CC/NUMBER 34
AUGUST 23, 1982

# This Week's Citation Classic

**Cohen J.** Multiple regression as a general data-analytic system.
*Psychol. Bull.* 70:426-43, 1968.
[New York University, New York, NY]

**Techniques are presented for using multiple regression and correlation analysis as a versatile and powerful data-analytic procedure that generalizes the analysis of variance and covariance. These include the representation of nominal scales and methods for handling interactions, curvilinearity, missing data, and covariates. [The *Science Citation Index*® (*SCI*®) and the *Social Sciences Citation Index*® (*SSCI*®) indicate that this paper has been cited in over 480 publications since 1968.]**

Jacob Cohen
Department of Psychology
New York University
New York, NY 10003

June 8, 1982

"Having flunked calculus in my freshman year at City College of New York, my vocational goal of teaching high school math (as well as further math training) was abandoned. Twelve years and one world war later, I was a full-fledged clinical psychologist PhD (1950) and all that remained of my earlier aspirations was a fascination with quantitative research methods. I drifted into teaching and research consulting in statistics and psychometric theory, and fiddling with methodological problems in the social and behavioral sciences. Without formal mathematical training, all I could do was make intuitive leaps and check them out numerically. Proofs were (and still largely are) beyond me.

"Multiple regression/correlation analysis (MRC), I had been taught, was a method for relating a group of variables to a single variable, and was largely used in psychotechnological applications, e.g., for predicting college grade point average from entrance examination scores; the resulting regression equation was used as a selection tool to aid in college admission decisions. My intuition and idea fragments wrested from barely comprehended technical publications led me to see MRC as potentially of far wider applicability. During the early-1960s, stumbling along, I constructed an MRC-based general data-analytic system which encompassed the analysis of variance and covariance as special cases. I had 'discovered' the general linear model. (An advantage of lack of training is the excitement of making such 'discoveries.') I prepared the article and sent it to some experts in the field for review. Their responses were detailed and conscientious, but not encouraging. They found the exposition 'fuzzy' and 'nonrigorous,' and objected most strongly to those of its aspects of which I was most proud. Since I did not understand all they wrote, I could not be sure, but as far as I could tell, they were not challenging the validity of the MRC system I presented, but rather its style, which certainly differed from theirs. With some trepidation, I nevertheless sent the manuscript off to *Psychological Bulletin* which, luckily, accepted it.

"I believe that part of the reason it has been so highly cited lies at the heart of the difference between the experts and me. The article was addressed to an audience of researchers and provided them with methods they could intelligently apply. What was 'fuzzy' and 'nonrigorous' in the mathematical-statistical framework of the experts was conceptually clear and intuitively accessible to behavioral scientists. There is an obvious object lesson here in scientific communication.

"The rash of reprint requests encouraged me to write (with a newly acquired wife) a textbook that expanded on the system[1] which has also been successful; we are currently winding up a second edition which features a causal models outlook.

"Unregenerate, I have recently completed a multivariate generalization of MRC called set correlation.[2] It had a similar reception from some of the experts who saw preprints. (Indeed, it was rejected by *Psychological Bulletin*.) This gives me reason to hope that some day it too will be a *Citation Classic*!"

1. **Cohen J & Cohen P.** *Applied multiple regression/correlation analysis for the behavioral sciences.* Hillsdale, NJ: Lawrence Erlbaum Associates, 1975. 490 p.
[The *SSCI* indicates that this book has been cited in over 535 publications since 1975.]
2. **Cohen J.** Set correlation as a general multivariate data-analytic method. *Multivariate Behav. Res.* 17:301-42, 1982.