

## A note on the growth of the use of statistical tests in *Perception & Psychophysics*

SCOTT PARKER

The American University, Washington, D.C.

Articles in *Perception & Psychophysics* that include no statistical tests have become rarer over the journal's first 24 years of publication. This change does not seem to represent an obvious improvement in the journal's content, but it may be a symptom of an unfortunate trend in the teaching and conduct of research.

There has been, over the years, a pronounced increase in the proportion of articles in the Psychonomic Society journal *Perception & Psychophysics* that use statistical significance tests in data analysis. I scanned the entirety of the journal from Volume 1 through Volume 46 (1966 through 1989). Ignoring those that did not have data collected by the authors (theoretical articles, reviews, biographical notes, discussions of other people's work, etc.), I counted the number of articles that did and did not report significance tests. An article was counted as reporting a significance test even if the only mention was, for example, a suggestion in passing that the separation between open and closed circles in a figure was not significant. I computed the percentage of articles in each volume that reported at least one significance test. I then reread Volumes 1-10 as a reliability check, computing the percentage of articles reporting tests in the second reading. I computed the 10 differences between percentages based on the first and second readings. Of those 10 differences, 1 was 8%, 8 were 3% or less, and 1 was 0.

The percentage of articles reporting significance tests in each volume is shown in Figure 1. The abscissa is year of publication; each data point represents a volume. Notice that there was only one volume in each of 1966 and 1967 and that there have been two volumes per year since then. For the data of Volumes 1-10 (the ones I recounted), the higher of my two counts is shown. Notice that the percentage of articles containing significance tests appears to increase over time, being around 65 to 70 in the first few volumes and around 85 recently. From 1966 to June 1978 (Volumes 1-23), in only five volumes did the percentage with tests reach 80, whereas the percentage exceeded 80 in all but five of Volumes 24-46.

For the modern reader, the value of Kendall's tau between percentage and year is 0.56 ( $p < .0001$ ) and the value of the Pearson product-moment correlation coefficient

is .77 ( $p < .0001$ ). (The correlations between percentage and volume differ from those between percentage and year only in the third decimal place.)

We may compare this journal's performance with that of the *Journal of Abnormal Psychology*. Sedlmeier and Gigerenzer (1989) report that the 1984 volume of that journal contained 56 articles, of which 55 included significance tests and the other discussed previous research. Cohen (1962) analyzed the *Journal of Abnormal and Social Psychology* for 1960 and found that 70 of the 78 articles contained statistical tests. (The total number of statistical tests reported in those 70 articles was 4,829.) This suggests that *Perception & Psychophysics* is catching up with some other areas of psychology. (A wonderful history of hypothesis testing and its adoption by psychology is given in chapter 1 of Gigerenzer & Murray, 1987.)

One may wonder why the change is occurring. Surely the rapid increase in the availability of computational power and statistical software has enabled researchers to

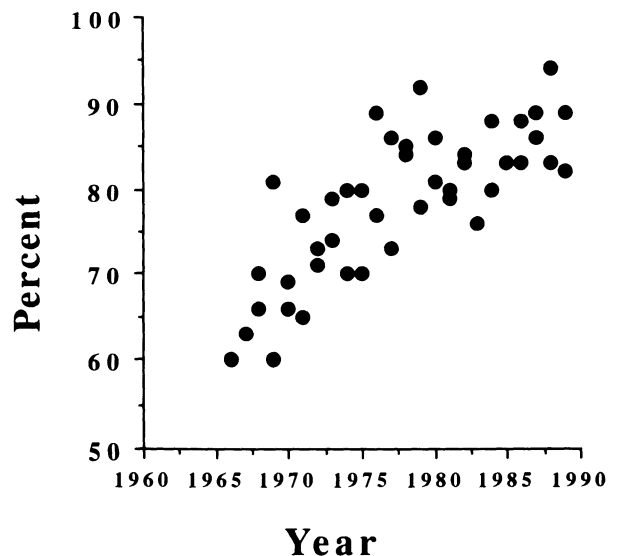


Figure 1. Percentage of articles reporting the performing of statistical tests in each volume (one volume in 1966 and 1967; two volumes per year thereafter) by year.

Thanks are extended to Bruce Schneider and Debra Zellner for helpful comments on earlier versions and to Tracy Aquino and Bryan Fantie for making the figure. Correspondence should be sent to Scott Parker, Department of Psychology, The American University, Washington, DC 20016.

do easily and rapidly statistical tests that were once difficult and time-consuming. So perhaps the rise in the rate of statistical testing is an overdue improvement. However, during my reading of the journal, I was not struck by the need for statistical tests in the early volumes. The early volumes' reports did not seem less competent or less informative than those published later. So I doubt that the change is simply a matter of tightening up standards or filling obvious needs.

I suspect rather that the change should be seen in the context of a parallel change in the training of experimental psychologists. The way students in psychology are trained in research now emphasizes statistical testing more than it used to. The old experimental psychology texts (Andreas, 1960, 1972; Osgood, 1953; Woodworth & Schlosberg, 1954) did not have chapters devoted to methodology and data analysis. More current texts (e.g., Calfee, 1975; d'Amato, 1970; Snodgrass, Levy-Berger, & Haydon, 1985) commonly do. Some (e.g., Christensen, 1988; McBurney, 1990) are devoted exclusively to methodological discussions. Methodology is now accorded a status at least equal in importance to that of research content. The old textbooks introduced experiments that were topically interesting without pointedly identifying them as examples of experimental arrangements that students might find useful. Recently, though, textbooks have come to treat experiments primarily as examples of particular experimental arrangements rather than as the means by which something of interest was learned. Inventorying arrangements, rather than surveying an area of study, has become the organizing principle for the introductory research course. Discussion of method and research design has taken on a life of its own, and its content is largely statistical. This gets people's thoughts "channeled into the routine of statistics" (Bolles, 1988, p. 79). It gets people to think of experimentation as an enterprise that involves statistical testing as a natural ingredient.

Perhaps we should blame Sir Ronald Fisher (1935) for using the phrase *the design of experiments* to designate the analysis of variance and its variants. The insinuation is that the supply of statistical analyses defines the supply of well-designed experiments. So the books on "experimental design" read by psychology graduate students tend to be largely texts on statistics (e.g., Edwards, 1985; Keppel, 1982; Kirk, 1982; Myers, 1979; Winer, 1971). All the training we give our students in statistics and in methodology conspires not only to make experimenters likely to perform statistical tests but also to make the journal's ad hoc reviewers likely to demand them. Further-

more, as research funds become scarcer, experimenters may increasingly prefer well-established research strategies that provide standardized ways, specifiable in advance, to draw defensible conclusions from experimental outcomes.

Some would say that the increased use of statistical tests is not beneficial but, rather, that it reveals the growth of a bad habit (see, e.g., Bolles, 1988). At the least, it seems that in this respect the research and analysis reported in *Perception & Psychophysics* have become more uniform. That need not be a bad thing, but in this instance it probably is.

## REFERENCES

- ANDREAS, B. G. (1960). *Experimental psychology*. New York: Wiley.
- ANDREAS, B. G. (1972). *Experimental psychology* (rev. ed.). New York: Wiley.
- BOLLES, R. C. (1988). Why you should avoid statistics. *Biological Psychiatry*, **23**, 79-85.
- CALFEE, R. C. (1975). *Human experimental psychology*. New York: Holt, Rinehart & Winston.
- CHRISTENSEN, L. B. (1988). *Experimental methodology* (4th ed.). New York: Simon & Schuster.
- COHEN, J. (1962). The statistical power of abnormal-social psychological research: A review. *Journal of Abnormal & Social Psychology*, **65**, 145-153.
- D'AMATO, M. R. (1970). *Experimental psychology: Methodology, psychophysics and learning*. New York: McGraw-Hill.
- EDWARDS, A. L. (1985). *Experimental design in psychological research* (5th ed.). New York: Harper.
- FISHER, R. A. (1935). *The design of experiments*. Edinburgh: Oliver & Boyd.
- GIGERENZER, G., & MURRAY, D. J. (1987). *Cognition as intuitive statistics*. Hillsdale, NJ: Erlbaum.
- KEPPEL, G. (1982). *Design & analysis* (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall.
- KIRK, R. E. (1982). *Experimental design* (2nd ed.). Monterey, CA: Brooks/Cole.
- MCBURNAY, D. H. (1990). *Experimental psychology* (2nd ed.). Monterey, CA: Wadsworth.
- MYERS, J. L. (1979). *Fundamentals of experimental design* (3rd ed.). Boston: Allyn & Bacon.
- OSGOOD, C. (1953). *Method and theory in experimental psychology*. New York: Oxford University Press.
- SEDLMEIER, P., & GIGERENZER, G. (1989). Do studies of statistical power have an effect on the power of studies? *Psychological Bulletin*, **105**, 309-316.
- SNODGRASS, J. G., LEVY-BERGER, G., & HAYDON, M. (1985). *Human experimental psychology*. New York: Oxford University Press.
- WINER, B. J. (1971). *Statistical principles in experimental design* (2nd ed.). New York: McGraw-Hill.
- WOODWORTH, R. S., & SCHLOSBERG, H. (1954). *Experimental psychology* (rev. ed.). New York: Holt, Rinehart & Winston.