The Sectioned Density Plot

Dale J. COHEN and Jon COHEN

Effective graphical presentation efficiently summarizes, exposes, and communicates patterns in data. Here we describe a new plot, the *sectioned density plot*, that compares full distributions across multiple groups. We designed the sectioned density plot to exploit the visual system's natural ability to interpret occlusion and intensity variation as changes in depth. By incorporating depth into the graphical display, we were able to combine the ability of the boxplot to display trends in variance and central tendency with the ability of a histogram/kernel density plot to present distribution shape.

KEY WORDS: Boxplot; Comparing distributions; Graphical presentation; Kernel density; Perception.

1. THE SECTIONED DENSITY PLOT

Comparing distributions across multiple groups presents a significant design problem. There are at least three important dimensions of the data that should be displayed (the values of the data on the continuous variable, the frequency or density of those values, and the values of the grouping variable), and more often than not, this information is presented in two-dimensional x-y space. Because there are fewer spatial dimensions than data dimensions, the display of at least one of the data dimensions must be compromised.

Compromising the display of the data values or density limits the information conveyed about the distribution shape. Tukey's (1977) boxplot is an example of such a graphic. Tukey's boxplot summarizes the shape of a distribution by displaying only key characteristics of the distribution (i.e., median, interquartile range, skew, and outliers). This allows many distributions to be presented side by side so the reader can compare these characteristics across groups. Two disadvantages result from summarizing distribution shape. First, because a symbol system is required to display the key characteristics of the distribution, the boxplot requires a skilled interpreter. Second, and more importantly, the boxplot can render very different distributions similarly. This disadvantage is highlighted in Figure 1.

Compromising the display of the values of the grouping variable inhibits the detection of trends across groups. The kernel density plot (or equivalently, the histogram) is an example of such a graphic. The kernel density plot approximates the probability density function of the population from which the data were sampled and plots the contour of that function. By doing so, the kernel density plot retains the ability to present the precise shape of a distribution. To compare distributions across multiple groups, one can plot many of these contours on the same axis and use a legend to identify groups. Such a solution, however, requires the reader to shift his or her attention repeatedly between the graph and the legend which inhibits the ability of the viewer to detect trends across the groups. This difficulty is highlighted in Figure 2.

We can stack kernel density plots to obtain a more effective representation (see Figure 3). Although the stacking removes some clutter, the viewer must still integrate information across graphs, which requires noticeable cognitive effort. The extra

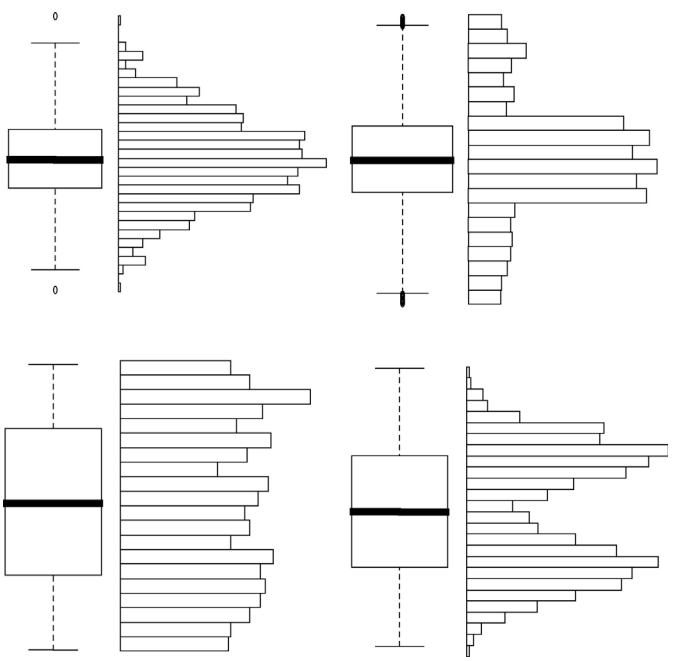

Figure 1. Normal, square wave, uniform, and bimodal distributions depicted using vertical boxplots and histograms. Note how the shapes of the original distributions are unrecoverable from the boxplots alone.

Dale J. Cohen is Professor, Department of Psychology, University of North Carolina Wilmington, Wilmington, NC 28403 (E-mail: *cohend@uncw.edu*) and Director of Research, Memory Assessment and Research Services, 1241A Military Cutoff Rd., Wilmington, NC 28405. Jon Cohen is Vice President and Director of Assessment, American Institutes for Research, Washington, DC 20007 (E-mail: *jcohen@air.org*). This project has been funded at least in part with Federal funds from the U.S. Department of Education, National Center for Education Statistics under contract number RN95127001. The content of this publication does not necessarily reflect the views or policies of the U.S. Department of Education, National Center for Education Statistics, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.

 DOI: 10.1198/000313006X108972

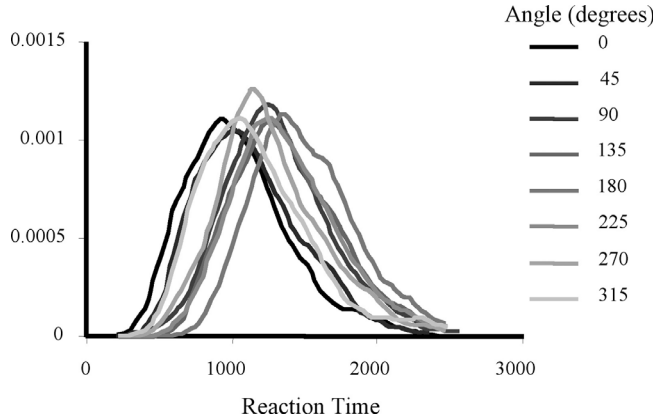

Figure 2. A kernel density plot of reaction time (RT) data from a traditional mental rotation task. To ease interpretation, the intensity of the line color correlates with angular disparity. One cannot readily see the typical inverted "V" where RTs peak at 180 degrees of angular disparity. It is the distribution overlap and the requirement of a legend that make it difficult to identify trends in the central tendencies of the data.

effort is likely because the densities of each plot lie on separate axes. As a result, to compare densities across distributions, the observer must read and store in memory the density in Distribution A, then read the density in Distribution B, and compare the two. One can partially alleviate this problem by imagining the densities rising into the third dimension. This allows the perceptual system to compare densities by comparing heights on a single, imagined z-axis. Nevertheless, this exercise in imagination takes effort, and once achieved, fades quickly. Thus, although the variations of the kernel density plot alleviate some of the difficulties associated with comparing distributions across groups, the variations are not ideal.

The *sectioned density plot* combines the ability of a histogram/kernel density plot to present shape information with the boxplot's ability to present trends in variance and central tendency. The attempt to combine the advantages of kernel density plots and boxplots is not new: Hintze and Nelson (1998) did so with some success when designing their violin plot. The violin plot adds rotated kernel density plots on either side of a boxplot. The sectioned density plot presents similar information in a more cohesive image by equating the spatial and data dimensions. This is accomplished by making effective use of the third dimension (z-space).

The visual system automatically and effortlessly interprets many two-dimensional visual features as distance in a third dimension (Cutting and Vishton 1995). The most widely adopted perceptual cue to the third dimension in graphical methods is motion (Pastizzo, Erbacher, and Feldman 2002; Wainer and Velleman 2000). Although it is possible to exploit motion for graphics designed to be viewed interactively on a computer, motion is not an option for printed scientific journals. There are, however, two other salient features that the visual system automatically and effortlessly interprets as distance in a third dimension that are appropriate for the printed page: occlusion and intensity variation.

Occlusion refers to the phenomenon whereby objects closer to the viewer will obscure objects farther from the viewer if the two objects share positions in two-dimensional space. Unlike superimposition, in which the closer object may be semi-transparent, for occlusion to occur the closer object must be opaque so that no light from the further object can pass through the closer object to reach one's eye. For example, most viewers perceive the image in Figure 4(a) as a square in front of a circle. Here, the only cue to the depth interpretation is occlusion. Interestingly, occlusion is only implied: there is no direct evidence that the shape perceived as a circle is actually a circle. That is, although it is possible to perceive the image as a square next to a "Pac-Man" like figure, this interpretation is almost never spontaneously realized. Cutting and Vishton (1995) summarized the data addressing the many cues to perceiving spatial depth and conclude that occlusion is the strongest cue to depth of all those reviewed. Although occlusion is a powerful cue to spatial depth, statistical graphics virtually never use it.

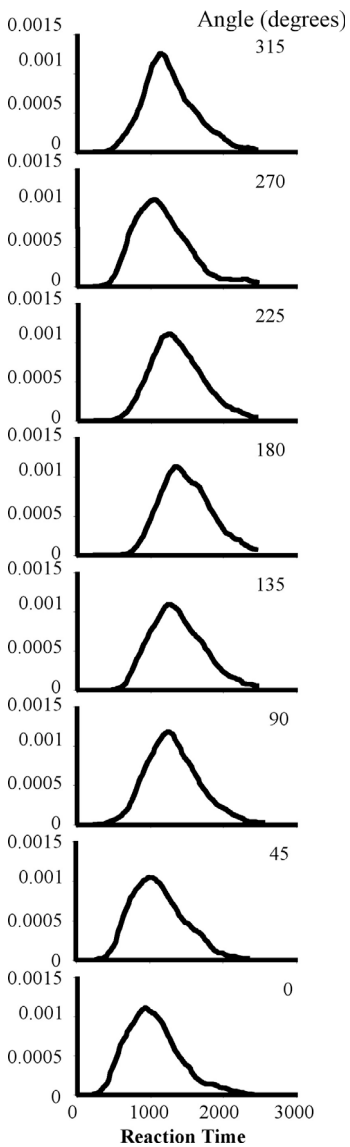

Figure 3. Stacked kernel density plots of the same data presented in Figure 2. This variation improves one's ability to detect the trend in central tendency present in the data, but to do so requires noticeable cognitive effort and significant page space.

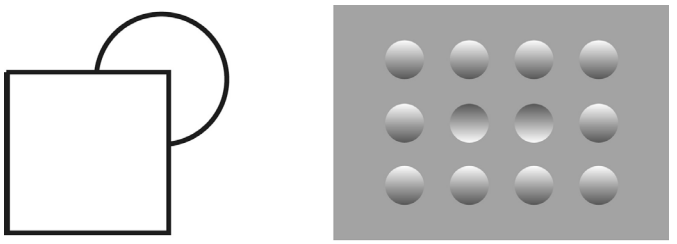

Figure 4 (a) An example of occlusion implying depth. (b) An example of how intensity variation specifies three-dimensional shape: the circles appear as three-dimensional shapes. The visual system assumes a single light source (generally from above), so the outer circles appear convex and the inner circles appear concave. If one rotates the picture 180 degrees, then the polarity of the shapes will reverse.

Intensity variation refers to the gradual change of intensity that is related to the three-dimensional shape of an object [also referred to as shape from shading: e.g., Mingolla and Todd (1986); Ramachandran (1988); Todd and Mingolla (1983)]. Ramachandran (1988) demonstrated that the brain automatically and effortlessly interprets intensity variation as changes in depth even when the intensity variation does not strictly conform to what would occur naturally. In this demonstration, Ramachandran created a display where circles were filled with a flat transition between light and dark (see Figure 4(b)). Most observers perceived the displays as hemispheres even though a naturalistic presentation of a hemisphere requires that the shadows follow the hemisphere's contours. Wainer and Francolini (1980) demonstrate that when most viewers read a graphic they quickly and accurately interpret changes in intensity as changes in amount—even without the benefit of a legend. This ability, however, relies on the systematic ordering of intensities. Legge, Gu, and Luebker (1989) showed that randomly ordered intensity variations are difficult to interpret.

The *sectioned density plot* displays information in three-dimensional space by exploiting the natural ability of the visual system to interpret occlusion and changes in intensity as depth. Although either intensity variation or occlusion alone would provide aid in perceiving the third dimension, we feel that the

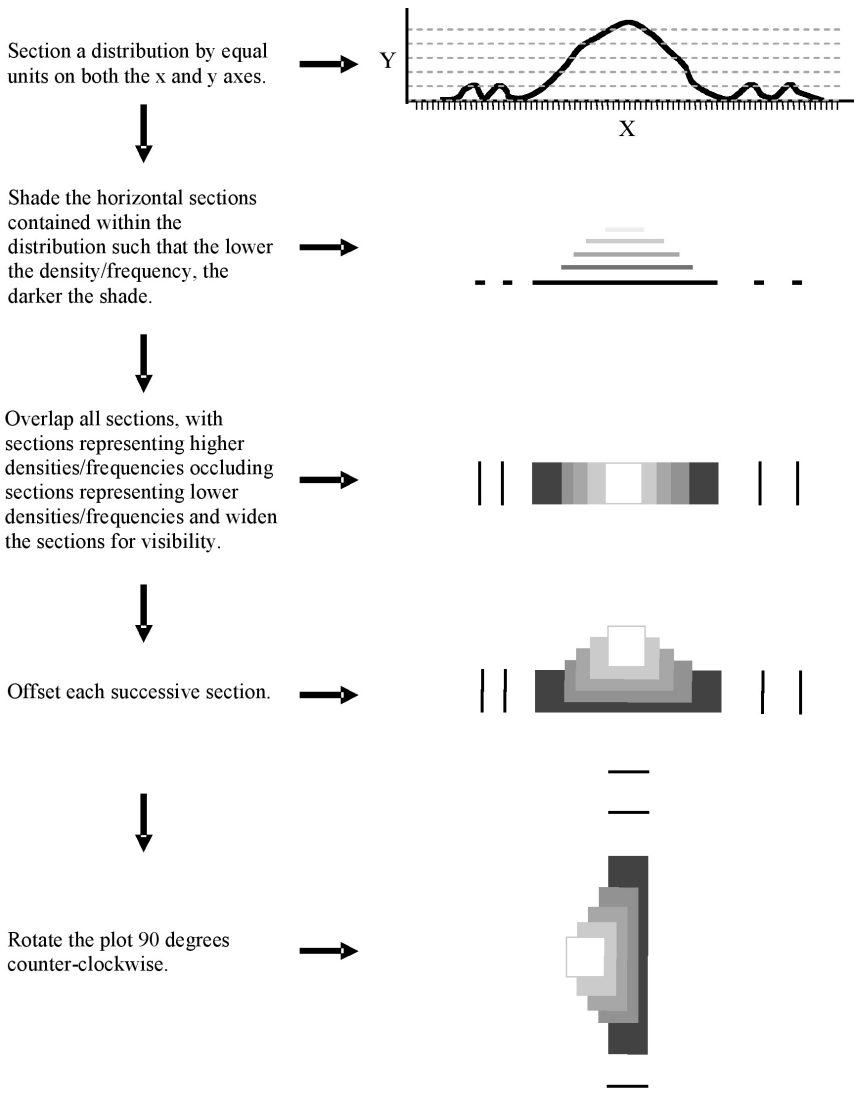

Figure 5. A stylized example of how to create a sectioned density plot.

combination of cues is particularly effective. We further encourage the eye to interpret lighter colors as higher values by placing the plot on a dark background.

The sectioned density plot presents information comparable to that presented in a kernel density plot/histogram, but does so in less space and in a way that exploits automatic perceptual processes (see Figure 5). The sectioned density plot graphs (1) the levels of the grouping variable on the x-axis, (2) the values of a continuous dependent variable on the y-axis, and (3) the relative frequency or density of the values of the dependent variable on an implied third dimension (see Figure 6). In essence, the sectioned density plot presents a distribution as if it were a building (with height corresponding to density/frequency) positioned vertically and viewed from above. Any observer who is familiar with the abstract concept of a distribution should perceive the shape of the distribution in the sectioned density plot with little training.

As with a histogram, the sectioned density plot begins by partitioning the range of the target variable (say, x) into a finite number of fixed-width intervals i. Fixed-width intervals facilitate clear and simple interpretation: the data bins are not perceptible, so varying widths would change the meaning of the graph without any visual cues. Letting $f(x_i)$ represent the relative frequency within that interval, and $f^* \max(f(x_i))$, partition the relative frequencies/density into K intervals of width f^*/K. For each interval i on x, the relative frequency/density interval

$$\delta_i = k \quad \text{if} \quad (k-1)\frac{f^*}{K} < f(x_i) \leq k\frac{f^*}{K} \quad \text{for} \quad k = \{1, 2, \ldots, K\}.$$

The values of δ_i are plotted with intensities that increase monotonically with k against values of x_i, which appear on the vertical axis. Each section of the sectioned density plot representing density interval δ_i (1) is shifted slightly to the left of, (2) occludes, and (3) is brighter than the section representing density interval δ_{i-1}. This plotting gives the appearance of depth, effectively graphing the relative frequency/density on the implied third dimension. Once again, fixed-width density intervals simplify interpretation. Varying the width of the density intervals would require varying the spacing between sections, and such variations are difficult to perceive.

The sectioned density plot has two free parameters: (1) the number of fixed-width data bins (similar to a histogram), and (2) the number of fixed-width density intervals (i.e., the number of "sections" in the implied depth dimension). The number of fixed-width data bins and density intervals will influence the appearance of the distribution. Thus, similar to a histogram, it is important to adjust these two free parameters for optimum display of the data. Figure 7 shows sectioned density plots of the same dataset for multiple choices of the number of fixed-width data bins and density intervals.

When choosing the number of data bins, it is important to consider the range of data across groups (in addition to the general considerations one makes when choosing the number of data bins for an ordinary histogram). For example, if one prefers 30 data bins per distribution, and the plot will contain two groups that have little overlap, one should set the number of data bins to 60 or 70. This relatively large number of data bins is required because only a few data bins will contain data from both distributions. In contrast, if the two distributions have a high degree of overlap, then fewer data bins will be necessary because many of the data bins will contain data from both distributions. We recommend starting with a relatively large number of data bins (e.g., 75 or more) and then adjusting this number after viewing the first graph.

The number of density intervals should not be greater than the number of data points in the largest data bin (i.e., the number of density intervals should not exceeding the resolution of the data). Because small datasets are likely to have few data points per data bin, the number of density intervals should be low (e.g., 5). Similarly, because larger datasets will likely have many data points per data bin, the number of density intervals should be greater (e.g., 10). We recommend starting with 10 density intervals and then adjusting this number after viewing the first graph.

To increase the amount of information available to the observer, we recommend that distributional information about the combined data be presented on the ordinate (see Figure 6). Here we present the median and interquartile range of the data

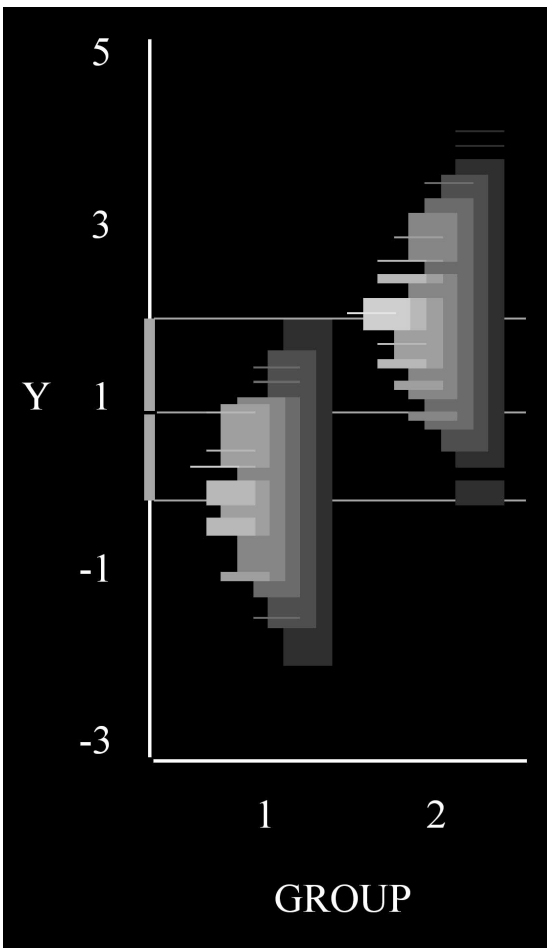

Figure 6. A sectioned density plot depicting the distributions from two groups of data. Both distributions approximate the normal and the mean of Group 2 is greater than that of Group 1. The ordinate of the graph contains graphical indicators of the locations of the 25th, 50th, and 75th percentiles of the data collapsed over group.

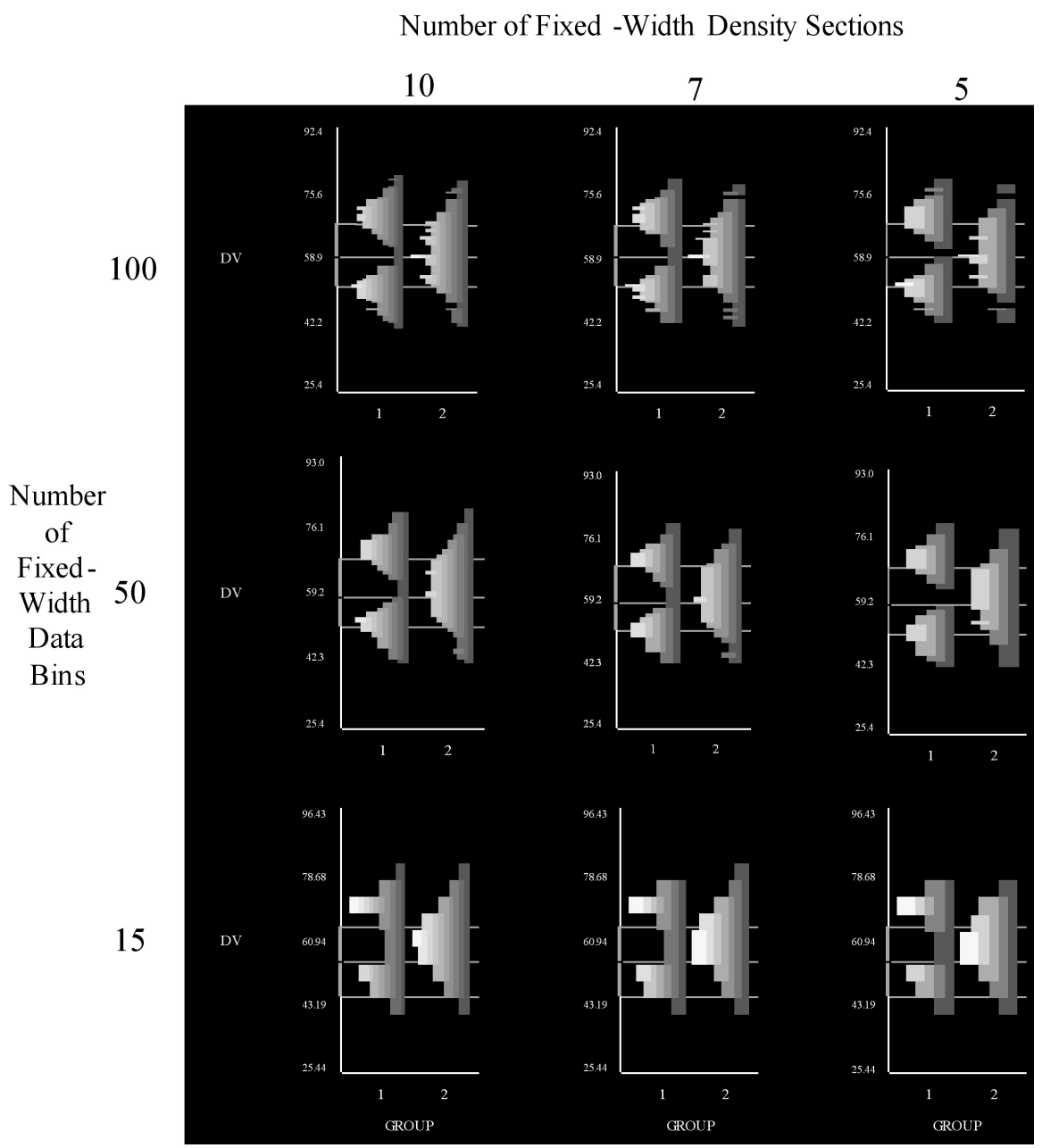

Figure 7 This figure demonstrates the effects of varying the number of fixed width data bins (along the vertical dimension) and density intervals (along the horizontal dimension) on the resulting sectioned density plots of the same data set. This dataset contains two groups with equal means. Group 1 is bimodal and Group 2 approximates the normal. Too many intervals reveals noise in the data, while too few intervals hide meaningful data.

combined across all groups (similar to a boxplot). For example, if one presents sectioned density plots comparing males and females, the ordinate will represent the combined male and female data. Three grid lines located at the median and interquartile range of the combined data provide a reference to assess distribution shift. Because there are only three gridlines, and these lines are low contrast, we minimize their likelihood of distracting the viewer. Finally, the values on the y-axis are labeled at fixed intervals as is common in most graphics. Figure 8 presents sectioned density plots representing the same distributions presented in Figure 1. In contrast to the boxplot representations in Figure 1, the viewer can readily visualize the shapes of the distributions from which the sectioned density plots are derived.

Figures 3 and 9 present the same data. Because of the structure of the sectioned density plot, many distributions can be presented side by side, similar to a boxplot. This allows for the easy comparison of distributions across groups. In addition, because the sectioned density plot presents density on an implied third dimension (z-axis), the reader can compare densities across distributions by directly comparing the heights of the distributions. This ability is enhanced by sectioning the distributions at equal

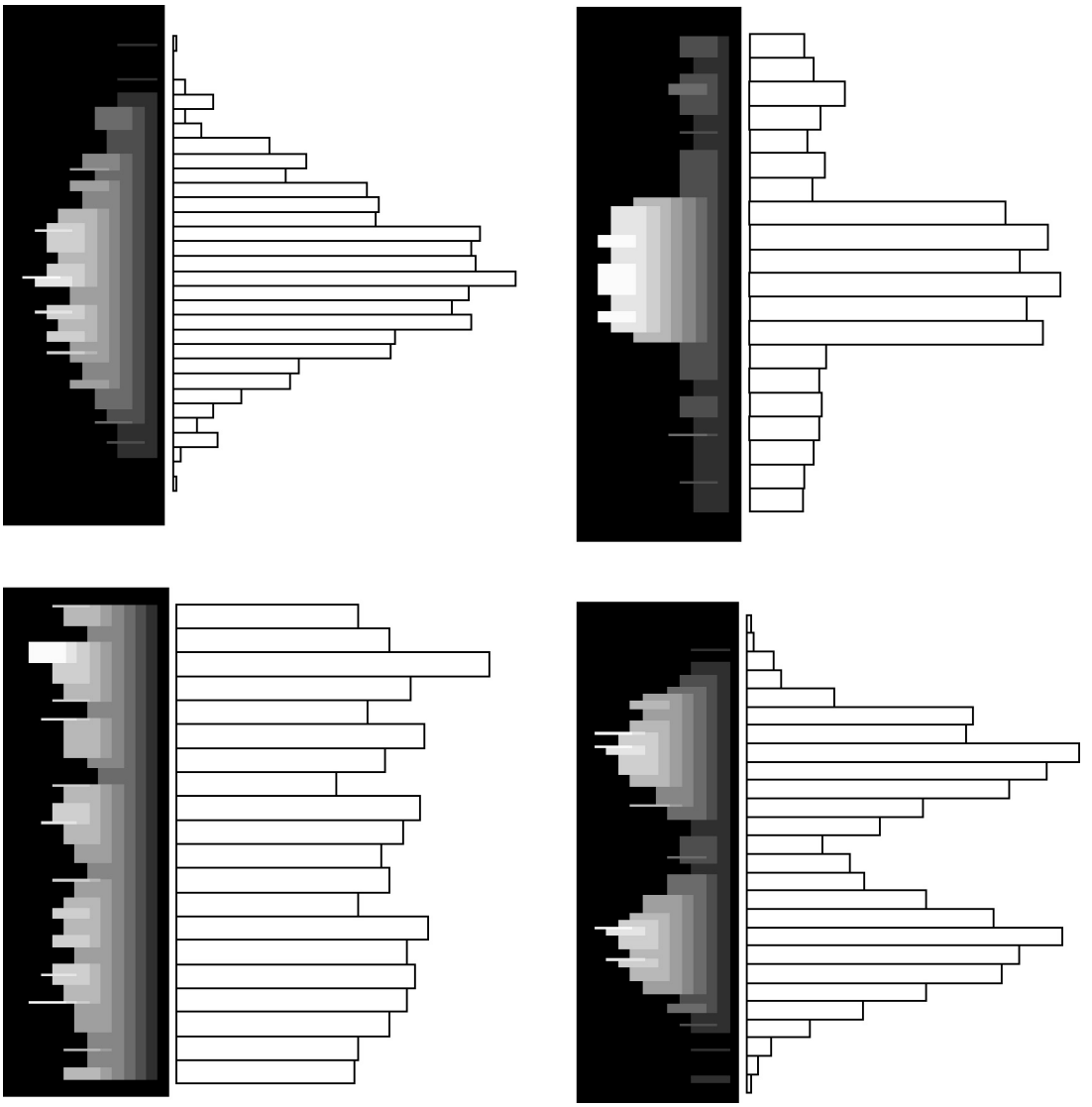

Figure 8. This figure presents the same four distributions depicted in Figure 1. Each distribution is depicted by a sectioned density plot and a histogram oriented vertically.

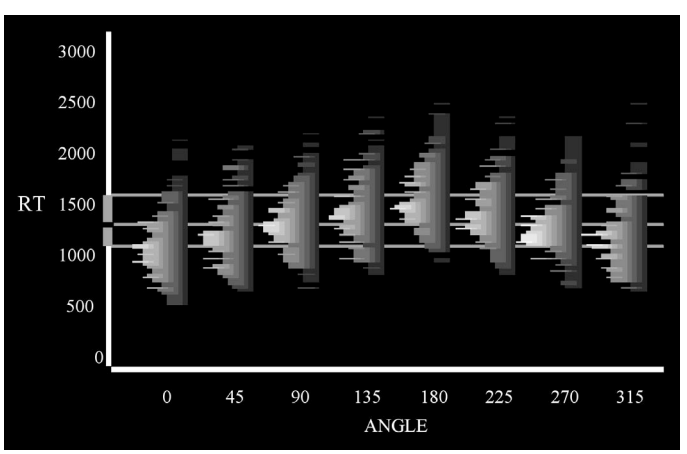

Figure 9. Sectioned density plot of same data presented in Figures 2 and 3. Notice that both the inverted "V" trend in the data and the shape of the distributions can be readily perceived.

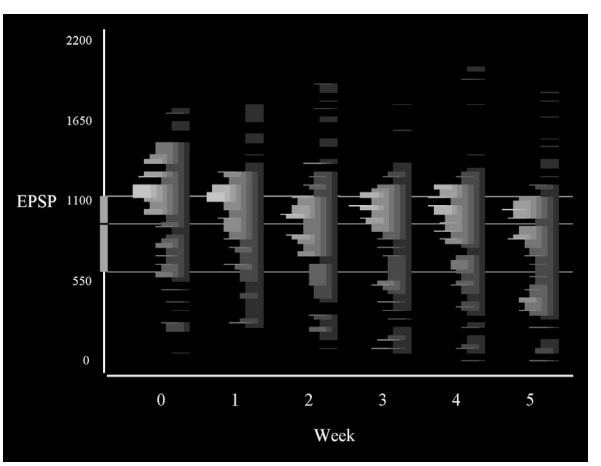

Figure 10. Sectioned density plot of electrophysiological responses (extracellular excitatory post synaptic potentials) from the dentate gyrus region of the hippocampus to perforant path stimulus pulses (200 μA) in vivo in rats over the course of several weeks. Both the trend of decreased EPSPs over time as well as the emerging bimodality should be visible.

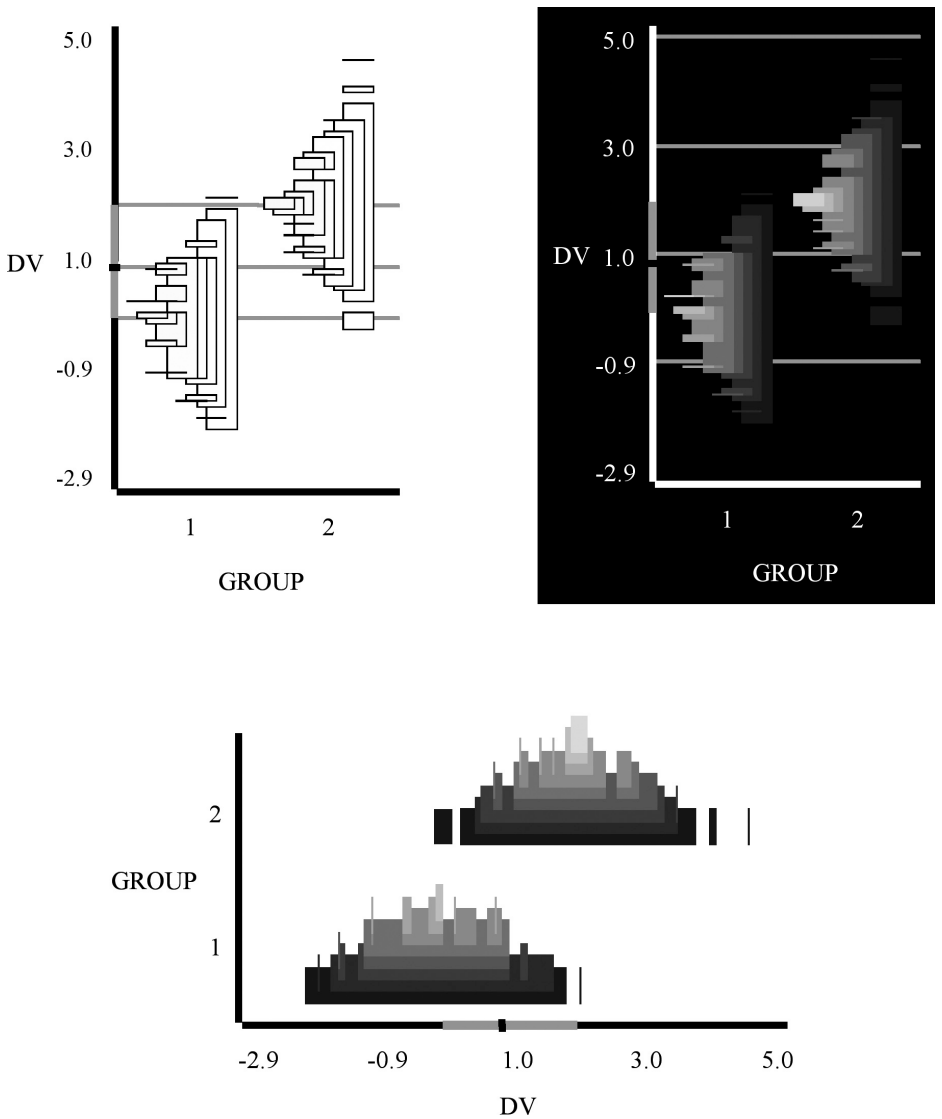

Figure 11. Various presentations of the sectioned density plot.

intervals, allowing for direct the comparison of densities across distributions. In the end, the viewer can integrate the shape, variance, and placement information across distributions with little or no effort.

Finally, Figure 10 demonstrates the advantage of being able to simultaneously visualize the shape and tendencies of distributions across multiple groups. Figure 10 presents electrophysiological responses (extracellular excitatory post synaptic potentials, EPSPs) from the dentate gyrus region of the hippocampus to perforant path stimulus pulses (200 μA) in vivo in rats. These rats were injected daily with 5 mg/kg of the antidepressant fluoxetine for several weeks. As one can see from the graph, EPSPs from the fluoxetine group decreased over time. Importantly, the graph also reveals the emergence of bimodality as weeks progress, suggesting that fluoxetine dramatically affects a subset of responses. Such an effect could only be revealed with a graphic that displays the entire distribution.

Because a key criterion for the acceptance and use of a graphical method is the ease with which researchers can use the method, we have programmed the sectioned density plot in *AM*, a freely distributed statistical software package downloadable at *http://am.air.org/*. This program contains a variety of parameters associated with the sectioned density plot that may be adjusted to meet the user's specific needs (see Figure 11). For example, the graph can be rotated so the x and y axes are switched with group presented on the y-axis and the continuous variable presented on the x-axis. In addition, one may choose to eliminate the dark background, change or remove the intensity variation across sections, add color variation, and so on. One can also change the characteristics of the x or y axes, the spacing between scales, and the spacing and number of grid lines. Finally, similar to the kernel density plot, one can plot the estimated probability density function of the population from which

the data was sampled. Thus, the sectioned density plot can be customized to fit the users' needs.

In sum, graphical presentation is central to the scientific discovery process (Smith and Prentice 1993; Tufte 1970, 1983; Tukey 1974, 1977; Tukey and Wilk 1970; Wainer and Thissen 1993). We introduce a new plot, the sectioned density plot, that combines the superior ability of a histogram/kernel density plot to present shape information with the boxplot's superior ability to present trends in variance and central tendency. The sectioned density plot is currently programmed in *AM*, a freely distributed statistical software package downloadable at *http://am.air.org/*.

[Received July 2005. Revised February 2006.]

REFERENCES

Cutting, J. E., and Vishton, P. M. (1995), "Perceiving Layout and Knowing Distances: The Integration, Relative Potency, and Contextual Use of Different Information About Depth," in *Handbook of Perception and Cognition, Vol. 5; Perception of Space and Motion*, eds. W. Epstein and S. Rogers, San Diego, CA: Academic Press, pp. 69–117.

Hintze, J. L., and Nelson, R. D. (1998), "Violin Plots: A Box Plot-Density Trace Synergism," *The American Statistician*, 52, 181–184.

Legge, G. E., Gu, Y., and Luebker, A. (1989), "Efficiency of Graphical Perception," *Perception and Psychophysics*, 46, 365–374.

Mingolla, E., and Todd, J. T. (1986), "Perception of Solid Shape From Shading," *Biological Cybernetics*, 53, 137–151.

Pastizzo, M. J., Erbacher, R. F., and Feldman, L. B. (2002), "Multidimensional Data Visualization," *Behavior Research Methods, Instruments, & Computers*, 34, 158–162.

Ramachandran, V. S. (1988), "Perception of Shape From Shading," *Nature*, 331, 163–166.

Smith, A. F., and Prentice, D. A. (1993), "Exploratory Data Analysis," in *A Handbook for Data Analysis in the Behavioral Sciences: Statistical Issues*, eds. G. Keren and C. Lewis, Hillsdale, NJ: Erlbaum, pp. 349–390.

Todd, J. T., and Mingolla, E. (1983), "Perception of Surface Curvature and Direction of Illumination From Patterns of Shading," *Journal of Experimental Psychology: Human Perception and Performance*, 9, 583–595.

Tufte, E. (1970), "Improving Data Analysis in Political Science," in *The Quantitative Analysis of Social Problems*, ed. E. Tufte, Reading, MA: Addison-Wesley, pp. 437–449.

Tufte, E. (1983), *The Visual Display of Quantitative Information*, Cheshire, CT: Graphics Press.

Tukey, J. W. (1974), "Mathematics and the Picturing of Data," in *Proceeds of the International Congress of Mathematics*, Vancouver, Canada.

——— (1977), *Exploratory Data Analysis*, Reading, MA: Addison-Wesley.

Tukey, J. W., and Wilk, M. B. (1970), "Data Analysis and Statistics: Techniques and Approaches," in *The Quantitative Analysis of Social Problems*, ed. E. Tufte, Reading, MA: Addison-Wesley, pp. 370–390).

Wainer, H., and Francolini, C. M. (1980), "An Empirical Inquiry Concerning Human Understanding of Two-Variable Maps," *The American Statistician*, 34, 81–93.

Wainer, H., and Thissen, D. (1993), "Graphical Data Analysis," in *A Handbook for Data Analysis in the Behavioral Sciences: Statistical Issues*, eds. K. Gideon and L. Charles, Hillsdale, NJ: Erlbaum, pp. 349–390.

Wainer, H., and Velleman, P. F. (2000), "Statistical Graphics: Mapping the Pathways of Science," *Annual Review of Psychology*, 52, 305–335.