Ten Statisticians and Their Impacts for Psychologists

Daniel B. Wright

Florida International University

ABSTRACT—***Although psychologists frequently use statistical procedures, they are often unaware of the statisticians most associated with these procedures. Learning more about the people will aid understanding of the techniques. In this article, I present a list of 10 prominent statisticians: David Cox, Bradley Efron, Ronald Fisher, Leo Goodman, John Nelder, Jerzy Neyman, Karl Pearson, Donald Rubin, Robert Tibshirani, and John Tukey. I then discuss their key contributions and impact for psychology, as well as some aspects of their nonacademic lives.***

The disciplines of psychology and statistics have grown up as close companions; they are "inextricably bound together" (Stigler, 1992, p. 60). Although each can track their intellectual genealogies back millennia, as academic disciplines they are only about 100 years old. Early on, single individuals like Galton and Spearman could make substantive impacts in the core of both fields, but academic specialization and the development of both disciplines make this more difficult in modern times. Many psychologists like Abelson, Cohen, Cronbach, and Meehl have crossed into statistics, but their efforts have been mainly to bring statistical science into psychology. Others like Cattell, Coombs, Luce, Shepard, Stevens, Swets, Thurstone, and Tversky have focused on measurement and psychometrics. Measurement is fundamental for science (Borsboom, 2006), and psychologists have arguably done more for measurement theory than any other discipline (Hand, 2004). The people listed above would be on a list of notable psychologists who have influenced statistics.

The focus of this article is to compile a list of 10 statisticians who have done important work that psychologists should be aware of. The fact that psychologists take statistics courses throughout undergraduate and graduate training and that every empirical article includes statistics are testament to the importance of statistics within psychology. Unfortunately, statistical techniques are often taught as if they were brought down from some statistical mount only to magically appear in SPSS. This gives a false impression of statistics. Statistics is a rapidly evolving discipline with colorful characters still arguing about some of the core aspects of their discipline.

The aim of this article is to present a list of the 10 statisticians whom I think psychologists should know about. By *statistician*, I am referring to people who would likely be in a statistics department in the 21st century. I am not listing names prior to 1900, though interested readers should consult Stigler (1986). Limiting the list to 10 was difficult, and I based the list in part on personal preferences. The list is a mix of founding fathers and more recent stars, but the statistical advances had to be ones of importance to psychologists. The list is all male—however, the gender imbalance has been found in other lists of statisticians. For example, there are only four females (Gertrude Cox, F.N. David, Florence Nightingale, and Elizabeth Scott) in Johnson and Kotz's (1997) compendium of over 100 prominent statisticians. The proportions of females in the American Statistical Association and the International Statistical Institute are both increasing, so there should be less of a gender imbalance in future lists. One may also note that, in comparison with, for example, Johnson and Kotz's list, there are more British and American representatives in my list. A longer list would have included statisticians from Russia (e.g., Andrei Kolmogorov, Andrei Markov), India (e.g., Prasanta Chandra Mahalanobis, Calyampudi Radhakrishna Rao), and elsewhere. Efron (2007) describes how the center of mass for statistics is moving toward Asia.

This article has three main sections. First, three of the founding fathers of modern statistics are listed. Next, I discuss three statisticians who I have found particularly influential for my own understanding of statistics. Although this trio is based on my own preferences, their impact has, by any measure, been immense. In the third section, I describe techniques that are of importance to psychologists and list the statistician associated with that technique. It is not always easy to identify a single person to associate with each technique, but this was done to simplify the list.

Address correspondence to Daniel B. Wright, Psychology Department, Florida International University, 11200 S.W. 8th Street, Miami, FL 33199; e-mail: dwright@fiu.edu.

A JOURNAL OF THE ASSOCIATION FOR PSYCHOLOGICAL SCIENCE aps

PERSPECTIVES on Psychological Science

FOUNDING FATHERS

Founding the Discipline: Karl Pearson

The assumption during Karl Pearson's upbringing was that he would follow his father's path into law, but his interests were more broad and scientific. After finishing a mathematics degree at Cambridge, he received his legal credentials and traveled around Europe broadening his knowledge in fields as diverse as poetry and physics. During his travels, he became more interested in socialism. Although the belief that he changed the spelling of his name from Carl to Karl in homage to Karl Marx is probably incorrect (see Porter, 2004, p. 78), he did write to Marx to offer to translate Marx's *Das Kapital*. Marx declined the offer. Lenin described Pearson as "this conscientious and honest opponent of materialism" (as cited in Plackett, 1983, p. 59). When he returned to London, he formed the Men and Women's Club to discuss ideas of sexual equality. His politics led him to turn down a knighthood when it was offered in 1935. He also was a great admirer of Francis Galton and his eugenics, and Galton reciprocated by backing Pearson financially. In contemporary hindsight, it is difficult to reconcile Pearson's socialist and egalitarian ideals with Galton's eugenics.

Karl Pearson eventually settled at University College London lecturing about statistics. His writings were voluminous and, during his early years, included statistical works, plays, and political thoughts. His 1892 book *The Grammar of Science* was his first momentous achievement (K.P. Pearson, 1892/2007). The book greatly influenced Albert Einstein's theories of relativity and, as discussed later in this article, influenced Jerzy Neyman's views toward authority. In 1911, he founded the first statistics department at University College London, which became the hub of statistics. He founded and edited one of the most influential journals in the discipline (*Biometrika*; he also founded *Annals of Eugenics*, now *Annals of Human Genetics*). He laid some of the groundwork for principal component analysis and several measures of variation. He developed two of the statistical tests most used by psychologists today: Pearson's correlation and Pearson's χ^2. It is worth noting that the degrees of freedom he advised for χ^2 was in error (nk-1 for an $n \times k$ contingency table and the null of no association) and was later corrected by Fisher. Even the paper including this error has had a large impact on the discipline (Stigler, 2008). Walker (1958, p. 22) summarizes Pearson's contributions: "Few men in all the history of science have stimulated so many other people to cultivate and to enlarge the fields they had planted." One of the people most stimulated, even if at times it was to show where Pearson was wrong, was Ronald Fisher.

Mathematical Framework: Ronald Fisher

Although Karl Pearson is the founding father of the discipline, Ronald Fisher's mathematical rigor produced the scaffolding onto which the discipline was built. According to another person on my list, Bradley Efron: "Fisher should be everybody's hero because we were incredibly lucky to get a mentality of that level in our field" (Holmes, Morris, & Tibshirani, 2003, p. 275). Efron (1998, p. 95) describes how Fisher's impact on the field would have been the envy of Caesar and Alexander. In fact, Fisher published one of the most important books of science, *Statistical Methods for Research Workers* (Fisher, 1925). Psychologists will know of his work on experimental design, analyses of variance, and the test statistic, *F*, that bears his initial. Many psychologists will also be aware that he was instrumental in the development and use of *p* values as part of evaluating evidence. He had arguments about the particular meaning of *p* with other statisticians (in particularly, Neyman), but it is Fisher's use of *p* that, according to Gigerenzer (1993), is most similar to how psychologists currently do null hypothesis significance testing and is similar to how future analytic methods might be (Efron, 1998). Consider Fisher's beliefs about *p* values from his 1925 book as listed in Box 1 (adapted from Wright, 2006a).

BOX 1

Fisher's Beliefs Regarding p Values

p value	Fisher's statements
.1 to .9	"Certainly no reason to suspect the hypothesis tested" (p. 79)
.02 to .05	"Judged significant, though barely so . . . these data do not, however, demonstrate the point beyond possibility of doubt" (p. 122)
below .02	"Strongly indicated that the hypothesis fails to account for the whole of the facts" (p. 79)
below .01	"No practical importance whether p is .01 or .000001" (p. 89)

Fisher is sometimes blamed for contemporary science's fascination with a single value: .05. Although one can point to a few quotes in the 1925 book about .05 (e.g., "convenient to take this point as a limit in judging whether a deviation is to be considered significant or not. . . . We shall not often be astray if we draw a conventional line at .05," pp. 47, 79), he believed that different *p* values reflect different amounts of evidence against the null hypothesis. The value ".05" only gained prominence when researchers needed to make tables based on a few *p* values before the advent of computers and when they used *p* within a binary decision-making framework.

Within mathematical statistics, Fisher developed the notions of sufficiency, Fisher information, and most notably maximum likelihood (Stigler, 2007). Like Pearson, he was drawn to the great intellectual question of the day: genetics. His impact on genetics was also great (some argue as great as his impact on statistics), as noted by his daughter, Joan Fisher Box, who wrote a definitive biography of him (J.F. Box, 1978). Efron's (1998) look into the future of Fisher's influence on modern statistics is insightful and shows how Fisher's ideas have particular relevance to the data explosion in many scientific fields. Within psychology, this data explosion is most evident within neuroimaging.

Proving the Approach Was Right: Jerzy Neyman

Jerzy Neyman's life coincided with some of the key moments in European and North American history from the last century. From Polish descent, he spent his early years living in different parts of Russia, and in 1912 he went to the University of Kharkov. The first World War and the Russian revolutions of 1917 led to the end of Tsarist Russia and the beginning of the USSR under Vladimir Lenin. Although many of the Bolshevik education reforms were to Neyman's liking and show parallels to the educational philosophy he espoused in his later career, the conditions were harsh. For example, he was forced to trade matches for food with local farmers and was arrested for it. While in prison, he heard the man who had been in the cot next to him being executed (C. Reid, 1998).

After reading and being enamored with Karl Pearson's *The Grammar of Science*, Neyman went to London. Neyman was not impressed with the mathematical rigor of Karl Pearson, but he did meet Pearson's son, Egon, who became a close friend and collaborator. They produced the framework of null and alternative hypotheses that dominates the hypothesis testing approach in statistics. In his early career, Neyman also revolutionized sampling with stratified random sampling and created the notion of confidence intervals. In the 1930s, both Fisher and Neyman were on the faculty at University College London, and they were both critical of each other's theories. This dispute continued throughout their lives. Although there are at least two sides to any dispute, most historical accounts paint Neyman as the less mean spirited of the two. With the second World War approaching, Neyman moved to the University of California at Berkeley, a move prophesized by Fisher.[1]

At Berkeley, Neyman continued making statistical advances in different sciences like biology, astronomy, and political science (elections), in addition to his research for the military. He created the Statistics Laboratory and later the Department of Statistics, which became one of the top statistical departments in the world. He was an "ironhanded father-figure" (C. Reid, 1998, p. 216). This ironhanded aspect sometimes went too far, like when he forced Erich Lehmann to resign his editorship of *Annals of Mathematical Statistics*. This created a brouhaha in the department and the statistics community, which Lehmann said was an overreaction (Lehmann, 2008, p. 86). Usually the father-figure aspect reigned and this is evident throughout Lehmann's semi-autobiographical account as a student and then colleague of Neyman's.

Part of the Bolshevik's education reforms was the education of the working classes, who were denied access to much education in Tsarist Russia. Segregation in the 1960s United States also acted to deny educational opportunities to many people, and Neyman strived to lessen this. He was influential getting Berkeley to start scholarships for those from poor backgrounds. He felt strongly for Martin Luther King's movement and spurned many academics to contribute to King's organization. The 1960s were turbulent at Berkeley for other reasons as well. The administration were forcing faculty to sign a loyalty oath, and police were arresting students for free speech. Although Neyman signed the oath, he chaired a group that helped get legal funds for student and faculty when it became necessary. The Vietnam War had divided America, and Neyman was part of a group of professors who took out antiwar advertisements—the military eventually threatened to withdraw his funding.

Within psychology Neyman will be best remembered for the framework he and Egon Pearson developed for hypothesis testing, confidence intervals, and methods in sampling. Within statistics, he will be most remembered for applying a mathematical rigor to the questions and solutions posed by the likes of Fisher and for bringing mathematical statistics to the United States.

Summary of Founding Fathers

Limiting the list to three founding fathers is harsh. I did not chose anyone prior to Karl Pearson, but Pearson himself (E.S. Pearson, 1978) made clear how he was influenced by earlier writings. Stigler's (1986) description of pre-Pearson statistics shows the long past of statistics (with notaries like De Moivre, Gauss, and Laplace), but the impact of these people predated the start of both disciplines. I have also not included people who I do not consider primarily statisticians, like Gustav Fechner and Francis Galton. Although Fechner developed several methods for psychology, as the founder of psychophysics his place belongs in a list of physicists who psychologists should know about. He may be credited as the first psychologist to make use of statistical techniques (Stigler, 1992). Galton's profession is more difficult to pinpoint. He traveled through Africa charting areas previously unexplored by Europeans, brought measurement into meteorology, was one of the first psychometricians, and developed a theory of heredity, but he is probably most known (and disliked) for his views on eugenics (Brookes, 2004).[2] In the field of statistics, he developed correlation, regression, and associated graphical methods, and he used these within his anthropometric laboratory; however, his nonstatistical work makes him less of a statistician and more of a polymath.

MY THREE STATISTICAL HEROES

Exploring Data and Robust Estimation: John Tukey

When I mention "John Tukey," many psychologists go "yeah, I've heard of him, he did some post hoc tests that are even in

[1]When Neyman taught statistics at University College London, he questioned Fisher's approach and chose not to use Fisher's book as a textbook in his classes. Fisher told Neyman that this was not appropriate and suggested that he would be better suited somewhere far away—flippantly suggesting California (C. Reid, 1998, p. 126).

[2]Albert Gifi, Galton's loyal servant for over 40 years, had an indirect effect on statistics. Galton left his fortune to the Eugenics Society. Members of the Department of Data Theory, University of Leiden, were doing quite a bit of work on scaling and categorical data. They opted to use his name to publish as a group "as a far too late recompense for his loyalty and devotion" (Gifi, 1990, p. *x*) and to show the feelings of equality among the authors (van der Heijden & Sijtsma, 1996).

SPSS," but this does not begin to cover his accomplishments (see Wainer, 2005, pp. 117–124). He made numerous advances in many nonpsychology areas of statistics (e.g., the fast Fourier transform that is widely used in physics), public policy (e.g., as an adviser for Presidents Eisenhower, Kennedy, and Johnson and for developing a system for projecting election results with Robert Abelson in the 1960s), and was instrumental in the development of the jackknife (Tukey, 1969, p. 91), but the two areas that psychologists should most associate with him are graphing data and robust methods. Prior to Tukey, the science of displaying quantitative information was built on two premises: graphs should be used to lie about your data and your data are so boring that you have to add chartjunk to interest your audience (Tufte, 2001). As Tufte pointed out, "it is hard to imagine any doctrine more likely to stifle intellectual progress in a field" (2001, p. 53). In 1977, Tukey published *Exploratory Data Analysis* (*EDA* or the orange book) that legitimated graphing as part of academic science. This remains the definitive reference for graphs like boxplots and stem-and-leaf diagrams, which are methods for summarizing information about a variable's distribution.

In 1960, Tukey wrote a chapter for *Contributions to Probability and Statistics* in which he showed that even small deviations from normality could greatly affect the precision of estimates derived assuming the normal distribution (Tukey, 1960). This was critical, as it showed that deviations from normality too small to be detected with histograms, quantile–quantile plots, and statistical tests could still cause problems for methods that assumed normality. This was a catalyst for many statisticians to work on robust methods and Tukey, among them, made many significant contributions (Huber, 2002). These robust estimators often lessen the weight of outliers compared with the traditional least-squares approach. Figure 1 shows the

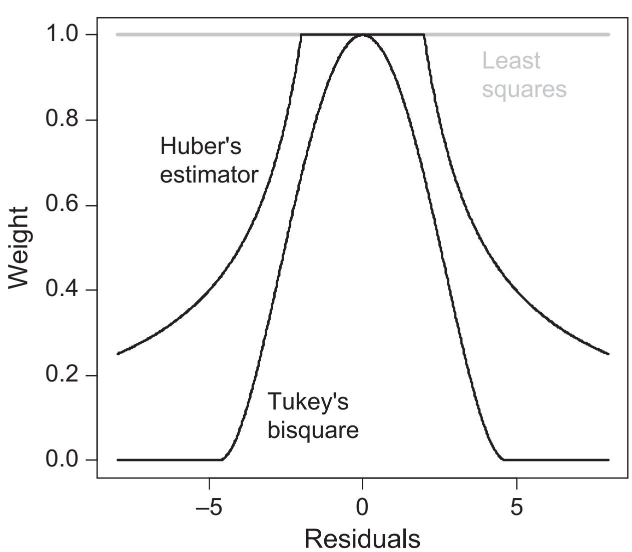

Fig 1. **The weight function for Huber's robust estimator and Tukey's bisquare (or biweight) compared with least squares estimation.**

weight function for Huber's estimator and Tukey's bisquare (or biweight) estimator. Huber's function works like least squares until the residuals reach some magnitude, and then their influence drops off. The influence of residuals for Tukey's estimator drop off immediately, and after a point they have no influence (Fox, 2002).

Graphical methods and robust statistics are rapidly growing areas of statistics and psychology. Although Tukey did not invent good scientific graphs (see Wainer's, 2005, description of Playfair, 1876/2005) and robust methods like least absolute value estimation actually predate least squares (Stigler, 1986), Tukey made these active areas of research. Both sets of methods are emphasized in the APA Task Force on Statistical Inference (Wilkinson & The Task Force on Statistical Inference, APA Board of Scientific Affairs, 1999). And finally, for Trivial Pursuit enthusiasts, Tukey coined many terms including ANOVA for analysis of variance, software, and bit.

Attributing Causation: Donald Rubin

My second hero is Donald Rubin. He arrived in 1965 at Harvard University to complete a Ph.D. in psychology only to be told that his undergraduate psychology degree at Princeton lacked enough statistics. Rubin had dabbled too much with physics as an undergraduate and, rather than take what he called a "baby" statistics course (Rubin, 2006, p. 2), he opted for a Ph.D. in statistics and has spent most of the last 20 years as Chair of the Statistics Department there. Despite this initial treatment as a graduate student, he has given back knowledge to psychology in four main ways. I will briefly describe the first three and then describe the fourth as an overarching topic.

First, many psychologists will already be aware of Rubin's work on effect sizes and meta-analysis with Robert Rosenthal (Rosenthal & Rubin, 1978). Second, Rubin is known for the EM (which stands for expectation-maximization) algorithm (Dempster, Laird, & Rubin, 1977). It is an iterative algorithm used within psychology for missing data and latent variable models. Third, Rubin proposed methods for matching groups from nonexperimental studies (see papers collected in Rubin, 2006). In quasi-experiments, it is difficult to reach causal conclusions. Analyses of covariance are sometimes used, but they can be inflexible in many circumstances and problematic when there are many covariates. Rubin and colleagues (most notably Rosenbaum, 2002) devised methods for using background variables to predict who will be in which groups and then matched people in control and treatment groups accordingly. This method, known as propensity matching or propensity scores, is popular in medical and educational statistics and is growing in popularity among psychologists.

The fourth way in which Rubin contributed to psychology is through his model for attributing causation (1974). There are several theories about causality, but Rubin's has become particularly popular. The first time I read Rubin's causal model (see

Holland, 1986, for detailed coverage),[3] I did not know whether he was saying something obvious or insightful. It was like when you first see paintings by Roy Lichtenstein and wonder if they are art. After a while, many people came to believe that Lichtenstein was brilliant, and that is how I felt after closer examination of this model (see Wright, 2006b, in relation to psychology). Rubin's model defines causality closely linked to experiments: for something to be causal, you have to be able to think about manipulating it. Then, causality can be attributed and you can estimate the causal effect in two ways.

1. You can determine how the control group, who received no treatment, would have performed if they had received the treatment, and compare this estimate with how the control group actually performed.
2. You can determine how the treatment group would have performed if they had not received the treatment, and compare this estimate with how the treatment group actually performed.

The problem is that each of these comparisons involves a hypothetical group: the control group receiving the treatment and the treatment group not receiving the treatment. Rubin's model makes explicit the need to estimate values for these unobserved groups in order to make causal attributions. This makes many of the assumptions explicit and has lead to solving many problems of causal inference, for example, Lord's Paradox (Wainer, 1991).

Statistics in the Computer Age: Brad Efron

Bradley Efron is best known for the bootstrap (Efron, 1979). This is a computer-intensive technique that allows people to assess the precision of estimates and to make confidence intervals in a general way without having to rely on mathematical formulas that often either do not exist or make unrealistic assumptions. A simple bootstrap works in the following steps. First, a sample is drawn from the original sample with replacement and is called the first bootstrap sample. Because the sample is drawn with replacement, some people's data may be included several times in this sample and some people's data may not be included at all. Next, you calculate the test statistic of interest—for example, the median. You repeat this a few thousand times and find the distribution for the median from all the bootstrap samples. To estimate the 95% confidence interval for bootstrap samples, researchers often use the middle 95% of the observed values. Efron and others have created other ways to estimate the confidence intervals, and these are included in most of the statistics packages that include bootstrapping. The beauty of the bootstrap is how it switches the emphasis from mathematical problems that have always been hard (but sometimes possible) to those that used to be impossible but are now very practical computing problems. The bootstrap and its developments are currently used for estimating parameters and constructing confidence intervals for these estimates for problems where no mathematical solutions exist, but in the future they may be used for estimation in many more problems because they perform better than the traditional mathematical solutions when the assumptions of these solutions only approximately fit the situation.

BOX 2
Efron's Interpretations of Achieved Significance Levels

Achieved significance level (ASL)	Interpretation
ASL < .10	Borderline evidence against hypothesis
ASL < .05	Reasonably strong evidence against hypothesis
ASL < .025	Strong evidence against hypothesis
ASL < .01	Very strong evidence against hypothesis

Efron's Web site (http://www-stat.stanford.edu/~brad/) includes several papers detailing his vision for the future of science and statistics. He describes how statistics has become the language of science. For several decades, he has argued that empirical Bayes methods should become more popular. In psychology, some people argue about the philosophical differences between frequentist and Bayesian methods of statistics.[4] Efron is much more pragmatic and problem-focused, arguing that empirical Bayes methods take the best bits of both approaches by taking some of the subjective decisions away from the classical Bayesian statistician and answering them with data. Efron saw Fisher as his statistical hero and in many ways they are similar. Both nestled themselves between the Bayesians and classical frequentists. Compare how Efron interprets the observed *p* value (or achieved significance level; Efron & Tibshirani, 1993, p. 204) in Box 2 with Fisher's observations in Box 1.

Since arriving at Stanford in 1960 as a graduate student, Efron has had a remarkable career. Stanford now praises the National Medal of Science award winner and former Chair of Faculty Senate and Associate Dean of Science, but it was not always that way. They expelled him in the 1960s for a parody of *Playboy*,

[3]This is sometimes called the Rubin–Holland model for Holland's excellent exposition of the theory (see Holland, 1986). If Paul Holland had been on my list, his achievements would have included work on educational testing and learning the banjo from Jerry Garcia (Wainer, 2005, p. 160).

[4]Several people have asked whether Thomas Bayes should be included for the theorem that bears his name. The problem is that it is not clear whether Bayes came up with Bayes's theorem. The theorem was published from Bayes's papers 3 years after his death. However, Stigler (1983) notes that this theorem appears in print 12 years before Bayes's death. Through meticulous detective work and applications of Bayes's theorem, Stigler concludes that the more likely developer of the theorem was Nicholas Saunderson, a blind professor of mathematics at Cambridge. Other prominent Bayesians, like economist John Maynard Keynes and the philosopher Richard Jeffrey were excluded for the list because they are nonstatisticians. Others like Bruno de Finetti, Dennis Lindley, and Leonard Savage were considered and would have been included if the list was slightly longer.

which in his own words went "a little too far" (Holmes et al., 2003, p. 296).

PARTICULARLY USEFUL TECHNIQUES AND ASSOCIATED STATISTICIANS

This next set of statisticians was chosen to highlight specific methods that are, or should be, important to psychologists. As most methods are developed by several people, I will mention more than one person for each but will focus on one person and describe some biographical information. The methods described are transformations, categorical data analysis, generalized linear models, and data reduction. In addition, the importance of statistics computing packages and some of the people associated with specific packages are discussed. Several other topics could have been chosen. If a method was discussed in relation to one of the previous statisticians it is not repeated here.

Transforming Data: David Cox

Transforming data within psychology should always be considered. For example, is it better to analyze reaction time data in seconds or in its reciprocal, speed (Hand, 2004)? Within statistics transformations are often used to make the data correspond more closely with the assumptions of the model and a simple transformation was described over 40 years ago which remains useful. George Box and David Cox (1964) proposed the following power transformation:

$$new\,X = \begin{cases} ([X+a]^b - 1)/b & \text{if } b \neq 0, \\ \ln(X+a) & \text{if } b = 0. \end{cases}$$

This has become known as the Box–Cox transformation.

Their approach was to try different values of a and b (they labeled them λ_1 and λ_2) until the model had a simple structure (i.e., no interaction terms), approximately constant variance, and residuals that were approximately normal. With modern computers, this trial and error procedure is easily automated (e.g., Venables & Ripley's, 2002, Box–Cox function in S-Plus/R).

The decision of whether to choose Box or Cox for the list is difficult because both made important contributions to statistics in other areas. George Box contributed to many areas, designing experiments (G.E.P. Box, Hunter, & Hunter, 2005), time series methods (G.E.P. Box, Jenkins, & Reinsel, 2008), and Bayesian methods (G.E.P. Box & Tiao, 1992), but for this list David Cox gets the nudge. He is best known for the proportional hazards model for analyzing survival data. This breakthrough paper (Cox, 1972) has been cited over 25,000 times. He also did much work on binary data beginning in the late 1950s that is summarized in Cox and Snell (1989). He describes how much of the motivation for this was psychology research done by his then colleagues at Birkbeck College in London (N. Reid, 1994).

Categorical Data Analysis: Leo Goodman

Psychologists are taught several techniques for working with response variables that are assumed to be continuous, but often the data from psychological research are categorical. In the typical undergraduate statistics course, psychology students are taught only the χ^2 test for categorical data (with no mention of Pearson's error described above), and they are sometimes told to do a series of χ^2 tests for multivariate examples. Although there was some work on categorical data analysis (CDA) in the first half of the last century, it was not until the middle of that century that there was a more concentrated effort (Agresti, 2002, pp. 619–631). Cox and Snell's book on binary data (1989) has already been mentioned, and Rasch's item response model (and further developments) have caught on within educational measurement (e.g., Embretson & Reise, 2000). However, to represent the strides made in CDA, Leo Goodman makes the top 10 list. In one of the best textbooks on categorical data analysis, Agresti (2002) summarizes Goodman's contribution:

> Over the past 50 years, Goodman has been the most prolific contributor to the advancement of CDA methodology. The field owes tremendous gratitude to his steady and impressive body of work. (p. 627)

Beginning in 1954, Goodman wrote a series of papers with William Kruskal (of the Kruskal–Wallis test) looking at bivariate comparisons (reprinted in Goodman & Kruskal, 1979). An example of their work together is the Goodman and Kruskal γ (gamma), which is used to measure the association between two ordinal variables.

Throughout the 1970s, Goodman produced a series of papers on log-linear models, log-multiplicative models, and latent class modeling that allow researchers to consider multiple categorical variables simultaneously (many reprinted in Goodman, 1978). Log-linear models involve estimating the log of the frequencies of a multiway contingency table with a linear model of the categories for that table. It can be viewed as a general way to perform χ^2 tests when you have more than two variables. Log-multiplicative models allow the categories to be multiplied together to predict the log of the frequencies. This relates to correspondence analysis and other scaling procedures for categorical variables. Latent class models are appropriate when you assume that categorical latent variables underlie responses to categorical observed variables. For a recent description, see Goodman (2007).

Although Goodman is known within psychology, his impact has thus far been greater in sociology. This is due in part to his substantive interests in social mobility and in part to traditional survey instruments (the bread and butter of sociology) producing more categorical data than many psychology research instruments. However, it is also due in part to less coverage of CDA methods, other than the χ^2 test, in introductory psychology statistics.

Generalized Linear Model: John Nelder

The generalized linear model is a method developed by John Nelder and colleagues (McCullagh & Nelder, 1989; Nelder & Wedderburn, 1972) for analyzing data that do not fit the normal distribution assumptions. In summarizing the value of this method, Hoffman (2004) writes, “We are most fortunate to be living in a time when the statistical tools for analyzing regression models no longer require that dependent variables follow a continuous, normal distribution” (p. viii). In psychology, the variable of interest is often a frequency, a proportion, or a dichotomy. These often follow Poisson, binomial, and Bernoulli distributions, respectively. Nelder and Wedderburn (1972) showed that models for a set of common problems (those with distributions from the exponential family) can be analyzed efficiently.[5] There are three main parts of a generalized linear model. First, there is a link function. The link function can take many forms, but the natural logarithm and the logit—$ln(x)$ and $ln(x/(1-x)$, respectively—are two of the most common. They are used to map the predicted responses onto the second part, the linear model. Finally, the analyst has to choose the distribution assumed for the residuals. This approach is called the *generalized linear model* (GLM) because it involves mapping a linear model onto a response that has been transformed via a link function. GLMs are now part and parcel of our arsenal for attacking data—the normal regression and the logistic regression are the most commonly used—and McCullagh and Nelder's (1989) book on GLMs has become a classic.

The most confusing aspect of the phrase *generalized linear model* for most psychologists is that the term is similar to *general linear model*, which usually refers to the standard normal regression that allows interaction terms and dummy variables. Nelder wished he came up with a fancier name for his approach to distinguish them better (Senn, 2003, p. 127). As an aside, in the same interview, he tells how the groundbreaking 1972 paper was rejected with no opportunity to resubmit from a statistics journal before being accepted at another journal.

John Nelder has influenced statistics both before and after the advent of GLMs. He codeveloped the Nelder–Mead simplex method (Nelder & Mead, 1965), which is a method for finding the best solution to complex problems and is very influential outside of psychology. He spent much of the first part of his career working on statistics for genetics and agriculture research and encountered many studies with confounded designs. To help analyze such messy designs, he developed the general balance algorithm in which the user describes the design and the algorithm takes into account the design's structure. He wrote the computer package GENSTAT (http://www.vsni.co.uk/products/genstat/) to encapsulate this approach. The package has reached Version 10 and is one of the best comprehensive statistics packages. Nelder was also the chairman of the group that designed the package GLIM, which was popular among statisticians but is no longer being produced. This package was based on Wedderburn and his 1972 GLM paper. In recent years, Nelder has been working on extensions to GLMs using random variables (including multilevel models), which are now incorporated into several statistical packages.

Data and Model Simplification: Robert Tibshirani

The next topic is data reduction. This is one of the largest areas of statistics and is the most difficult topic for me to single out a corresponding individual. One of the main goals of statistics is to reduce masses of complex data into simple and comprehensible patterns of information. Data reduction problems can be split into two types, often called *unsupervised learning* and *supervised learning* (Hastie, Tibshirani, & Friedman, 2001). These phrases deserve further explanation.

Unsupervised learning occurs when there is no “right” answer. The data are reduced from a high dimensionality (e.g., 20 variables) to a lower dimensionality (e.g., two components) and the value of the reduction is assessed by how much important information is maintained from the original data (often the covariance matrix or some other similarity matrix). Examples include principal component analysis, multidimensional scaling, and some latent variables models. Each of these topics is important, but they do not place an additional person onto my list. Principal component analysis is the most widely used of these techniques in statistics and would have been chosen, except that one of its originators, Karl Pearson (1901), has already been covered. After Pearson, several people including Hotelling (1933) and Eckart and Young (1936) made important advances. Multidimensional scaling was not chosen because many of the important developments were by psychologists (e.g., Shepard and Young) and the sociologist Guttman. This is not meant to downplay either the importance of the technique or of the individuals. One statistician worth noting is Joseph Kruskal, brother of the statistician who worked with Goodman (there was a third brother, Martin, who made several breakthroughs in mathematics and physics). Latent variable techniques (including factor analysis, item response modeling, and structural equation modeling) are common in psychology. The first use of this technique was in Spearman's *g* (1904) paper and massive developments were made by Thurstone (1947). Although many statisticians and psychologists have made great strides in these methods, Thurstone had the greatest impact on psychologists. So although unsupervised learning, in all of its guises, is important for psychology, it does not add a new statistician to the list.

In introductory psych-stats terminology, supervised learning occurs when there is a dependent variable on the left side of the equals sign. The data reduction can be assessed by comparing the predicted values from the model on right side of the equals sign with the dependent variable. For simple problems, this can be done with a correlation. A common way to assess the fit of

[5]Robert Wedderburn passed away when he was 28 years old, so his potential impact was never realized. Before his death, he extended generalized linear models so that they could encompass a much wider range of problems using quasi-likelihood (Wedderburn, 1974).

more complex models is to "supervise" the model while it is learning part of the data and then see how well the estimated model performs with the rest of the data. This is why the second branch of data reduction is called *supervised learning*. The typical problem involves making the model as simple as possible, and there are many advances that are particularly useful in sciences with massive amounts of data, like bioinformatics. In psychology, the common situation involves a multiple regression in which the researcher wants to eliminate variables that have little predictive value. This can be done with the stepwise methods available in popular statistics programs, but these are disliked by methodologists. Hastie et al. (2001; see Wright & London, 2009, for an introduction designed for psychologists) describe new state-of-the-art methods in this area. Each of these authors—Hastie, Tibshirani, and Friedman—has made significant advances to the field to warrant being considered for the list, but with only 10 places I limited the choice to one person.

The choice is for the lasso and Robert Tibshirani (1996). The lasso is a conceptually simple technique that is now also quickly solved thanks to Efron, Hastie, Johnstone, and Tibshirani (2004). The lasso constrains the sum of the absolute values of the standardized coefficients (the βs in a standard regression) to some value until the solution looks fairly good, as assessed by cross-validation or adjusted R^2 (for example). As the value decreases, some βs become 0 and therefore drop out, creating a more parsimonious model. Other βs become smaller. Because estimated coefficients for correlated predictor variables are often too large in magnitude, this lessens the variability of the estimates. In recent years, there have been several extensions to the basic lasso. Many of these are designed for situations in which there are many more variables than subjects (a common situation in bioinformatics).

Tibshirani has made impacts in other areas. With Trevor Hastie, he developed generalized additive models (Hastie & Tibshirani, 1990). Rather than having individual predictors included in the model in a linear fashion, a series of polynomial curves (*splines*) can be pieced together smoothly to allow more flexible relationships between predictor variables and the response variable (see Wright & London, 2009, for introduction). In addition, Tibshirani wrote the definitive monograph on the bootstrap with Bradley Efron (Efron & Tibshirani, 1993). Tibshirani is the youngest person on this list. His inclusion will be controversial and is in part due to my prediction for his future impact.

Summary and Statistical Computing Packages

There are many other topics I could have listed. One topic that deserves special mention is statistical computing packages. The computer has been integral to statistics over the past 50 years. For example, it allows computer intensive estimation, like the jackknife and bootstrap, that was not available before computers were available to make rapid calculations. Another advance is the creation of statistical packages that have allowed nonstatisticians, sometimes with minimal training, to manage statistics. This has advantages and disadvantages, but clearly the automation of techniques in statistical packages has had the largest impact on psychologists of all statistical advances in recent decades.

It would not be wise to single out any individual for statistical computing packages, but it is worth mentioning a few of the main packages and what they are best known for. I will discuss four packages (SPSS, SAS, MATLAB, and R) in chronological order.

One of the most used packages within academic psychology is SPSS (now called PASW). It began in the late 1960s (http://www.spss.com/corpinfo/history.htm) when Stanford University graduate students Norman Nie and Dale Bent and recent graduate student Hadlai Hull began writing statistical functions aimed at allowing social scientists to conduct their own statistics. Initially, this was just for researchers at Stanford, but when Nie and Hull moved to University of Chicago operations were scaled up and they began distributing the package. They initially distributed the package for free and made money by selling the manual. Since then, SPSS has continued to grow, being one of the first to release DOS and Windows versions.

The SAS package was developed soon after SPSS by Jim Goodnight and colleagues to analyze agricultural data. It has become a large company, and SAS is used across the academic and business community. Their corporate history shows numerous awards related to statistical and computing advances, but what stands out is a corporate philosophy and employee-friendly perks, such as Wednesdays being declared M&M days. They have won awards in several countries for being a good place to work (www.sas.com).

Another package that is popular in some areas of psychology is MATLAB. This was originally produced by Cleve Moler in the late 1970s to help his students perform matrix operations, but Jack Little and Steve Bangert saw its potential as commercial software for engineers (Moler, 2006). It has an advantage that people can create add-on toolboxes for particular types of analysis. For example, neuroimaging is very popular in psychology and there is a toolbox called SPM 5 (Statistical Parametric Mapping, Version 5) for MATLAB (http://www.fil.ion.ucl.ac.uk/spm/).

The final package to discuss is called R. It is recent, but its popularity is rapidly increasing. It grew out of the S-Plus system, which grew from the S language of AT&T Bell Laboratories (now Lucent Technology). S/S-Plus was written by John Chambers and colleagues (Becker, Chambers, & Wilks, 1988; Chambers, 2008; Chambers & Hastie, 1992). This is an object-oriented language for statistical programming and data analysis that won the 1998 Software System Award from the Association of Computing Machinery. This is the only statistics software to win this prestigious award (other winners include Java, Unix, and the World-Wide Web). Object orientation and the development of powerful classes of objects create a powerful statistics environment (Chambers, 2008).

R (R Development Core Team, 2009) works in a very similar way to S/S-Plus, and many of the same functions run in both

TABLE 1
Statisticians That Psychologists Should Know About

Name	Accomplishments	Year: university (adviser of doctorate)
David Cox	Box–Cox transformation, proportional hazard model	1949: University of Leeds (Henry Daniels)
Bradley Efron	Bootstrap	1964: Stanford University (Rupert Miller, Jr.)
Ronald Fisher	Analyses of variance, maximum likelihood, *p* values	1926: University of Cambridge (James Jeans)
Leo Goodman	Log-linear models, categorical data analysis	1950: Princeton University (John Tukey)
John Nelder	Generalized linear models, simplex method, GENSTAT and GLIM	1950: University of Cambridge (diploma in statistics)
Jerzy Neyman	Hypothesis testing, confidence intervals, sampling	1924: University of Warsaw (Wacław Sierpiński)
Karl Pearson	First statistics lab, correlation, chi-square	1879: University of Cambridge (Francis Galton)
Donald Rubin	Expectation-maximization algorithm, causation, matching	1971: Harvard University (William Cochran)
Robert Tibshirani	Lasso, generalized additive models, bootstrap	1985: Stanford University (Bradley Efron)
John Tukey	Exploratory data analysis, robust methods	1939: Princeton University (Solomon Lefschetz)

Note. Nelder completed a diploma rather than a doctorate. After his diploma, he received training from Frank Yates and others at Rothamsted.

packages. R was originally written by Robert Gentleman and Ross Ihaka (who both thanked John Chambers). One important difference is that R is can be freely downloaded and you can download updates whenever convenient (www.r-project.org). Further, when statisticians develop new procedures, they will often produce packages in R and store them on CRAN, the Comprehensive R Archive Network (cran.r-projects.org; a similar site exists for S-Plus, http://csan.insightful.com/). Most of the papers over the last few years in the *Journal of Statistical Software* are about R packages, including a special issue on psychometrics (de Leeuw & Mair, 2007). These all can be downloaded freely.

SUMMARY

Table 1 shows the statisticians on the list; some of the techniques associated with them; and their doctoral year, location, and adviser. Any list like this will have many near misses. For example, should Kolmogorov be included for his work on probability? Should any of the Kruskal brothers be listed? Should Bruno de Finetti or someone else be included for the Bayesian approach? Should Jöreskog be included for structural equation modeling? Should multilevel modeling and meta-analysis be listed as topics? Why were ranking procedures not included? It seems unfair to George Box, Trevor Hastie, and others not to include them when including their collaborators. Many statisticians have had profound impact on groups of psychologists through lectures, and many have written textbooks. Should these people be included? Not including quantitative psychologists will also seem a mistake too many, but given the numbers of quantitative psychologists, this group could be the focus of another article (see the lists given at the start of this article). Many issues were considered and there remains much space for heated discussion about the most important statisticians for psychologists. It is likely that any two psychologists reading through these sources and considering their own personal experiences would construct their own lists.

Another question is why a midcareer psychologist like myself should compose a list of 10 statisticians rather than, for example, a statistician or a quantitative psychologist with much more experience. There have been lists composed by statisticians (e.g., Johnson & Kotz, 1997; Lehmann, 2008), and the historian of statistics, Stephen Stigler, has provided many lucid descriptions of statisticians. There are also interviews of contemporary statisticians in the journals *Statistical Science* and *Journal of Educational and Behavioral Statistics*. These sources of were valuable for providing information about statisticians. I believe the perspective of a psychologist who is still learning about statistical procedures is probably most similar to the perspective of typical readers of *Perspectives on Psychological Science*.

Finally, the central assumption of this paper is that learning about a small group of statisticians will improve understanding of statistics. The list is meant to introduce some of the main statistical pioneers and their important achievements in psychology. Readers are encouraged to learn more about these people and others. It is hoped learning about the people behind the statistical procedures will make the procedures seem more humane than many psychologists perceive them to be.

Acknowledgments—Thanks to Susan Ayers, Andy Field, Ruth Horry, and Elin Skagerberg for comments on an earlier draft.

REFERENCES

Agresti, A. (2002). *Categorical data analysis* (2nd ed.). Hoboken NJ: Wiley.

Becker, R.A., Chambers, J.M., & Wilks, A.R. (1988). *The new S language*. London: Chapman & Hall.

Borsboom, D. (2006). The attack of the psychometricians. *Psychometrika*, *71*, 425–440.

Box, G.E.P., & Cox, D.R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society, Series B*, *26*, 211–246.

Box, G.E.P., Hunter, J.S., & Hunter, W.G. (2005). *Statistics for experimenters: Design, innovation, and discovery.* New York: Wiley.

Box, G.E.P., Jenkins, G.M., & Reinsel, G.C. (2008). *Time series analysis: Forecasting and control* (4th ed.). New York: Wiley.

Box, G.E.P., & Tiao, G.C. (1992). *Bayesian inference in statistical analysis*. New York: Wiley.

Box, J.F. (1978). *The life of a scientist*. New York: Wiley.

Brookes, M. (2004). *Extreme measures: The dark visions and bright ideas of Francis Galton*. New York: Bloomsbury Publishing.

Chambers, J.M. (2008). *Software for data analysis: Programming with R*. New York: Springer.

Chambers, J.M., & Hastie, T.J. (1992). *Statistics models in S*. London: Chapman & Hall.

Cox, D.R. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society, Series B*, *34*, 187–220.

Cox, D.R., & Snell, E.J. (1989). *Analysis of binary data* (2nd ed.). London: Chapman & Hall/CRC.

de Leeuw, J., & Mair, P. (2007). An introduction to the special volume on "Psychometrics in R." *Journal of Statistical Software*, *20*(1). Retrieved from http://www.jstatsoft.org/v20/i01

Dempster, A.P., Laird, N.M., & Rubin, D.B. (1977). Maximum likelihood from incomplete data via EM algorithm. *Journal of the Royal Statistical Society, Series B*, *39*, 1–38.

Eckhart, C., & Young, G. (1936). The approximation of one matrix by another of lower rank. *Psychometrika*, *1*, 211–218.

Efron, B. (1979). Bootstrap methods: Another look at the jackknife. *Annals of Statistics*, *7*, 1–26.

Efron, B. (1998). R.A. Fisher in the 21st century. Invited paper presented at the 1996 R.A Fisher lecture. *Statistical Science*, *13*, 95–114.

Efron, B. (2007). *The future of statistics*. Retrieved from http://stat.stanford.edu/ ~ brad/talks/future.pdf

Efron, B., Hastie, T., Johnstone, I., & Tibshirani, R. (2004). Least angle regression. *Annals of Statistics*, *32*, 407–499.

Efron, B., & Tibshirani, R. (1993). *An introduction to the bootstrap*. New York: Chapman & Hall.

Embretson, S.E., & Reise, S.P. (2000). *Item response theory for psychologists*. Mahwah, NJ: Erlbaum.

Fisher, R.A. (1925). *Statistical methods for research workers*. London: Oliver & Boyd.

Fox, J. (2002). *An R and S-Plus companion to applied regression*. Thousand Oaks, CA: Sage.

Gifi, A. (1990). *Nonlinear multivariate analysis*. Chichester, United Kingdom: Wiley.

Gigerenzer, G. (1993). The superego, the ego and the id in statistical reasoning. In G. Keren & C. Lewis (Eds.), *A handbook for data analysis in the behavioral sciences: Methodological issues* (pp. 311–339). Hove, United Kingdom: Erlbaum.

Goodman, L.A. (1978). *Analyzing qualitative/categorical data*. Cambridge, MA: Abt Books.

Goodman, L.A. (2007). Statistical magic and/or statistical serendipity: An age of progress in the analysis of categorical data. *Annual Review of Sociology*, *33*, 1–19.

Goodman, L.A., & Kruskal, W.H. (1979). *Measures of association for cross classifications*. New York: Springer-Verlag.

Hand, D.J. (2004). *Measurement theory and practice: The world through quantification*. London: Edward Arnold.

Hastie, T., & Tibshirani, R. (1990). *Generalized additive models*. New York: Chapman & Hall.

Hastie, T., Tibshirani, R., & Friedman, J. (2001). *The elements of statistical learning: Data mining, inference, and prediction*. New York: Springer.

Hoffmann, J.P. (2004). *Generalized linear models: An applied approach*. Boston: Pearson Education.

Holland, P.W. (1986). Statistics and causal inference. *Journal of the American Statistical Association*, *81*, 945–960.

Holmes, S., Morris, C., & Tibshirani, R. (2003). Bradley Efron: A conversation with good friends. *Statistical Science*, *18*, 268–281.

Hotelling, H. (1933). Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*, *24*, 417–441.

Huber, P.J. (2002). John W. Tukey's contributions to robust statistics. *Annals of Statistics*, *30*, 1640–1648.

Johnson, N.L., & Kotz, S. (Eds.). (1997). *Leading personalities in statistical sciences*. New York: Wiley.

Lehmann, E.L. (2008). *Reminiscences of a statistician: The company I kept*. New York: Springer.

McCullagh, P., & Nelder, J.A. (1989). *Generalized linear models* (2nd ed.). London: Chapman & Hall.

Moler, C. (2006, January). *The growth of MATLAB and The MathWorks over two decades*. Retrieved from www.mathworks.com/company/newsletters/news_notes/clevescorner/jan06.pdf

Nelder, J.A., & Mead, R. (1965). A simplex-method for function minimization. *Computer Journal*, *7*, 308–313.

Nelder, J.A., & Wedderburn, R.W.M. (1972). Generalized linear models. *Journal of the Royal Statistical Society Series A*, *135*, 370–384.

Pearson, E.S. (Ed.). (1978). *The history of statistics in the 17th and 18th centuries against the changing background of intellectual, scientific and religious thought. Lectures by Karl Pearson given at University College London during the academic sessions 1921–1933*. New York: Macmillan.

Pearson, K.P. (1901). On lines and planes of closest fit to a system of points in space. *Philosophical Magazine (6th series)*, *2*, 557–572.

Pearson, K.P. (2007). *The grammar of science*. New York: Cosimo. (Original work published 1892)

Plackett, R.L. (1983). Karl Pearson and the chi-squared test. *International Statistical Review*, *51*, 59–72.

Playfair, W. (2005). *The commercial and political atlas and statistical breviary*. New York: Cambridge University Press. (Original work published 1876)

Porter, T.M. (2004). *Karl Pearson: The scientific life in a statistical age*. Princeton, NJ: Princeton University Press.

R Development Core Team (2009). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Reid, C. (1998). *Neyman*. New York: Springer-Verlag.

Reid, N. (1994). A conversation with Sir David Cox. *Statistical Science*, *9*, 439–455.

Rosenbaum, P.R. (2002). *Observational studies* (2nd ed.). New York: Springer.

Rosenthal, R., & Rubin, D.B. (1978). Interpersonal expectancy effects: The first 345 studies. *Behavioral and Brain Sciences*, *1*, 377–386.

Rubin, D.B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, *66*, 688–701.

Rubin, D.B. (2006). *Matched sampling for causal effects*. Cambridge, United Kingdom: Cambridge University Press.

Senn, S. (2003). A conversation with John Nelder. *Statistical Science*, *18*, 118–131.

Spearman, C. (1904). "General Intelligence": Objectively determined and measured. *American Journal of Psychology*, *15*, 201–293.

Stigler, S.M. (1983). Who discovered Bayes's Theorem? *American Statistician*, *37*, 290–296.

Stigler, S.M. (1986). *The history of statistics: The measurement of uncertainty before 1900*. Cambridge, MA: Harvard University Press.

Stigler, S.M. (1992). A historical view of statistical concepts in psychology and educational research. *American Journal of Education*, *101*, 60–70.

Stigler, S.M. (2007). The epic story of maximum likelihood. *Statistical Science*, *22*, 598–620.

Stigler, S.M. (2008). Karl Pearson's theoretical errors and the advances they inspired. *Statistical Science*, *23*, 261–271.

Thurstone, L.L. (1947). *Multiple factor analysis*. Chicago: Chicago University Press.

Tibshirani, R. (1996). Regression shrinkage and selection via the Lasso. *Journal of the Royal Statistical Society, Series B*, *58*, 267–288.

Tufte, E.R. (2001). *The visual display of quantitative information* (2nd ed.). Cheshire, CT: Graphics Press.

Tukey, J.W. (1960). A survey of sampling from contaminated normal distributions. In I. Olkin, S. Ghurye, W. Hoeffding, W. Madow, & H. Mann (Eds.), *Contributions to probability and statistics* (pp. 448–485). Stanford, CA: Stanford University Press.

Tukey, J.W. (1969). Analyzing data: Sanctification or detective work? *American Psychologist*, *24*, 83–91.

Tukey, J.W. (1977). *Exploratory data analysis*. Reading, MA: Addison-Wesley.

van der Heijden, P.G.M., & Sijtsma, K. (1996). Fifty years of measurement and scaling in the Dutch social sciences. *Statsitica Neerlandica*, *50*, 111–135.

Venables, W.N., & Ripley, B.D. (2002). *Modern applied statistics with S/S-Plus* (4th ed.). New York: Springer.

Wainer, H. (1991). Adjusting for differential base rates: Lord's paradox again. *Psychological Bulletin*, *109*, 147–151.

Wainer, H. (2005). *Graphic discovering: A trout in the milk and other visual adventures*. Princeton, NJ: Princeton University Press.

Walker, H.M. (1958). The contributions of Karl Pearson. *Journal of the American Statistical Association*, *58*, 11–22.

Wedderburn, R.W.M. (1974). Quasi-likelihood functions, generalized linear models, and the Gauss-Newton method. *Biometrika*, *61*, 439–447.

Wilkinson, L., & The Task Force on Statistical Inference, APA Board of Scientific Affairs. (1999). Statistical methods in psychology journals: Guidelines and explanations. *American Psychologist*, *54*, 594–604.

Wright, D.B. (2006a). The art of statistics: A survey of modern statistics. In P.A. Alexander & P.H. Winne (Eds.), *Handbook of educational psychology* (2nd ed., pp. 879–901). Mahwah, NJ: Erlbaum.

Wright, D.B. (2006b). Causal and associative hypotheses in psychology: Examples from eyewitness testimony research. *Psychology, Public Policy, and Law*, *12*, 190–213.

Wright, D.B., & London, K. (2009). *Modern regression techniques in R*. London: Sage.