DISCUSSION ARTICLE

Comment

Michael FRIENDLY and Ernest KWAN

We consider Gelman's claims about the relative merits of tables versus graphs from a psychological perspective that emphasizes the role of data displays in the communication of quantitative results from authors to readers or viewers. From this perspective, we consider these claims in relation to a cognitive distinction between graph people and table people.

1. GRAPH PEOPLE VERSUS TABLE PEOPLE

We are grateful to Andrew Gelman for what can best be described as thought-provoking *chutzpah*, in its most positive sense. To reply in like manner, we write this in the first-person, non-royal *I*.

Thus, I will begin with a bald assertion: there are two kinds of people in this world—graph people and table people. If you sit in your local Starbucks, or even in a departmental faculty meeting and gaze around, you will have trouble at first distinguishing them by sight. But trust me—I have a Ph.D. in quantitative and cognitive psychology, so I should know what I am talking about. With a little training, you can do this, too.

Establishing this assertion scientifically is similar to what psychologists have done for over 100 years, using techniques of principal components analysis, factor analysis, and more recently, structural equation modeling, almost all of which we developed. As a result (e.g., McCrae and Costa Jr. 1987) we now know that nearly all the aspects of your personality can be summarized along five dimensions: the so-called Big 5: *Openness* to experience (appreciation for art, adventure, curiosity, ...); *Conscientiousness* (self-discipline, act dutifully, ...); *Extraversion* (positive emotions, seeking the company of others, ...); *Agreeableness* (compassionate and cooperative toward others); *Neuroticism* (experience unpleasant emotions easily, such as anger, anxiety, depression, ...).

With only a little bit of training, you will easily be able to classify Aunt Bertha, cousin Charles, your department chair, and others along such dimensions. If pressed, I will even

Michael Friendly is Professor, Psychology Department, York University, Toronto, ON, Canada (E-mail: *friendly@yorku.ca*). Ernest Kwan is Assistant Professor, Sprott School of Business, Carleton University, Ottawa, ON, Canada (E-mail: *ernest_kwan@carleton.ca*).

Journal of Computational and Graphical Statistics, Volume 20, Number 1, Pages 18–27
DOI: 10.1198/jcgs.2011.09166b

admit that my academic forefathers established all this primarily with tables, though they did use graphic methods of factor rotation extensively before the advent of analytic methods. As psychologists, we even invented a convenient acronym for you to remember this: *OCEAN*.

So, my assertion is first that there is another underlying dimension, not of personality, but rather of cognition, underlying both the *presentation* of quantitative information in tables and graphs by authors and the *understanding* of this information by readers and viewers. What works best depends most strongly on the match between the requirements of a given task on the one hand, and the skills and orientation of reader or viewer on the other.

The second part of my assertion is that these distributions, in the general population, are at least strongly bimodal, if not fully two-point, discrete distributions—graph people versus table people. Thanks to the recent important developments in Bayesian computational statistics, I do not have to address the stronger, two-point claim here, given my prior.

My initial diagnostic impressions from reading Gelman's article are recorded in my case notes: "As clear a case of graphic-denial as I've ever observed; ask about tabular-tendencies of parents and mentors; should we try the penile-erectile test with brief visual presentations of tables and graphs? What is the *real* question?"

1.1 WHAT IS THE QUESTION?

Gelman raises the question of why tables might be better than graphs (or not) with tongue firmly planted in cheek to stimulate discussion, and this is a worthy goal. As he overstates his case, this tabular-centric view invites the conclusion that almost any form of tabular presentation will suffice, as long as it is factually correct.

But in any debate, it is useful to know exactly "what is the question?" In as fine an example of the shifting-sands school of rhetoric as I have seen in a while, Gelman frames the comparison of tables and graphs in different ways, each nicely illustrated with ad hominem arguments: Graphical methods (cute toys) versus statistical modeling (serious statistics); use of graphs in the statistical literature (ignored or under-used) versus data visualization (eye-catching fluff); use of graphs in applied social science (little serious role).

Part of this debate has a long history, largely centered on the nature of the *task* (look up a precise value? make comparisons? detect trends, differences or anomalies?); see the article by Gelman, Pasarica, and Dodhia (2002, sec. 2.1) for a brief summary. Here, I just want to call attention to a brief note by Karl M. Dallenbach (1963), the editor of the *American Journal of Psychology* from 1926–1967. Publication of graphs in this journal had always been difficult and deprecated (requiring expensive "line cuts"), and Dallenbach had been largely a table person. In this note, he reports an epiphany: a long-undetected error from an earlier article caused him to "deduce from that error some evidence regarding the relative value of tables and graphs in the presentation of experimental result." From his evaluation of this case he morphed to a graph person. He concluded:

> All the evidence obtained from the reproduction of the study mentioned here indicates that the graphic method is 'better' than the tabular. Tables, since graphs are based on them, are necessary, but they are like background rocks, heavy and uninteresting. Graphs, on the other hand, spice the reports; clarify them, and make them interesting and palatable. (Dallenbach 1963, p. 702)

I describe this here to give some comfort to Gelman and other table- or crypto-table people reading this. Although the evidence on the Big 5 traits of personality suggests that they are relatively immutable over one's lifetime and may even have some genetic component, cognitive capacities are more mutable, so even a predisposition as a table person *is* subject to change.

1.2 MODES OF COMMUNICATION: WORDS, NUMBERS, PICTURES

A good deal of confusion disappears when one considers a graph or table as an act of (or attempt at) communication, similar to using words. Then, more interesting questions arise:

- What is the communication goal?
- Who is the audience?
- Was the communication effective?

Thus, a given scientific result or statistical analysis can be conveyed in different forms—words, numbers (p-values, parameter estimates, or tables), or pictures (graphs or diagrams)—for different purposes (analysis or presentation), to achieve different communication goals (exploration, detection, comparison, aesthetics, or rhetoric). Moreover, this view suggests that communication is an activity directed from a source (author) to a target (reader or viewer) and therefore the communication *mode* should be tailored to the audience in order to achieve the desired goal.

In fact, we can take this further, and consider the proposition that the majority of human communication involves, to a fair approximation, different relative proportions of the three primary ingredients: words, numbers, and pictures. Figure 1 shows some examples that help to position graphs and tables in a wider context. I freely admit that this is just a cute toy and it is not based on any data. But it does show that (in my view) tables occupy a rather lonely position, and I would have been foolish to try to present this view in a table.

Most statisticians and applied researchers know implicitly how this works. In the analysis stage, you use a collection of statistical and graphical methods both to *summarize* the data (often in tables of numbers) and to *expose* it (in graphs); you make notes (in words) regarding what you have seen and come to believe. At this stage, you are both the author and viewer, and the communication goal is the "Aha!" experience—you have found something noteworthy. In a conference presentation, you have only 20 minutes to convince your audience that what you have found is indeed interesting, so you need to focus on the "Wow!" experience. If you are smart, you will use a larger proportion of incisive visual displays, not all of the words you could have used on your slides, and tables of numbers only as necessary. Finally, you want to publish your findings, so you need to think of the communication goals and audience anew, but particularly the editors and reviewers who will decide whether you publish or perish. This goal should help decide the mix of words, numbers, and pictures in what you write and submit. Of course, through all of this is another layer: are you a graph person, or a table person? What about your audience?

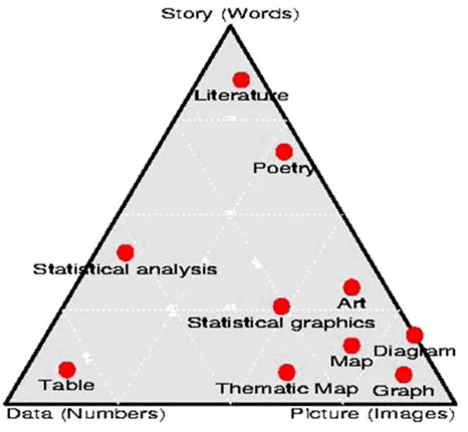

Figure 1. Modes of communication, as composed of words, numbers, and pictures, displayed in trilinear coordinates. Each point shows the (fictitious) composition of a given communication form, referred to the vertices representing 100%. The online version of this figure is in color.

2. USE OF GRAPHS AND TABLES IN SCIENTIFIC PUBLICATION

Gelman asserts that "graphs tend to be ignored or underused in much of the literature of statistics and applied fields," but this view is highly selective and ignores a growing body of research on the role of graphs (and of tables) in the construction and communication of science, as well as trends in the history of data visualization. From the previous section, it should be clear that use of graphs or tables in journal publication represents just a slice of the communication goal—intended audience tableau, but let us see where this takes us.

In a classic article Cleveland (1984) surveyed the use of graphs in 14 disciplines for the years 1980–1981, selecting 57 journals (4–5 in each area), with 50 articles selected randomly from each journal. Given that page space in journals is a limited resource, he measured the "fractional graph area" (FGA), or proportion of the total area of all journal pages devoted to graphs. Cleveland was careful in his tabulations, excluding figures such as apparatus illustrations, theoretical diagrams, etc.: "a figure was judged to be a graph if it had scales and conveyed quantitative information," so the FGA measure represented the amount of text displaced by graphs.

The results, presented in a dotplot (Cleveland 1984, fig. 3) compared journals in natural science, mathematical science, and social science. For example, the average graph use in natural science journals (chemistry: 0.18; physics: 0.17) vastly exceeded that in social science journals (economics: 0.025; sociology: 0.01).

More recently, other authors have taken up the more detailed study of the use of graphs versus tables across and within disciplines. Noteworthy here are articles by Smith, Best, and collaborators (Smith et al. 2000, 2002) where they asked respondents to rate each of the disciplines from the Cleveland study on a 1–10 scale, distinguishing "soft" science at the low end from "hard" science at the high end. As shown in Figure 2(A), use of graphs (in terms of FGA) across disciplines was nearly perfectly correlated with the rated hardness of the discipline.

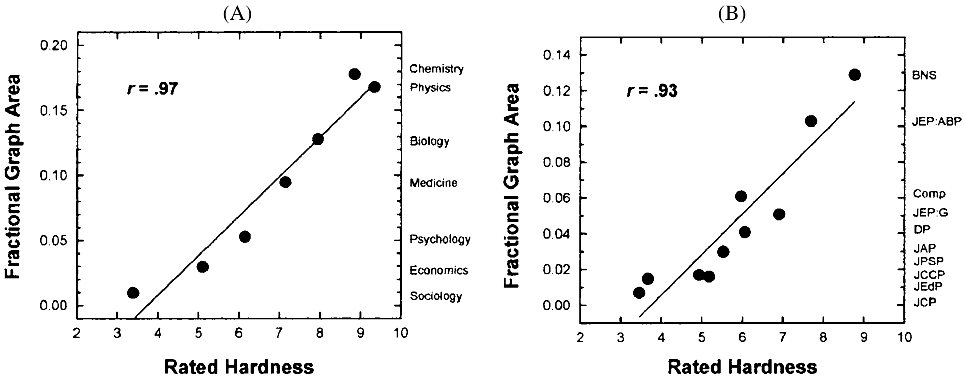

Figure 2. Proportion of journal page area devoted to graphs, in relation to rated hardness. (A) For seven scientific disciplines; (B) for 10 psychology journals. (Source: Smith et al. 2002, fig. 1.)

It should therefore not be surprising to Gelman that graph use in political science (somewhere between sociology and economics in hardness) is at the lower end of the continuum. If he is to "take a lead from our most prominent social science colleagues," he would do somewhat better to follow the exemplars set in psychology than in the dismal science of economics. As well, he (Gelman, Pasarica, and Dodhia 2002) and others in political science (Kastellec and Leoni 2007) have amply demonstrated some impressive ways in which tabular displays of even complex statistical models and model comparisons can be turned into graphic ones that preserve the essential information and make the results far more apparent.

What might be surprising is that this strong positive relation between graph use and "hardness" also applies *within* subfields of a given discipline. In psychology, journals range from the soft side (*J. Counseling Psychology*, *J. Educational Psychology*) to the harder side (*J. Experimental Psychology*, *Behavioral Neuroscience*). In a parallel study, Smith et al. (2000) obtained ratings of hardness for 10 psychology journals and also calculated graph use (FGA) for 156 articles distributed across these journals. Their results (Figure 2(B)) show nearly as strong a relation between hardness and graph use within psychology as the relation across disciplines.

The icing on this cake is shown in Figure 3, which shows the comparison between use of graphs and tables across these subfields of psychology. As rated hardness increases, area devoted to tabular displays decreases. This inverse relation is not unexpected, but the magnitude of the effect might be: in the two softest journals, the ratio of graph use to table use was about 1:10; among the hardest-rated journals, this ratio approached 10:1. It is also noteworthy that the total space devoted to data displays (tables and graphs) was more nearly constant, averaging about 14% of total page area; nevertheless, there was a smaller tendency for data display area to increase with rated hardness ($r = 0.35$).

As you will clearly see, all of this is easily explained by the cognitive distinction I introduced at the outset: The harder sciences (and harder subfields within a given area) are disproportionally populated by graph people—those whose significant phenomena are often so striking that they think it most natural to display their results in visual form. Those

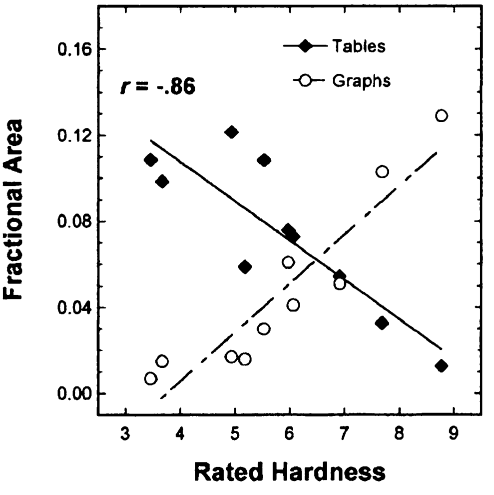

Figure 3. Proportion of page area devoted to tables and graphs in 10 psychology journals. (Source: Smith et al. 2002, fig. 3.)

who gravitate to the softer sciences often find that they have much smaller effects to deal with and so are often drawn to tabular displays, which conceal them nicely.

You can see from this description that there is a bit of a chicken-and-egg problem here: do graph people tend to gravitate to harder science, or does harder science shape researchers more as graph people? Are the data or phenomena studied in the softer sciences more complex or variable and less law-like that they resist portrayal in simple graphs? I leave these questions as open problems for future research.

2.1 WHEN TABLES FAIL

One of Gelman's major premises is that the results of 'serious' modeling are better presented in tables of coefficients ("the authors are... giving you their betas") rather than in graphs ("less form, more content"). I referred above, in passing, to the article by Kastellec and Leoni (2007) illustrating how such tables of coefficients could be rendered more cogently as graphs.

There is one more important point to be made here: In many cases either the coefficients in fitted models are meaningless without graphical display or their interpretation is exceedingly difficult to understand. For examples, consider complex generalized regression models with transformations of predictors, polynomial terms, spline functions, not to mention interactions of the above, non-identity link functions, polytomous responses, generalized additive models, etc. Such models in tabular presentation often leave me gasping for air, but I am comforted that I can understand the terms in almost any model through an *effect display* (Fox 1987, 2003)—a plot of predicted values for a term, absorbing its lower-order relatives, and averaging over other terms in the model. Maybe it is just that I am a graph person, but I know for sure that even pure table people cannot extract any sunlight from such tabular cucumbers.

3. COMBINING TABLES AND GRAPHS: SEMI-GRAPHIC DISPLAYS AND TABLEPLOTS

Of course, it does not have to be either–or: tables *versus* graphs. There is a long history to what Tukey (1972) gave the name "semi-graphic" displays, integrating exact numbers into visual displays that also showed something more, or helped understand the numbers in a wider or more coherent way. Our historical ancestors (Playfair, Guerry, Minard, Nightingale, etc.) in the development of statistical graphics were usually cognizant of the fact that graph viewers might also want to know the numbers on which the graphs were based. Sometimes, they included the numbers directly on the graphs as annotations, sometimes in separate tables, a tip-of-the-hat to the table people among their readers. Sometimes, as in the case of Mendeleyev, simple quantitative data appeared unwieldy and resisted understanding until a method to represent them in a semi-graphic display was found.

It is no accident that Francis Galton (1886) developed the ideas of regression, starting from a table of the joint frequency distribution of characteristics of parents and their offspring, and overlaying lines connecting the means of $(Y|X)$ and $(X|Y)$ as well as contours of constant frequency. The remarkable visual insights he derived from this table turned into a graph became the foundation for correlation, regression, the bivariate normal distribution, and, ultimately, a huge chunk of modern statistical methods. Karl Pearson (1920, p. 37) would later say, "…that Galton should have evolved all this… is to my mind one of the most note-worthy scientific discoveries arising from analysis of pure observation." Clearly, Galton was a graph person.

Gelman clearly knows all the deficiencies of tables from the perspective of communication, and we will not belabor these here. He is also aware of the literature on how to make tables better, either by focusing more clearly on the message they should convey (Wainer 1993, 1997) or by reformulating them as graphs (Friendly and Kwan 2003; Gelman, Pasarica, and Dodhia, 2002; Kastellec and Leoni 2007).

3.1 TABLEPLOTS

Here, I would like to introduce another idea to help bridge the gap between table people and graph people: the tableplot (Kwan 2008; Friendly and Kwan 2009), designed as a semi-graphic display in the form of a table with numeric values, but supplemented by symbols with size proportional to cell value(s), and with other visual attributes (shape, color fill, background fill, etc.) that can be used to encode other information essential to direct visual understanding.

To illustrate, consider the statistical evidence that was used to develop the Big 5 dimensions of personality and to establish their stability over an individual's lifespan and cross-cultural replicability, and how the precise predictions of such theories can be exposed by tableplots (Kwan, Lu, and Friendly 2009). As commonly operationalized, the five dimensions are measured by 240 items grouped into 30 sub-scales ("facets") of the Revised NEO Personality Inventory (NEO PI-R; Costa Jr. and McCrae 1992), with six facets measuring each of the five dimensions. Using either exploratory factor analysis followed by rotation, or confirmatory factor analysis, researchers attempt to determine the extent to which the

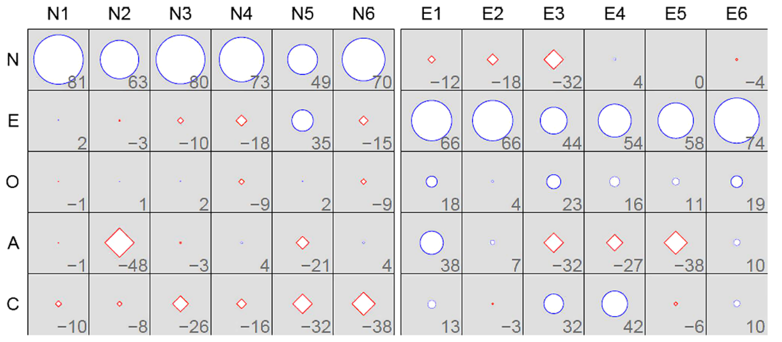

Figure 4. Tableplot of the first 12 facets from the normative NEO PI-R factor pattern measuring Neuroticism and Extraversion (Costa Jr. and McCrae 1992). Symbols (blue circles for positive loadings; red diamonds for negative loadings) scaled to |maximum| of 1; cell labels in hundredth decimal. (Source: Kwan, Lu, and Friendly 2009, fig. 2.) The online version of this figure is in color.

factor loadings conform to the five-factor theory or the extent to which two or more samples can be said to exhibit the same factor structure. Almost invariably, such results are presented in tables of factor loadings, sometimes accompanied by standard errors or other numerical comparison measures, but the model comprises 150 parameters (30 facets × 5 factors).

As an example of this graphic method, Figure 4 shows a portion of a tableplot of the facets for two factors from the normative study by Costa Jr. and McCrae (1992) to which other results are often compared. Each cell shows the factor loading (×100) as a number, and as a circle (for positive loadings) or diamond (negative loadings), scaled to have a maximum size at |loading| = 1. Confirmation of the five-factor theory requires that all target loadings approach 1.0 and non-target loadings approach 0. The tableplot display could be supplemented by other annotations to indicate p-values or the results of significant tests, but the message from this example is clear enough: The Neuroticism factor appears to be relatively well-measured by its facets, while the Extraversion factor has smaller target loadings and some possibly troubling non-target ones.

This could be just another "cute toy," unnecessary in serious statistical modeling. A strength of the tableplot, however, is that it allows an easy detailed diagnosis of the fit between predicted and estimated factor patterns or between estimated results from multiple samples. Figure 5 shows one example, in which the normative results from Costa Jr.

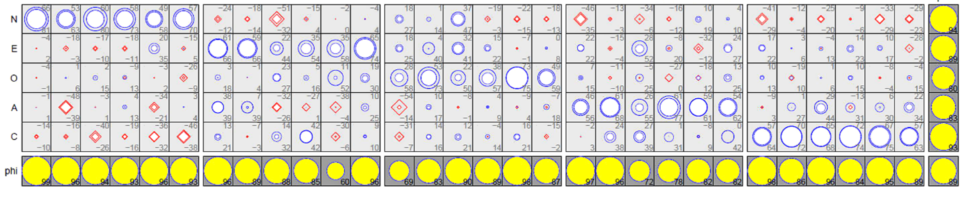

Figure 5. Superimposed tableplot of the normative (Costa Jr. and McCrae 1992) and Shona (Piedmont et al. 2002) NEO PI-R factor patterns, augmented by congruence coefficients (ϕ). Symbols scaled to |maximum| of 1; cell labels in hundredth decimal. (Source: Kwan, Lu, and Friendly 2009, fig. 4.) The online version of this figure is in color.

and McCrae (1992) are compared to those from a cross-cultural Shona-speaking sample from Zimbabwe (Piedmont et al. 2002). Such comparisons are extremely difficult in tabular displays, and so the similarities and differences are often summarized in congruence coefficients ($-1 \le \phi \le 1$) indicating the degree of similarity either within or across factors. Figure 5 shows these for the two samples on the tableplot margins. It is clear that overall, agreement is very strong, yet there are a few facets (E5, O1, A3) for which the evidence of identical factor structure is less compelling.

In case the use of tableplots or related semi-graphic displays is either unfamiliar in the context of factor analysis or unconvincing for closet table people who long to be trans-tabled, I provide examples in the supplementary materials illustrating how the results of more traditional models can be better understood and explained though tableplots than through the "background rocks, heavy and uninteresting" of tables described by Dallenbach. A **tableplot** package for R software is *http://cran.r-project.org/package=tableplot*.

4. WHY I LIKE (SOME) BAYESIAN STATISTICIANS TALKING ABOUT GRAPHS

What do Bayesian statisticians think about while they are doing an analysis and what do they think about when trying to communicate their findings to others? I do not think that those of the Bayesian or non-Bayesian persuasion are as fundamentally different as graph people versus table people in any way I could describe for your basic Starbucks test. Perhaps this is an area for future psychological research.

I would like to frame this part of the discussion in terms of this prescription for the central goal of applied statistical analysis: Tell an *A*ccurate, *C*redible, *U*nderstandable, and *I*nteresting story about some *R*eal problem (mnemonics: *ACUIR*, or *AURIC*, the gold standard).

By this standard, what Bayesian statisticians might lose on the understandability front they try to more than overcome on the side of credibility. In most cases, the point estimates from frequentist and Bayesian analyses are quite similar, or nearly identical with uninformative priors. What differs is how credibility is assessed and conveyed: by standard confidence intervals and p-values versus posterior distributions and their summaries.

Yet, perhaps because of the complexities of Bayesian modeling, particularly with complex models of significant practical importance, I find that I am often impressed with the pains that Gelman and some others (e.g., Gelman and Hill 2007) take to make their results comprehensible, not infrequently with graphic displays tuned to be responsive to the *AURIC* standard. The bottom line is this: however much Gelman professes otherwise, deep down both he and I know that he is really a closet graph person masquerading in table person garb. If he likes, we can always try the penile-erectile tests for tabular and graphic displays.

ADDITIONAL REFERENCES

Cleveland, W. S. (1984), "Graphs in Scientific Publications," *The American Statistician*, 38, 261–269. [21]

Costa Jr., P. T., and McCrae, R. R. (1992), *Revised NEO Personality Inventory (NEO PI-R) and NEO Five-Factor Inventory (NEOFFI) Professional Manual*, Odessa, FL: Psychological Assessment Resources. [24-26]

Dallenbach, K. M. (1963), "Tables vs. Graphs as Means of Presenting Experimental Results," *The American Journal of Psychology*, 76 (4), 700–702. [19]

Fox, J. (1987), "Effect Displays for Generalized Linear Models," in *Sociological Methodology, 1987*, ed. C. C. Clogg, San Francisco: Jossey-Bass, pp. 347–361. [23]

——— (2003), "Effect Displays in R for Generalised Linear Models," *Journal of Statistical Software*, 8 (15), 1–27. [23]

Friendly, M., and Kwan, E. (2003), "Effect Ordering for Data Displays," *Computational Statistics and Data Analysis*, 43 (4), 509–539. [24]

——— (2009), "Where's Waldo: Visualizing Collinearity Diagnostics," *The American Statistician*, 63, 56–65. [24]

Galton, F. (1886), "Regression Towards Mediocrity in Hereditary Stature," *Journal of the Anthropological Institute*, 15, 246–263. [24]

Gelman, A., and Hill, J. (2007), *Data Analysis Using Regression and Multilevel/Hierarchical Models*, Cambridge, U.K.: Cambridge University Press. [26]

Kwan, E. (2008), "Improving Factor Analysis in Psychology: Innovations Based on the Null Hypothesis Significance Testing Controversy," unpublished doctoral dissertation, York University, Toronto, ON. [24]

Kwan, E., Lu, I. R. R., and Friendly, M. (2009), "Tableplot: A New Tool for Assessing Precise Predictions," *Zeitschrift für Psychologie/Journal of Psychology*, 217 (1), 38–48. [24,25]

McCrae, R. R., and Costa Jr., P. T. (1987), "Validation of the Five-Factor Model of Personality Across Instruments and Observers," *Journal of Personality and Social Psychology*, 52, 81–90. [18]

Pearson, K. (1920), "Notes on the History of Correlation," *Biometrika*, 13, 25–45. [24]

Piedmont, R. L., Bain, E., McCrae, R., and Costa Jr., P. T. (2002), "The Applicability of the Five-Factor Model in a Sub-Saharan Culture: The NEO PI-R in Shona," in *The Five-Factor Model of Personality Across Cultures*, eds. R. R. McCrae and J. Allik, New York: Kluwer Academic/Plenum, pp. 155–173. [25,26]

Smith, L. D., Best, L. A., Stubbs, D. A., Archibald, A. B., and Roberson-Nay, R. (2002), "Constructing Knowledge: The Role of Graphs and Tables in Hard and Soft Psychology," *American Psychologist*, 57 (10), 749–761. [21-23]

Smith, L. D., Best, L. A., Stubbs, D. A., Johnston, J., and Archibald, A. B. (2000), "Scientific Graphs and the Hierarchy of the Sciences: A Latourian Survey of Inscription Practices," *Social Studies of Science*, 30 (1), 73–94. [21,22]

Wainer, H. (1993), "Tabular Presentation," *Chance*, 6 (3), 52–56. [24]

——— (1997), "Improving Tabular Displays, With NAEP Tables as Examples and Inspirations," *Journal of Educational and Behavioral Statistics*, 22 (1), 1–30. [24]