficient visualization is limited to three component visualizations. It is not possible to create an efficient four component visualization by adding a second retinal variable: you could not, for example, add a component for country of birth by coding each mark with shape reflecting nationality. Both the three component limitation and the concept of "efficiency" are central to Bertin's theory, so I will return to them shortly.

Different classes of visualization can be distinguished by the correspondences which may be portrayed on the plane. A diagram, such as shown in Figure 2, can depict correspondence between all members of different components, but not between members of a single component. Figure 3a shows a network, where the nodes are the component members and the lines show correspondences. The defining property of a network is this additional possibility of portraying any and all correspondences between elements of a single component. In this case, and in the case of maps as well, the two planar dimensions represent a single component. Addition of a second component would require a retinal variable. In Figure 3b, the lines have different dashing, a line implantation of the texture variable.

Since there is only one planar variable, Bertin says that networks and maps are therefore limited to two components. It is not possible to simply use a second retinal variable as a substitute for the second planar variable. Bertin implies that there is a fundamental difference between planar and retinal variables and that they are not interchangeable.

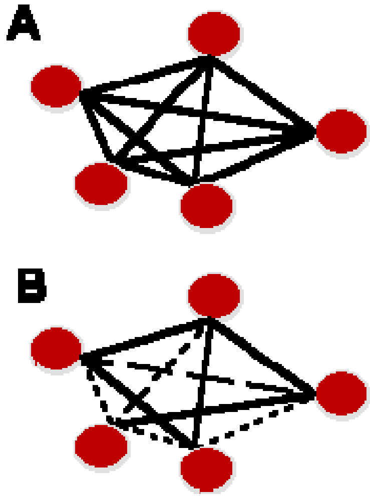

Figure 3

3.3 Good Visualizations Use Variables with Proper "Length" and "Level of Organization"

As we shall see, a clue to this fundamental difference lies in Bertin's notion that there must be a match in both "length" and "level of perceptual organization" between a component and the visual variable used to represent it. Bertin believes that the major problem of most data visualizations is the choice of visual variables with inappropriate length or level of organization.

3.3.1 Variable "Length" is the Number of Perceptible Steps

Length refers to the number of categories or steps (distinguishably different colors, brightness levels, etc.) A visual variable must have a length equal to or greater than the component it represents. If the number of data categories exceeds the number of distinguishable steps in the visual variable, then the data could not be properly perceived. For example, if a scale permitted 10 different data values and the visual variable provided only 8 distinguishable steps, the observer must perceive some of the different data categories as being identical.

A major problem for the visualization designer is to match the data to a visual variable with the correct length. Moreover, the designer must be sure that the steps in the visual variable (the brightness increments, etc.) were all distinguishable. Bertin's solution to this problem is to create of log scales. As will be discussed later, however, the usefulness of this approach is unclear because length varies with context.

3.3.2 "Level of Organization" Is the Level of Data which a Variable May Represented

Level of perceptual organization specifies the type of data scale, nominal, ordered, or ratio, which each visual variable can portray (Table I). That is, suppose the goal is to allow the viewer to extract ratios from the visualization, e. g., to immediately see that one value is twice another. The component must then be represented by a visual variable in which a doubling the variable's physical magnitude produces a doubling of perceived magnitude. Brightness could not be used, for example, because doubling intensity produces only about a 1.4 factor increase in perceived brightness. Bertin suggests that such failure to properly match the component and visual variable level of organization *is the major single source of error in visualization design*. An important aspect of his theory is that this error cannot occur when using the planar variables; they can be perceived at all levels of organization and therefore represent any component.

Table I: Original Bertin

	Associative	Selective	Ordered	Quantitative
Planar	Yes	Yes	Yes	Yes
Size		Yes	Yes	Yes
Brightness		Yes	Yes	
Texture	Yes	Yes	Yes	
Color	Yes	Yes		
Orientation	Yes	Yes		
Shape	Yes			

Associative Organization

Bertin subdivides the nominal scale level into two subcategories, associative and selective.

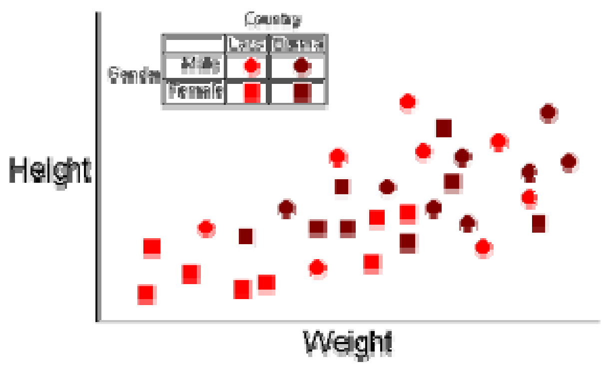

Figure 4

Associative perception is the lowest organization level, allowing grouping of all elements of a variable in spite of different values. For example, Figure 4 shows another version of the height/weight graph, with additional components represented by marks of different shapes (gender) and brightness (nationality). Shape is associative because the observer automatically groups different shapes together. If you look away and then glance quickly back at the graph, data points group by brightness, and shape differences are ignored. This makes the relationship among height/weight and nationality readily apparent. On the other hand, brightness is not associative because observers cannot easily ignore different values. In Figure 4, it is much harder to distinguish different shapes and to ignore brightness differences. As result, viewers will find it much more difficult to visualize the relationship among height, weight and gender. According to Bertin, planar dimensions, texture, color, orientation (point and line implantations only) and shape are associative while size and brightness are not.

Selective Organization

The next higher level, "selective perception," is the flip side of association. It permits the viewer to select one category of a component, perceive locations of objects in that category and ignore others. In Figure 2b, for example, the observer can visually select either the males or females by attending to either bright or dark marks. In Figure 4, observers cannot select shape, i. e., cannot select square or circle. Visual variables allowing selection include the planar variables, size, brightness, texture, color and point and line orientation implantations. Shape is the only variable which cannot be used for selection.

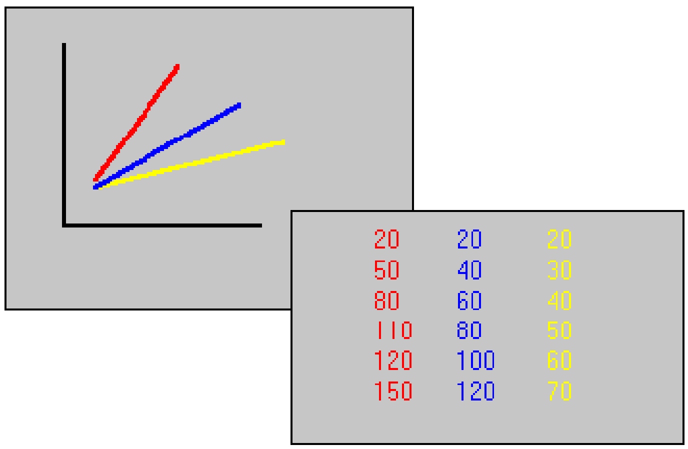

Figure 5

The concepts of associative and selective organization have implications for a wide range of issues in data visualization, such as the presentation of multiple data views. Figure 5 shows a typical example. One window shows data in graph form while the other shows a table. Which graph goes with which table? It's obvious if the two views are linked by color. If the informational view in the right window had a different color than it's corresponding object in the left view, the viewer will incorrectly form a nominal scale grouping. The message is clear: different views of the same data should be presented with the same values of variable at the selective level of organization.

Ordered Organization

Associate and selective perception provides only nominal scale classification. Bertin suggests that some variables permit an "ordered" organization, which allows the data to be ordinally ranked. Observers can see that one value of a variable represents a larger or smaller quantity than another. In Figure 2a, for example, increased age could be represented by data points with increased brightness. In this case, it would be possible to make judgments about relative height/weight of older and younger people. Bertin considers the planar variables, brightness, size and texture variables as ordered while excluding shape, color and orientation.

Quantitative Organization

The highest level of organization is "quantitative" which permits direct extraction of ratios, without need of consulting a legend, etc. That is, the ratio of variable values maps directly to the ratio of the data values. In figure 2a, if age were represented by data points of different size, which is a quantitative variable, the viewer could immediately see which individuals were twice as old as others. While ordered data permit only relative magnitude of the represented quantities, quantitative variables support perception of ratios. Only size and the planar dimensions are quantitative[3].

3.4 "Efficient" Visualizations are Single *Image*s

The relation of different visual variables to their levels of perceptual organization plays a critical role in Bertin's concept of *image*. Bertin emphasizes the difference between effortless and effortful organization. For a variable to offer a specific level of organization, a viewer must be able to perceive the level of organization without close scrutiny. Bertin notes that perception must be "immediate," and not necessitate scanning "sign by sign" through the marks. Throughout his work, Bertin stresses that good visualizations should permit "immediate" information extraction at a single glance with no need to move the eyes or attention. This quality of a visualization is its "efficiency" - how long it takes to perceive the required information. The best visualizations would permit critical information to be extracted at a glance, with no sign by sign scrutiny and ignoring all irrelevant correspondences. He uses the term *image* to refer to the portion of the visualization which "the eye can isolate ... during an instant of perception" and which allows the viewer to "disregard all other correspondences." As already noted, the *image* is limited to a maximum of three component correspondences in diagrams and two component correspondences in networks and maps.

Visualizations with four or more components are not efficient because they require integration across images. In Figure 4, for example, an observer could not immediately see information on Laotian males, since four components are involved: the two planar dimensions and two retinal components, brightness and shape. Such visualizations would be less than optimal because they contain two *image*s. The viewer would have to pick out a particular combination of shape and brightness, a task that cannot be performed with a glance but that would require effort and the movement of attention sign by sign.

To summarize, Bertin suggests that viewers extract information from visualizations through perception of correspondences between variables representing different data dimensions (components). There are only 8 variables, 2 planar and 6 retinal, available to portray components. Further, highly efficient visualizations, *image*s, are restricted to visualizations containing a maximum of two planar and one retinal variable. Lastly, as summarized in Table I, different variables support different levels of immediate perceptual organization. It is striking that only the planar variables permit all levels of organization, suggesting that spatial dimensions have a special role in data visualization.

Bertin bases his theory solely on his introspection and makes no attempt to provide any empirical data support. Moreover, it is a performance theory, with no grounding in perceptual or physiological research. It places the locus of control and explanation for behavior in the image rather than in the observer. In contrast, psychology and physiology explain perception by processes within the viewer. This subtle conceptual distinction is a major stumbling block to understanding the perceptual literature and to moving from a shallow to a deep theory of visualization. A basic knowledge of these processes would be helpful in predicting the behavior of a viewer to a new visualization.

Next, I will attempt to ground Bertin's theory in first principles of perception. The discussion focuses on data and theory from visual perception, psychophysics and physiology, which are most relevant to the problem of perceiving correspondences. I'll avoid much mention of external and internal identification since they both involve meaning and therefore higher level cognitive processes. Instead, I'll focus on two lower level aspects of Bertin's theory, each closely connected to a particular area of vision research. The first is Bertin's ideas about selective and associative organization and the 3 component limitation for constructing *image*s. This part of *Image* Theory parallels vision research in preattentive search and texture segmentation. The second is Bertin's views on ordered and quantitative organization and the method for determining length. The psychophysical scaling literature is concerned with virtually the same issues.

It will be clear that most of Bertin's intuitions, although vague in many respects, are well founded and, with some qualifications, run parallel to current research theory. My ultimate goal is to show how a grounding in basic vision research 1) permits a better, first principles, understanding of Bertin's theory, 2) suggests extensions and modifications and 3) can guide development of more efficient visualizations. Specifically, I will show that Bertin is incorrect about both the three component limitation to efficient visualization and the organization level permitted by several retinal variables. In addition, I will discuss the likely properties of additional retinal variables, flicker, motion and binocular disparity, which are currently available through more sophisticated computer graphics.

4.0 Vision Research Theory Supports Bertin

Most biological vision theories suggest that visual processing occurs in three steps: 1) formation of the retinal image, 2) decomposition of the retinal image information into an array of specialized representations and 3) reassembly of the information into object perception. The visual world is first coded as an image which produces activity in an array of retinal photoreceptors. The image is multidimensional in the sense that it contains information about a large number of image properties, such as color, shape, motion, etc. The processing of such complex data is an intractable problem (see Green, 1991 for a more detailed discussion), so the visual system decomposes the image into an array of simpler feature representations which can then be processed in parallel by different portions of the brain. At some later stage, the visual system reassembles the information to produce a coherent perception of the world. In short, vision researchers attempt to understand the brain, like any complex system, by means of a divide-and-conquer strategy: the brain as a collection of quasi-independent modules which perform different visual functions. Understanding visual processing reduces to a matter of specifying both the nature of the functional modules and their interactions.

Physiological data support the notion that the brain contains areas specialized for specific image features. Livingstone and Hubel (1988), for example, suggest that the brain contains 3 processing "streams," each containing neurons sensitive to different image attributes with one, for example, tuned to color, another to motion, etc. While neurons in each stream are tuned to a different image attribute, each has a spatially limited receptive field. Space (planar dimensions) is therefore the only attribute to which all neurons are sensitive. However, the picture presented by Livingstone and Hubel is probably far too simple, and the physiological basis for such distinct feature modules is becoming less compelling (deyoe and van Essen, 1988; Zeki, 1991; Schiller, and Logothetis, 1990; Merigan and Maunsell, 1993). However, it serves as a useful first approximation.

If the retinal image information is decomposed into separate feature representations, then the great mystery is how the visual system manages to reassemble features into an object. Attneave (1974) succinctly stated the commonly accepted answer when discussing a display containing a blue circle and a green triangle:

"If, as we have reason to believe, color and form are processed in separate parts of the nervous system, why does one not simply perceive circle, triangle, blue, green without knowing which form has which color? The simple answer, I think, is that blue and circle are tagged to the same spatial location." (page 109).

According to Attneave, then, spatial location is the attribute which holds retinal features together when time comes for reassembly. In Bertin's terms, space is special because location is the only attribute which all visual modules share, making it optimal for depicting correspondences.

4.1 Visualization Design Requires an Understanding of Preattentive Vision

Studies of visual search and image segmentation have also provided strong support for the notion that location glues together separate feature representations. This research has focused on one basic phenomenon: observers can instantaneously and effortlessly decompose some images into its constituent objects while other images require a slow and arduous scanning to locate objects. Why are some images instantly processed while others require time and effort and what does this reveal about the underlying organization of the visual system? To answer these questions, experimenters test observers with a variety of images in order to determine the conditions under which the effortless and effortful perception occurs. The studies have generated several theories of visual processing and visual representation.

In the basic paradigm, the observer detects the presence of a "target" element (or sometimes a group of elements - texture segmentation) embedded in a field containing various numbers of "distracters." The targets differ from the background distracters by one or more attribute, such as color, orientation, etc. The basic finding is that, under some circumstances, the observer seems able to detect the target (or segment the image) effortlessly - as if s/he were processing the entire visual field in a single automatic and parallel operation. This is often termed "preattentive" search because there is no need to focus attention on specific objects in the image: the target simply seems to "pop out." On other occasions, the observer must find the target by purposefully moving attention through space and serially scanning each object in the field.

The dichotomy of preattentive and attentive perception is clearly analogous to Bertin's distinction between immediate and sign by sign perception[4]. Moreover, other authors have similarly noted the importance of such distinctions. Cleveland (1984), for example, has noted that "elementary graphical-perception tasks," the preattentive perception basic graphical elements, underlies data visualization. He even created a list of basic tasks which overlaps significantly with Bertin's retinal variables. Abarbanel (1993) suggested that visualization can be defined as the substitution of "preconscious visual competencies" for "conscious thinking." Woods (1991) further said that "If the mental activities required to locate base data units are not automatic, but require mental effort or capacity, they may disrupt or divert...the main line of reasoning." The terms, "preattentive," "immediate," "preconscious" and "automatic" all highlight the necessity of designing displays so that the viewer can effortlessly perceive the fundamental visual elements, i. e., images.

A science of visualization will require a thorough understanding of the conditions which produce this preattentive, effortless processing. Fortunately (or unfortunately) there is a massive literature on this topic (see Green, 1991 for a partial review). To a first approximation, search is preattentive and parallel if 1) the target and distracters differ on a single "feature" such as color, orientation, etc. and 2) the difference in feature value (brightness, color, etc.) is great enough. If the target is defined by some combination of features (e. g., red and horizontal), then search becomes slow and requires effort.

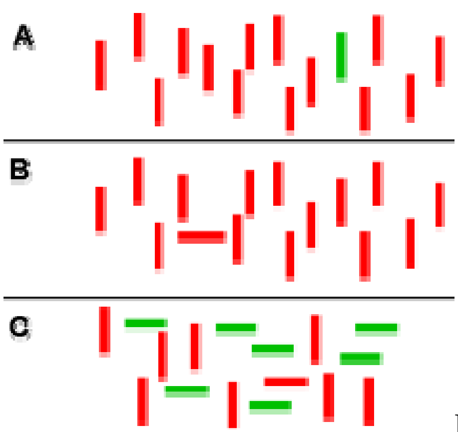

Figure 6

Figure 6 illustrates the difference between feature and conjunction search. The target (the odd item) is easy to find because it differs by a large amount in color (6a) in one case and orientation (6b) in the other. It is not even necessary to know in advance what the target is because of the preattentive pop out effect. On the other hand, search becomes serial and effortful if the observer must detect a "conjunction" of visual variables. In Figure 6c, the target is defined by a particular combination of orientation and color - horizontal and red. It no longer pops out, but requires a serial visual scan from item to item. The introspective reader will note that s/he had to move attention from location to location in order to find the target.

There a several theories which attempt to explain why some searches are preattentive and others require focal attention. The most influential, "feature-integration theory" (Treisman and Gelade, 1980) proposes special roles for space and attention. As suggested by Attneave, the brain represents individual features in different feature modules. That is, there is a color module, orientation module, etc. The feature modules contain no information on location, which resides in a "master map of locations." Other theories (Sagi and Julesz, 1985; Green, 1991, 1992), however, suggest that location is represented in the modules. Regardless of the exact theory, the generally held belief is that preattentive feature search can occur by examining only the contents of a single feature representation. Serial conjunction search, on the other hand, requires the observer to integrate features of a single object by reference to their common spatial coordinates. The integration is actually performed by focusing attention at the particular spatial location.

Figure 7 highlights key aspects of this theory. It shows the hypothetical feature modules for orientation and color, each represented in a different spatially mapped portion of the brain. Suppose the image has diagonal green lines, vertical red lines and a single diagonal red line. The visual system presumably decomposes the image into separate modules for orientation (A) and color (B). The viewer can always perform effortless feature selection with a single module. For example, the viewer can select the vertical marks without consulting the color module. In feature selection the target pops out, so there is no need to glue color and orientation together.

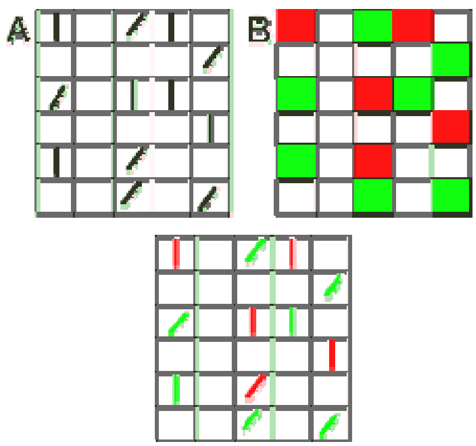

Figure 7

If, however, the selection must be based on a conjunction of features, such a horizontal and red, the task cannot be accomplished using either feature module alone. Instead, the viewer must combine the modules together before the target(s) become detectable.

It should be obvious by now that the most popular search theories run parallel to many aspects of Bertin's *Image* Theory. I've already noted the similarity of the immediate/sign-by-sign and preattentive/attentive dichotomies. Moreover, the difference between three and four component visualization parallels the distinction between feature and conjunction search. Recall Bertin's assertion that you can not create an *image*, a correspondence which can be perceived immediately, with more than three variables - two planar and one retinal. Although the search paradigm is not a exactly a visualization task, there are close similarities. Figure 6 does not have actual axes representing components, but locations of the objects still have implied components, the X and Y spatial coordinates. By extension, the top two panels would be analogous to Figure 2a. They are 3 component visualizations with the two planar variables (X and Y coordinates) and orientation (6a) and color (6b) as the retinal variable. It has been shown (Green, 1992) that if an observer can detect a target preattentively, s/he knows its location in XY space. In other words, s/he immediately/preattentively perceives the correspondence represented between the X and Y planar variables.

However, this is only possible with a single feature, i. e., 3 component, searches. They are single *image*s to Bertin because they use a single retinal variable. The conjunction search in 6c, which is analogous to the data visualization in Figure 4, requires effort because there are four components, 2 planar dimensions, orientation and brightness, so that the observer must integrate across two feature modules, or as Bertin would say, across two *image*s. In short, highest efficiency visualization only occurs when judgments can be made using preattentive vision, which in turn occurs only for feature (3 component) searches for diagrams.

To summarize, vision research provides first-principle explanations for many aspects of *Image* Theory. The three component limit is due to the way image features are represented in the nervous system and the difficulty of conjunction search. Planar and retinal variables are in fact different because spatial location ties all other attributes together and is hence naturally a better medium for depicting correspondences. Below, I add one more distinguishing attribute to the plane - it permits quantitative perception because it is seen as "veridical."

4.2 Psychophysical Scaling Explains Why Variables Have Different Levels of Organization

The creation of a single *image* requires that 1) the visualization contain no more than one retinal *image*/feature and 2) the visual variables have sufficient length and support the necessary level of perceptual organization. While the search/segmentation literature provided a first-principles explanation for the first point, there is another area of perception research, psychophysical scaling, which can similarly provide an empirical context for the second.

4.2.1 Stevens' Law Explains the Distinction Between Ordered and Quantitative Variables

Psychophysical scaling examines the psychometric function relating physical stimulus intensity to its perceived magnitude, or "sensation." Imagine, for example, that the goal is to learn the relationship between light intensity (physical variable) and apparent brightness (sensation). The observer views a series of lights differing in intensity and judges apparent brightness of each[5]. The results of such a study are typically expressed as a graph (Figure 8a) where sensation is plotted against physical intensity.

When psychophysicists perform such experiments with different physical dimensions, the curves typically fall into one of the three categories shown. Some curves are linear, meaning that sensation grows in direct proportion with physical intensity. In these cases, a doubling of physical intensity produces a doubling of sensation, so ratios are maintained. This is required for Bertin's quantitative perception where a doubling of the represented quantity should produce perception of doubled magnitude.

Unfortunately, linear curves are rare. Most physical dimensions produce nonlinear scales, with compressive functions being the more common result. In this case, doubling physical intensity does not produce proportional increases in sensation. As shown in Figure 8a, sensation grows more slowly than physical intensity, so that a doubling of intensity is perceived as a far smaller increase in brightness. Moreover, a doubling of intensity from 10 to 20 and one from 40 to 80 will be perceived as different relative changes. Physical variables producing such nonlinear scales would not permit quantitative level organization.

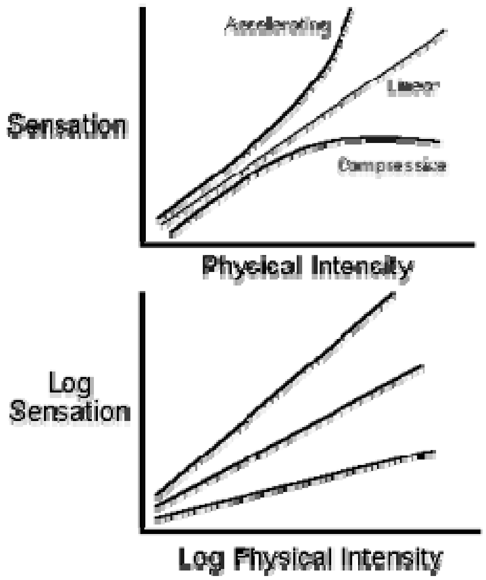

Figure 8

The relationship between sensation and intensity is frequently quantified by Stevens' Power Law:

$$\Psi = k\Phi^{\beta}$$

where psi is the perceived magnitude or "sensation," phi is the physical magnitude of the stimulus, *k* is a coefficient determined by the units of measurement and *beta* the exponent gives the slope of the line, when plotted on log-log coordinates, as in Figure 8b. This figure says that equal ratios of sensation are proportional to equal ratios of intensity[6].

The exponent and hence slope of the line in Figure 8b have a value of one when the relationship between sensation and physical magnitude is linear. Research (e. g., Gescheider, 1976; Spence, 1990) has shown that the only variable which reliably produces a slope of one is spatial extent, e. g., judgments of line length, etc., dovetails nicely with *Image* Theory, which states that the only quantitative variables are the planar dimensions and size. These are the only variables supporting quantitative organization because they are spatial extent judgments and hence are the only variables with psychometric power functions having an exponent of one, or "unbiased" in Cleveland's (1984) terminology. This is another special property of space.

The scaling data conversely explain why brightness is ordered but not quantitative. Studies always find brightness to be a compressive function. It is therefore not possible to directly perceive physical ratios - doubling the intensity of a mark does not double its perceived brightness. However, the brightness function is monotonic, so ordered organization is supported. The only caveat is that viewers must be able to discriminate the difference between brightness levels.

4.2.2 Weber's Law May Determine Length

Bertin notes that a visual variable must have sufficient length, i. e., be divisible into enough distinguishable intervals to represent the desired component. But how big must these steps be? How big a brightness difference is needed, for example, to permit preattentive selection based on intensity?

Bertin discusses methods for determining number of steps, advocating creation of log scales. At first glance, psychophysics seems to support this idea, since human discrimination threshold functions are generally described by Weber's Law:

$$K - \frac{\Delta I}{I}$$

where *I* is intensity of the physical stimulus, delta *I* is increment threshold needed to perceive a "just noticeable difference" (JND) and *K* is a constant fraction which differs across variables. This relationship suggests that JND's should be equal log steps and that higher K means larger absolute step sizes and shorter length[7]. To create a proper length scale, simply start at the lowest I value and add one JND for every data category.

Recall, however, that the step size must be preattentively perceptible. Most perceptual studies of Weber's Law are conducted under attentive processing, which does not correctly predict JND's needed for preattentive search. For example, discrimination of orientation requires only a 1^o to 2^o difference under focused attention, yet preattentive segregation requires orientation differences of about $12\text{-}15^o$ (Bergen and Julesz, 1982). It is not clear whether a JND function for preattentive search would necessarily follow Weber's Law as Bertin suggests. If it does, however, then the K value must be much greater than for attentive perception. Apparently the K value would vary with the type of judgment. Bertin implies this fact when noting that even for the same variable, length is shorter for associative/selective than for ordered/quantitative perception.

There is unfortunately little psychophysical work on the form of the psychophysical law for preattentive processing (but see Nagy and Sanchez, 1990, on color). However, the Bergen and Julesz orientation estimate is similar to that found using many other psychophysical techniques (e. g., Green, 1984) and is likely due to the bandwidth of cortical neurons with oriented receptive fields. Perhaps length and JND estimates for other variables can be can be obtained from existing psychophysical and physiological research.

In sum, the scheme suggested here predicts that the level of perception supported by a dimension depends on the shape of its psychometric function: 1) visual variables producing linear curves (exponent of 1) will permit both ordered and qualitative perceptual organization, 2) variables producing compressive (or accelerating) functions permit ordered perception but not quantitative perception and 3) variables which cannot be scaled for magnitude, such as shape, can be neither ordered not quantitative. Lastly, the accuracy of Weber's Law for determining length and step size is still unclear.

5.0 *Image* Theory Can be Extended Beyond Bertin

The picture painted so far suggests that vision research generally supports *Image* Theory. However, there are also research results to suggest that *Image* Theory requires some modification. First, Bertin's assignment of retinal variables to levels of organization is neither complete nor entirely accurate. Some visual variables can have levels of organization in addition to those suggested by Bertin. Moreover, advent of computer graphics permits use of several new visual variables which Bertin does not discuss. Second, there are several secondary perceptual effects which complicate direct application of *Image* Theory. Third, and perhaps most importantly, Bertin's belief that *image*s are always limited to three components is false.

5.1 Levels of Organization

In Table II, I use research data to propose corrections and extensions to Bertin's original level of organization assignments (Table I). Most of these modifications are based on the kind of search and scaling experiments described above.

	Associative	**Selective**	**Ordered**	**Quantitative**
Planar	Yes	Yes	Yes	Yes
Size		Yes	Yes	Yes
Brightness		Yes	Yes	Yes-if scaled
Texture	Yes	Yes	Yes	
Color (Hue)	Yes	Yes	Yes-limited	
Orientation	Yes	Yes		
Shape	Yes	Yes		
Motion:Velocity		Yes	Yes	Yes-if scaled
Motion:Direction		Yes		
Flicker:Frequency		Yes	Yes	Yes-if scaled
Flicker:Phase		Yes		
Disparity		Yes	Yes	

Table II: Updated Bertin

5.1.1 Shape can be selective

There has been much research (e. g., Treisman and Gormican, 1988; Julesz, 1984) designed to identify the primitive features which are represented in independent modules and which support preattentive, 3 component search. The results show that the list of primitives largely reflects those coded by the brain in primary visual cortex. Coincidentally, the list also agrees well, but not perfectly, with Bertin's list of retinal variables.

For example, Bertin claims that shape is always associative and never selective, but the literature is full of studies which contradict this assertion. While Figure 4 showed that squares and circles are indeed associative, Figure 9 shows a simple example of selection by shape based on "X" and "O". In fact, there are a large number of shapes which are selective but not associative (Julesz, 1984; Treisman and Gormican, 1988). Other data (Wang, Cavanagh and Green, 1994) further suggest that given sufficient practice and experience, almost any shape difference can support selectivity.

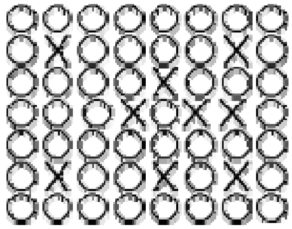

Figure 9

It is obvious why Bertin claims that shape was associative but not selective. He defined shape within a very narrow domain, solid shapes such as filled triangles, squares, circles, etc. These shapes are highly similar because most of their energy is in the low "spatial frequency" band, which dominates rapid perception (Green, 1984). Unfortunately, space do not permit much explanation here of the spatial frequency concept, but there are many references on the topic (e. g. De Valois and De Valois, 1988).

Why are some shapes associative while others are selective? While the answer is not entirely clear, the question has been approached by creating computer models (e. g., Graham, Sutter and Venkatesan, 1993) of the visual system. The image is typically processed by a series of cascaded linear (spatial frequency) and nonlinear (edge, etc.) filters which presumably correspond to neuronal levels in the brain. The decision as to whether two shapes/textures are associative or selective is determined by the similarity of their outputs in the model. In theory, such a model should make it possible to predict whether shapes/textures would be associative or selective. The applicability of such models to predicting immediate visualization of diagrams is unknown.

5.1.2 Color can be ordered

Color consists of three properties, hue, saturation and brightness (or hue, chroma and lightness). Bertin's discussion of color is incomplete because, while distinguishing brightness and color, he fails to separate hue from saturation. In contrast to Bertin's assertions, moreover, each variable permits ordered perceptual organization under some conditions.

First, it is true that hue is in general a nominal variable; red, green, blue, etc. do not form an ordered scale. However, over small ranges, hue can be ordered. For each hue, there is a specific example, called a "unique hue," which is perfect in that it is not tinged by any other hue. For example, unique yellow contains no trace of either red or green. In a sense, it is the most intense sensation of yellowness possible. An ordered scale of yellow could therefore be constructed starting around unique yellow and extending to either (but not both) unique red and unique green. Observers can then readily order hues along the yellow-green or yellow-red continuum. For example, there is a common color vision exam, the Farnsworth-Munsell Test, which requires observers to properly order colored chips in this way. Second, saturation, the amount of white mixed with a spectral, i. e., the purest, hue, is also an ordered variable. Viewers can readily order lights by increasing amounts of saturation.

5.1.3 Brightness can be "quantitative"

Brightness is ordered but not quantitative because it produces a psychometric function with an exponent less than one. If this idea is correct, then it should be easy, using existing psychophysical data, to promote brightness to a quantitative level of organization by rescaling the intensity axis.

The solution is to scale the X axis in Figure 8a so that the curved line becomes straight, i. e., make the steps at the low end of the scale larger. The exponent of the brightness power function varies somewhat with exact conditions but is usually somewhere around 0.5 - a square root law:

The nonlinear gamma built into most CRT monitors rescales brightness-grey level function to compensate for the visual brightness nonlinearity. However, degree of compensation is often overstated for two reasons. First, the visual system’s brightness exponent varies with conditions and can be as low as 0.33. Second, CRT monitor’s effective gamma varies with viewing conditions, such as ambient light and the exact monitor brightness and contrast settings. Never the less, brightness will approximate a quantitative variable on a computer screen but not on a printed page.

5.2 Additional visual variables

Although Bertin recognizes that computers present new opportunities for data visualization, his discussion of retinal attributes is limited to those generally available by print technology. The flexibility of computers and CRT displays opens opportunities for going beyond Bertin. First, I have already described how level of organization for some retinal variables can be upgraded by rescaling. Second, data can be displayed by means of several new visual variables, including motion, flicker and disparity (stereo depth). Using the analysis described above, psychophysical data should be able to predict the level of organization supported by these new variables (Table II).

5.2.1 Motion

One new retinal variable is motion, which can be split into two subvariables, velocity and direction. Both are undoubtedly selective, as demonstrated by studies showing preattentive search based on motion differences[8]. Further, the Gestalt principle of "common fate," says that humans interpret objects moving or flickering together as a single segment (e. g., Green and Odom, 1991). Marks which move together will easily be selected from those with differing motions.

It is unlikely that motion is associative. Objects with different motions are generally perceived as lying on different surfaces. Segmentation of an image into constituent surfaces is perhaps the earliest and most primitive perceptual function (Gibson, 1966). As a general rule, humans find it exceeding easy to select, but very difficult to group, objects appearing to lie on different surfaces. As a result, it is unlikely that marks can be associated across different motions[9].

Motion velocity is likely to be ordered, since it is a continuum of magnitude, and observers can readily discriminate steps of increasing value. Direction of motion is likely not ordered since it is not a variable of magnitude. Finally, velocity is not quantitative because it produces a compressive psychometric function. As noted above, however, velocity might be made quantitative by appropriate rescaling. Motion can also be used to portray depth by means of motion parallax, a highly potent depth cue.

5.2.2 Flicker

Flicker is another possible variable, but there is surprisingly little relevant research. Since the perceptual mechanisms which subserve motion and flicker overlap significantly (Green, 1984), it seems likely that flicker would support levels of organization similar to those possible with motion velocity.

Like motion, flicker also has two subvariables, frequency and phase. Flicker frequency, the speed of the on-off cycles, is selective by common fate, can be readily ordered, and could be promoted to quantitative by rescaling. There are, however, some problems with flicker frequency. First, it has very short length. Second, apparent brightness varies with frequency (the "Broca-Sulzer Effect"), making these somewhat "integral dimensions."

Phase refers to the relative point in the on-off duty cycle. For example, "in-phase" lights come on and off simultaneously. In "counter-phase" flicker, one light becomes bright while the other becomes dark and vice versa. Marks of different phase are readily selective by common fate. However, phase is not readily ordered. Neither flicker frequency nor phase is likely associative.

5.2.3 Binocular Disparity

A third new variable is binocular disparity, which can be created by giving the left and right eyes slightly different views of the same visualization. In normal viewing disparity is not an important cue for judging distance. In fact, humans get little absolute distance information from disparity, once objects are more than a few feet away. The classic study of Holway and Boring (1941), for example, even found that people judged distance more accurately under monocular conditions than under stereoscopic viewing.

Instead, disparity provides good *relative* depth information, revealing whether one object is closer or further than another, and is a very powerful cue for image segmentation. Its main role is to help break scenes into meaningful surfaces, so it is likely to be poor for association but ideal for selection. Since it provides only relative depth information, it is a poor choice to represent quantitative data but good for representing ordered data. Moreover, stereoacuity, the discrimination of small disparity differences, is one our sharpest senses, suggesting that disparity has extremely long length and is ideal then there are many data intervals.

Lastly, it would be hard to upgrade disparity by rescaling, since the relationship between disparity and perceived distance is complicated and varies with individual differences in factors such as interpupillary distance, the gap between the viewer's two eyes.

5.3 Factors Complicating *Image* Theory

Image Theory is meant to be a very general approach to data visualization. While research described above has validated it's broad strokes, there are many more recent findings which refute the simplistic view of both Feature-integration and *Image* Theory.

5.3.1 Visual variables are not processed independently

While feature-integration theory, a virtual analog to *Image* Theory, captures many aspects of visual search and segmentation, there are also contrary data. These are reviewed elsewhere (Green, 1991), but I'll just point out one particularly relevant problem. If features exist in independent computational modules, then preattentive search should not be impaired by orthogonal variation in an "irrelevant" dimension. That is, if a scene is to be preattentively segmented by texture, then any unsystematic color variation should be "irrelevant." Several studies, however, clearly show that orthogonal variation of irrelevant features does, in fact, impair preattentive processing. For example, it takes longer to select on the basis of orientation if there are color variations which must be ignored (Callaghan, 1989). Bertin does not acknowledge that such interaction between retinal variables can occur. In planning a visualization of optimal efficiency, variables should be chosen to minimize such crosstalk between dimensions.

Some dimensions interact so strongly that they cannot be separated without great effort. Garner (1974) suggested that different sets of visual image variables can be classified as "integral" or "separable." Separable dimensions are processed independently in preattentive vision. If asked to sort a sequence of pictures containing colored dots located in different parts of the pictures, observers can do so quickly using either color or location as the relevant variable and easily ignore the other. The planar are also dimensions are clearly separable. On the other hand, hue and saturation are integral - observers cannot ignore hue when trying to attend to saturation and vice versa. If a formalism maps one data dimension to hue and another to saturation, for example, then the visualization would not be efficient: users would need focal attention to separate the dimensions.

5.3.2 Association is influenced by "Pragnanz"

Bertin's discussion of association overlooks several factors which affect the way image elements group together. In the late 19th and early 20th century, the Gestalt school of perception suggested that human vision is based on the principle of "Pragnanz," which says observers have a strong innate predisposition to group image elements into the simplest "good form." Bertin notes that the planar and several retinal variables are associative. Gestaltists called these the laws of grouping by proximity and similarity.

There are several additional Gestalt grouping principles which can influence the way viewers group visual variables. Groupings can be created by "common fate," the tendency to move or flicker together, "good continuation", the tendency to perceive shallow curves, and symmetry.

5.3.3 Selective/Associate organization depends on variable values

Image Theory is complicated by secondary perceptual effects. A good example is the "asymmetry effect" in visual search. Research (Treisman and Gormican, 1988) has shown that the delectability of a target, whether it supports selective perception, depends on it's exact location on the value continuum.

An example will clarify. Suppose the display contains line segments which are oriented vertically or tilted by 45^o. Viewers find it easier to locate the tilted lines among vertical background items than vertical lines among tilted backgrounds items. In *Image* Theory terms, it is easier to select tilted than vertical lines. Therefore, whether a variable is associative or selective depends in part on the exact values chosen on a given visual variable.

5.4 Conjunction searches (four component images) are possible

Bertin's assertion that *image*s can contain no more than three components is seemingly supported by the general difficulty of conjunction search. However, studies have sometimes found that observers can preattentively find conjunctions, suggesting that four or more component *image*s are possible.

5.4.1 Motion and Disparity allow 4 component images

Studies show that a subset of visual variables permits conjunction search. Observers can effortlessly search for conjunctions of shape or color with disparity (Nakayama and Silverman, 1986) or motion (McLeod, Driver and Crisp, 1988). These searches have four components: 2 planar, 1 Bertin retinal variable plus disparity or motion, variables which Bertin does not include in his analysis.

Although there are perspective drawings in his book, Bertin says little about the possibility that depth can be an additional image component. The use of depth cues such as perspective, shading, transparency, motion parallax or binocular disparity, might admit the third dimension as a third spatial variable, allowing four component visualizations[10].

There are data which suggest that these depth cues permit *image*s with additional components. I have already noted that observers can search for conjunctions of binocular disparity and shape or color, which suggests that the third dimension can be used as an extra component for selection. It is unknown whether there are any other depth cues which could be similarly employed, or whether any could produce a quantitative level of organization for the third dimension. There is no reason to assume that the Z axis would act like the XY planar dimensions and support all levels because the depth spatial dimension is fundamentally different from the planar dimensions. The retinal image contains an explicit representation of the plane, but the third dimension is available only implicitly and must be computed by the various depth cues. As discussed above, depth cues are probably selective and ordered, but it seems unlikely that any of the retinal variables which could be used to portray depth would be quantitative. It might be best to think of disparity as providing only a "2 1/2 D", rather than a 3D, visualization.

5.4.2 Practice can create 4 or more component images

I've already noted evidence (Wang, et al., 1994) that arbitrary shapes can become selective with sufficient experience. Another study (Wolfe and Franzel, 1989) shows that observers can preattentively select almost any conjunction of Bertin's retinal variables given sufficient practice. Wolfe and Franzel further found that practiced observers can sometimes search faster with five components (conjunctions of three features) than with four. This suggests the counter-intuitive notion that higher numbers of components can produce higher efficiency visualizations.

6.0 Conclusion and References

I have outlined some basic research on human vision and suggested ways that this research could be used to guide data visualizations. Bertin's *Image* Theory has served as a valuable bridge to link the basic perceptual, psychophysical and physiological literature and visualization. Many of my proposals are only speculation because there has been little empirical research directly linking search/texture and scaling paradigms to data visualizations. However, if the field of visualization is to progress in a coherent manner, then it must be based on some sound empirical footing. Research on basic human vision seems a good candidate.

I should also reiterate that in the attempt to draw parallels between Bertin's theory and psychophysical data, I have (over)simplified many issues. The perception of images is a highly complex process, and it is therefore impossible to consider all the relevant factors. For example, I have ignored context effects, such as brightness and color contrasts, size constancy and other phenomena, which complicate the straightforward theorizing outlined above. The proper conclusion to draw is that vision research can be a helpful guide in designing visualizations, but cannot insure optimal visualizations. Perception is just too complicated, so usability testing will always be required.

Lastly, construction of a visualization must taken into account the viewer's goals and purposes. In choosing variables to represent different data components, for example, it is necessary to know what level of perceptual organization needs to be supported. This, in turn, depends on the kind of judgments that the user wants to make. Design of visualization is partly a knowledge engineering problem.

REFERENCES

Abarbanel, R. (1993) Problem solving with and without visualization. (Ed. G. Nielsen and D. Bergeron) *Proceedings of Visualization '93*, 343-344. IEEE Computer Society Press: Washington.

Attneave, F. Apparent motion and the what-where connection. *Psychologia*, **17**, 108-120.

Bertin, J. (1983) *The semiology of graphics*. Univ. Wisconsin Press: Madison, Wisc.

Callaghan, T. (1989) Interference and dominance in texture segregation: Hue, geometric form and line orientation. *Perception and Psychophysics*, **36**, 299-311.

Cleveland, W. (1984) *The elements of graphing data*. Wadsworth: Monterey, Ca.

De Valois, R. and De Valois, K. (1988). *Spatial Vision*. New York: Oxford Press.

De Yoe, E. A. & van Essen, D. C. (1988). Concurrent processing streams in monkey visual cortex. *Trends in Neuroscience*, **5**, 219- 226.

Garner, W. (1974) *The processing of information and structure*. Erlbaum: Hillsdale, N. J..

Gescheider, G. (1976) *Psychophysics: Method and theory*. Erlbaum: Hillsdale, N. J.

Gibson, J. J. (1966) *The Senses Considered as Perceptual Systems*. Boston: Houghton Mifflin.

Graham, N., Sutter, A., and Venkatesan, C. (1993) Spatial frequency- and orientation-selectivity of simple and complex channels in region segregation. *Vision Research*, **33**, 1893-1911.

Green, M. (1984) Masking by light and the sustained-transient dichotomy. *Perception and Psychophysics*, **34**, 617-635.

Green, M. (1991) Visual search, visual streams, and visual architectures. *Perception and Psychophysics*, **50**, 388-403.

Green, M. (1992) Visual Search: detection, identification and localization. *Perception*, 21, 765-777.

Hamming, R. (1962) *Numerical Methods for Scientists and Engineers*. McGraw-Hill: New York.

Holway, A. and Boring, E. (1941) Detriments of apparent visual size with distance variant. *American Journal of Psychology*, **54**, 21-37.

Julesz, B. (1984) Toward an axiomatic theory preattentive vision. In G. M. Edelman, W. Gall and W. Cowan (Eds.), *Dynamic Aspects of Neocortical Function* (pp 585-711). New York: Wiley.

Levkowitz, H. and Herman, G. (1992) Color scales for Image data, *IEEE Computer Graphics and Applications*, January, 72-80.

Livingstone, M. and Hubel, D. (1988) Segregation of form, color, movement and depth: anatomy, physiology, and perception. *Science*, **240**, 740-749.

Merigan, W. and J. Maunsell (1993) *Ann. Rev. Neurosci.* **16**, 369-402.

Merwin, D. and Wickens, C. (1993) Comparison of eight color and grey scales for displaying continuous 2D data. *Proceedings of the Human Factors and Ergonomics Society 37th Annual Meeting*, 1330-1334.

Montgomery, D. and Sorkin, R. (1993) *Proceedings of the Human Factors and Ergonomics Society 37th Annual Meeting*, 1325-1329.

Nagy, A. and Sanchez, R. (1990). Critical color differences determined with a visual search task. *Journal of the Optical Society of America A*, **7**, 1209-1217.

Nakayama, K.and Silverman, G. (1986). Serial and parallel processing in visual feature conjunctions. *Nature*, **320**, 264-265

Pun, T. & Blake, E. (1990) Relationships between image synthesis and analysis: towards unification. *Computer Graphics Forum*, **9**, 149-163.

Rogowitz, B. and Treinish, L. (1993) Data structures and perceptual structures. Human Vision, *Visual Processing and Digital Display IV*, 600-612.

Sagi, D. & Julesz, J. (1985). "Where" and "what" in vision. *Science*, **228**, 1217-1219.

Schiller, P. and N. Logothetis (1990). *Trends Neurosci.*,**1**, 1.

Schneider, G. (1969) Two visual systems. *Science*, 165, 895-902.

Spence, I. (1990) Visual psychophysics of simple graphical elements. *Journal of Experimental Psychology: Human Perception and Performance*, **16**, 683-692.

Treisman, A. and Gelade (1980) A feature-integration theory of Perception. *Cognitive Psychology*, **14**, 12-31.

Treisman, A. and Gormican, S. (1988). Feature analysis in early vision: Evidence from search asymmetries. *Psychological Review*, **95**, 15-48.

Zeki, S. (1989). Functional specialization and multi-stage integration in the visual cortex. *Perception*, **18**, 263.

Wang, Q., Cavanagh, P., and Green, M. (1994) Familiarity and pop-out in visual search. *Perception & Psychophysics,* **56**, 495-500.

Wolfe, J and Franzel, S. (1989). A modified feature-integration model for visual search. *Journal of Experimental Psychology: Human Perception and Performance,* **15**, 419-433.

Woods, D. (1991). The cognitive engineering of problem representations. In (G. Weir and J. Alty, eds.) *Human-Computer Interaction and complex systems*. New York: Academic Press.

FOOTNOTES

[*]Presented in part at *Vizualization '91*

[1] The word "image" will be used in two senses: 1) the more general sense of an array of light and 2) Bertin's sense as the fundamental unit of visualization perception. The avoid confusion, the word will appear in italics when referring to Bertin's *image*.

[2] By texture, Bertin means only marks made from the identical microelements which differ in size. Textures made from different microelements are area "implantations" of a shape difference. This distinction is not made in vision research terminology. Bertin also employs the term "value" instead of the more commonly used "lightness" or "brightness."

[3] Bertin, in discussing quantitative data, always gives examples which are ratio judgments. It's not clear whether he draws any distinction between interval and ratio data. This is an a significant oversight, since interval scales have no true 0 point and therefore do not permit direct extraction of ratios. Although the quantitative category should probably be subdivided in ratio and interval, I'll ignore this point and follow Bertin in assuming that quantitative equals ratio scale.

[4] The sharpness of the attentive/preattentive dichotomy is in some debate, but the issues are too involved to cover here. See Green, 1991, for a discussion.

[5] The methodology for obtaining brightness judgments is not relevant here, but is described in many books, e. g., Gescheider, 1976.

[6] There is debate on the canonical form of the psychometric function. In some cases, the relationship follows Fechner's Law, which predicts a straight line on a linear-log plot and says that equal ratios of intensity produce equal intervals of sensation. However, choice of psychophysical law does not affect any argument made here.

[7] The insightful reader might note that the log step JND's apparently conflict with the power functions described above; logarithmic changes in intensity (the X axis) would be linear with arithmetic changes in the sensation/JND (the Y axis). The discrepancy can be reconciled by assuming that JND's are themselves logarithmic and follow Weber's Law. The alternate is to adopt Fechner's Law, which is a log scale.

[8] Preattentive segmentation based on motion occurs for "short range" motion but not "long range" motion. This distinction, however, is beyond the scope of this discussion.

[9] The issue of surface perception underlies the concepts of associative and selective visualization. See Gibson, 1966 for a discussion of surfaces and perception.

[10] It is true that 3D graphs are commonly used for data visualizations. In order for them to allow 4 component *image*s, however, the correspondences among the components must be perceived without attentive scrutiny. It is unclear whether this actually occurs.