The work of Pitman (1937, 1939) showed clearly that the practice of analysis of variance need not depend on Gaussian (normal) distributions, that it made sense in much broader circumstances than those "close to Gaussian." In my experience, most applications of the practical analysis of variance are exploratory in nature, and deserve an inherently flexible approach. Informally, those who have made, or handled, many such analyses have learned to include flexibility in such matters as how many interactions to include—and when to combine two (or more) factors into a composite factor and when to leave them separate.

The formalization of such flexibility has lagged. Green and Tukey (1960) illustrated an approach that has not been widely followed. Johnson and Tukey (1987) have now taken this flexibility several steps further. (In his nearly completed Ph.D. thesis, Johnson is taking still further steps.)

There need be no conflict between Speed's important improvements in the mathematical description of a narrower process and the clarification and exposition of a broader one. I trust there will be none.

REFERENCES

GREEN, B. F., JR. and TUKEY, J. W. (1960). Complex analyses of variance—general problems. *Psychometrika* **25** 127–152.

JOHNSON, E. J. and TUKEY, J. W. (1987). Graphical exploratory analysis of variance illustrated on a splitting of the Johnson and Tsao data. In *Design, Data and Analysis by Some Friends of Cuthbert Daniel.* Wiley, New York. To appear.

PITMAN, E. J. G. (1937). Significance tests which may be applied to samples from any populations. *J. Roy. Statist. Soc. Suppl.* **4** 119–130.

PITMAN, E. J. G. (1939). Significance tests which may be applied to samples from any populations. III. The analysis of variance test. *Biometrika* **29** 322–335.

PRINCETON UNIVERSITY
FINE HALL
WASHINGTON ROAD
PRINCETON, NEW JERSEY 08544

REJOINDER

T. P. SPEED

CSIRO, Canberra, Australia

> ...I maintain that any writer of a book is fully authorised in attaching any meaning he likes to any word or phrase he intends to use. If I find an author saying, at the beginning of his book, "Let it be understood that by the word '*black*' I shall always mean '*white*,' and that by the word '*white*' I shall always mean '*black*,'" I meekly accept his ruling, however injudicious I may think it.
>
> Lewis Carroll [cited in Gardner (1960)]

In writing about models for the dispersion matrices of arrays of random variables, which are defined by equality constraints on the entries, and calling it anova, I tried to emphasise a unity between parts of theoretical statistics which I felt was not immediately apparent. Of course, the very wide variety of sums of squares decompositions, orthogonal decompositions of arrays of numbers and robust/resistant analogues of the latter such as median polish, are of enormous value in applied statistics and data analysis, and although many of these things are called the analysis of variance, they were not what I was talking about.

One of the minor miracles of modern mathematics for me is the way in which symmetry considerations can lead to Fourier-like decompositions rich in intuitive content and of great mathematical value. For a stochastic process indexed by the real line, temporal homogeneity implies that it is a mixture of uncorrelated sinusoids with random amplitudes and phases; for a random process on a sphere, homogeneity under the orthogonal group implies that it is a mixture of spherical harmonics with random amplitudes; and so on. All of these decompositions have exact or approximate forms for arrays of numbers, and can be motivated and derived without any considerations of probability or randomness, but I do think that the only way one can underline the unity I speak of is through the medium of dispersion models for arrays of random variables. I may be wrong in thinking this, as I may in thinking that this idea is helpful, but in taking the viewpoint that I do, I am certainly not undervaluing any of the other things called the analysis of variance.

To take a slightly different tack, I note that there are many ways in which statisticians or data analysts derive meaningful parameters, data reductions, hypotheses, decompositions,..., in short, analyses, whilst attempting to answer questions raised in the context of a given set of data. Some of the foregoing are achieved directly, at times with great insight, without any model assumptions; others are more transparently associated with assumed statistical models, and yet others are derived indirectly from more general structural considerations such as symmetry or invariance under a group. There are many interesting interconnections between these approaches, and my paper was intended to highlight one such interconnection relating to anova.

Even if I agreed with Lewis Carroll's view above, or with Humpty Dumpty: "When *I* use a word,..., it means just what I choose it to mean—neither more nor less," and I do not, I think I could still comment on the desirability of using the term analysis of variance to mean so many different things. In proposing the vote of thanks for the review paper Plackett (1960), Cox comments:

> Many practical applications of what is loosely called analysis of variance are really what Mr. Plackett calls "applications of regression analysis." That is, there is one error variance, a linear model with fixed parameters, and the main problem is to estimate these parameters and their standard errors. The analysis of variance is here an elegant way of setting out the least-squares calculations for the estimation of the error variance, and of providing tests for groups of parameters, where such tests are needed. Formally it may be helpful to consider expressions for expected mean squares in terms of "Components of variation," but at the practical numerical level this is not required. There is only one variance. As Mr. Plackett implies, there is a case for not using the expression "analysis of variance" for such applications.
>
> ...Problems of analysis of variance in the narrow sense may be said to arise when there are at least two variances in the problem that are of direct interest.
>
> Plackett (1960, pages 209–210)

I need hardly add that I would be pleased to see this view prevail. My own efforts towards this end began some years ago when I tried to introduce the term *anome*—analysis of means—for that part of the analysis of variance which is really regression analysis with a single error variance. I do not really expect this term to catch on but find it useful to make the distinction in teaching. (The similarity to anomie is entirely accidental.)

Having made these remarks about terminology, I would like to thank several of the discussants for taking the trouble to outline what the analysis of variance means to them. I have few comments to make about these expositions, however, because apart from possible terminological quibbles I generally agree with them. With reference to Graybill's remarks about the term *analysis* I should say that I *am* using it in a different sense from him; my analysis is a "resolution into simple elements," rather than a statistical analysis.

Diaconis "treats the problem data analytically" and outlines *his* view of spectral analysis, something which I have no difficulties with, although such an approach does not seem to go very far to me. Rather earlier than he, apparently, I need to think about the possible extension of my data set (populations), about whether the different factors have affected the means, the variances and covariances or both, of my data; thoughts which must be based upon the context of the data and the questions of interest associated with it. Only then would I turn to decompositions of my data; otherwise, I would not know what they tell me about my questions. Despite the difference of view—perhaps I am just not a data analyst—I enjoyed his comments and am very glad that he drew attention to Fortini (1977), for this thesis contains much fine work which is relevant to anova whatever one's viewpoint.

Although I am sorry that Kempthorne sees no more to anova than Pythagoras' theorem in finite-dimensional Euclidean spaces, I appreciate his comments and apologize for omitting reference to Zyskind (1962). My only excuse is that I made no attempt to give a complete list of references. Tjur's clarification of the relation between finite and infinite arrays was helpful, as were Hannan and Hesse's remarks on the restriction of anova models (in my sense) to subsets, something which I did not go into in this paper. Anderson surveys some facts about maximum likelihood estimation in these models, again something I did not mention, and I am sure that readers will find his comments on elliptically contoured distributions of particular interest. And I look forward to seeing the forthcoming thesis by Eugene Johnson referred to by Tukey.

A number of the discussants felt that my anova framework was too narrow, even given that it concerns only dispersion models for arrays of random variables. Hannan and Hesse ask why I restrict myself to cases where the identity representation of the isotropy group occupies a privileged place referring to Section 4 of the excellent paper Yaglom (1961)—which I should have cited—for more complex examples involving the vector case. Diaconis made a similar point in private correspondence, noting that Yaglom (1961)—see, especially, equation (3.9)—deals with situations far more general than Gel'fand pairs, contrary to a statement I made in Section 7. Yes, I suppose that these should be viewed as instances of anova, thus giving further evidence of my inability to draw a precise boundary; although I confess to being a little troubled having to include cases where noncommutativity of the dispersion matrices occurs and where matrix elements relative to arbitrarily chosen bases must be used in the representations. Brillinger suggests that we might broaden the notion of anova by referring to "some natural class of functionals" of the original array. Again his argument seems strong, but I note that the spectral representation and associated analysis of the variance of the process he mentions are instances of the theory given in Yaglom (1961, Section 4.4) for fields with random homogeneous increments. He also comments that one could dispense with any particular consideration of manova by using linear functionals, something which is done in Hannan (1970) for temporally homogeneous vector processes, and again I agree. However, I am not clear just what the general algebraic result that he alludes to should be. Lastly, Bailey invites me to consider as instances of anova (in the sense I used the term) certain models for dispersion matrices *not* defined by equality constraints on the entries. In the first class of models (Example 1) she mentions the variance is already analyzed, as the basic parameters of the model are the canonical variance components, whilst in her Example 2 one must take linear combinations of parameters to get the variance to be analyzed. I have tried to focus on that part of the theory which is common to all the instances of anova as I see them, and Bailey's examples do not seem, to me, to fit in.

In closing let me return briefly to the issue of terminology. Bailey, Nelder and Tobias remark that by focussing on dispersion models I am telling only half the story, but I think they would agree that what many currently call analysis of variance in the other half—models for mean values—concern little more than sums of squares and their associated linear subspaces. There are certainly open problems associated with the analysis of an array y under models such as $Ey \in \mathscr{T}$, $Dy \in \mathscr{V}$, where $\mathscr{T}$ is a linear space and $\mathscr{V}$ a class of nonnegative definite matrices, but surely no one wants to regard all such problems as concerning the analysis of variance. Of course, the mean value half is important, and the way in which the two halves fit together even more so, but it still seems to me that the half which has so much in common, both mathematically and scientifically, with the various types of spectral analysis, is the half relating to dispersion models.

Acknowledgments. I thank the Editor, the Associate Editor and all the discussants very much for their assistance, their interest and their very helpful comments.

REFERENCES

FORTINI, P. (1977). Representations of groups and the analysis of variance. Ph.D. dissertation, Dept. of Statistics, Harvard Univ.

GARDNER, M. (1960). *The Annotated Alice—Alice's Adventures in Wonderland and Through the Looking Glass* by Lewis Carroll. Clarkson N. Potter, Inc., New York.

HANNAN, E. J. (1970). *Multiple Time Series*. Wiley, New York.

PLACKETT, R. L. (1960). Models in the analysis of variance (with 9x). *J. Roy. Statist. Soc. Ser. B* **22** 195–217.

YAGLOM, A. M. (1961). Second order homogeneous random fields. *Proc. Fourth Berkeley Symp. Math. Statist. Probab.* **2** 593–622. Univ. California Press.

ZYSKIND, G. (1962). On structure, relation, sigma and expectations of mean squares. *Sankhyā Ser. A* **24** 115–148.

DEPARTMENT OF STATISTICS
UNIVERSITY OF CALIFORNIA
BERKELEY, CALIFORNIA 94720