# The Truthful Misrepresentation of Information in Scatterplots

"To present a useful and truthful picture, an accurate map must tell white lies," (Monmonier, pp 1). The visual display of graphical statistics also paradoxically requires misrepresentation in order for the viewer to glean crucial information which is actually present in the data. This can be illustrated in the display of scatterplots in which categorically distinct groups are presented within the same graph. Three organizational principles from vision science, by slightly misleading viewers, can actually help them efficiently perceive important information in the scatterplot by reducing the amount of time needed to process the display and directing them to important information. Specifically, the use of vernier offsets, occlusion, and the gestalt grouping principle of linearity can help viewers distinguish which observations belong to which category, and which categories have a significant linear relationship.

## Using vernier offsets and occlusion to emphasize category membership

Scatterplots become problematic when many observations have the same value on both variables. If the observations are accurately plotted, they will completely overlap, and assuming different categories use the same shape, it will be impossible to discern that there are multiple observations at a single position. A solution to this is to use small but thick line segments as the observations, which permit the use of vernier offsets (Figure 1) for overlapping observations. Vernier offsets not only eliminate the overlap problem by creating only partial rather than full overlap, but also expedite perception of small observations in the scatterplot, since the human visual system has higher vernier acuity than point acuity (Ware, 2004, pp 47-49).

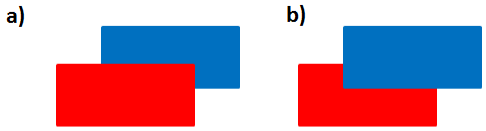


**Figure 1** Vernier offsets with chromatic abberation a) considered or b) not considered in the coloring.

If overlapping observations belong to different categories, the depth cue of occlusion can be used (Wolfe et al., 2000: pp 89-90) to depict different categories in addition to hue. Hue is often used to code for category since it is categorically rather than ordinally perceived (Cleveland, 2003, pp 229-230). When offsetting overlapping observations, however, it is important to consider which hue should be placed closer in depth. This is because different wavelengths of light naturally focus at different distances within the eye, known as *chromatic aberration* (Ware, 2003: pp 45-46). Thus, when a red and a blue color patch are placed next to each other, the red hue will naturally appear closer. By positioning observations categorically defined by red hues in front those defined by blue hues, occlusion and hue can jointly convey a

stronger impression that observations of different categories are positioned at different planes in depth (Figure 1a, cf. 1b).

## Gestalt grouping principle of linearity can encode linear significance

If the linear correlation is significant for one category but not another, gestalt grouping principles can be employed to make the scatterplot for this category "stand out", and actually seem more linear. This can be accomplished by considering the direction of the vernier offset for the significantly correlated observations. Since there are two directions in which each of the vernier stimuli can be moved, all overlapping observations of the category with a significant correlation can be moved toward regression line, whereas only half of the overlapping observations of the category with a non-significant correlation can be moved toward the regression line. This will increase the likelihood of perceiving a linear gestalt for the categories which have a significant correlation, but decrease it for the categories which do not.

## Conclusion

By offsetting observations away from their actual values, using depth to enhance categorical perception, and encouraging the idea of linearity, statistical graphics can slightly misrepresent unimportant aspects of the data in order to more easily convey those aspects which are most important.

## References

Cleveland, W. S. (1984). Elements of graphing data. *WADSWORTH ADV. BOOK PROG., 10 DAVIS DR., BELMONT, CA 94002, USA, 1984, 350*.

Monmonier, M. (1996). How to lie with maps. *Chicago and London*.

Ware, C. (2004). *Information visualization: perception for design*. Morgan Kaufmann.

Wolfe, J. M., Kluender, K. R., Levi, D. M., Bartoshuk, L. M., Herz, R. S., Klatzky, R. L., & Lederman, S. J. (2006). *Sensation & perception*. Sinauer Associates Sunderland (US); Basingstoke.