STATISTICAL INFERENCE FOR INDIVIDUAL ORGANISM
RESEARCH: SOME REACTIONS TO A SUGGESTION
BY GENTILE, RODEN, AND KLEIN¹

JACK MICHAEL

WESTERN MICHIGAN UNIVERSITY

In many of the papers published in this journal, the results of experiments are presented with little or no formal use of the logic of statistical inference. There is instead extensive reliance on relatively simple graphic presentation. Descriptive statistics such as per cents and means appear quite often but usually without an accompanying significance test or confidence interval. A recent issue of JABA, for example, (Winter, 1973) contained 13 experimental articles and only three reported any formal significance tests.

Suggesting that the major reason for this neglect is the "assumed inapplicability of statistical techniques to individual cases or small numbers of subjects" Gentile, Roden, and Klein (1972) proposed the use of an analysis-of-variance model for the analysis of data gathered with the reversal design using a single subject. Their article stimulated four further treatments of this topic that were submitted to this journal and that appear in this issue. What follows are some descriptive comments relating each of the four articles to the germinal paper by Gentile *et al.* and to each other, then each of the four papers, and finally an attempted behavioral analysis of the role of statistical methods in psychology and the implications of this analysis for the various proposals made by Gentile *et al.* and the other contributors.

All four of the follow-up articles contested the assumption made by Gentile *et al.* that suc-

cessively obtained observations within a treatment condition may be considered independent of one another, a seemingly critical assumption for the valid application of the analysis-of-variance model. Kratochwill *et al.* performed an experiment in which they demonstrated the rather extreme degree of dependence seen in successively obtained observations of exactly the type analyzed by Gentile *et al.* Hartmann provides a rational analysis of the same issue using a hypothetical set of data. Thoresen and Elashoff generally support Hartmann's criticism and describe what they see as some additional weaknesses in the assumed approach of Gentile *et al.*

Finally, Keselman, though not arguing the point, also identifies the nonindependence of successively obtained observations as the critical problem not adequately dealt with by Gentile *et al.*

When the assumptions underlying a statistical inference procedure cannot be met, however, the procedure is not necessarily useless. Sometimes, the test is relatively immune to any but the most extreme types of nonconformity to assumption, in which case one can view the probability value as a rough approximation and use the procedure anyway. This situation is made even more tolerable if the failures to satisfy assumptions can be seen to cause probability errors in a known direction—for example to render the test conservative. Then, if a significant finding is obtained one can be sure that if a proper test were used, the result would be even more significant. At the end of the section dealing with violations of the independence assumption, Gentile *et al.* state exactly this point, but it appears that they were addressing themselves to a different type of

¹This is one in a series of articles available for \$1.50 from the Business Manager, *Journal of Applied Behavior Analysis*, Department of Human Development, University of Kansas, Lawrence, Kansas 66045. Ask for Monograph #4.

nonindependence than that seen as most critical by the other authors. Hartmann describes two ways in which the nonindependence of successively obtained observations within treatments would be expected to produce a positive bias, and another factor that could well produce a negative (conservative) bias if the Gentile *et al.* procedure is used; Thoresen and Elashoff agree with Hartmann's interpretation.

It would appear, then, that one of the critical assumptions underlying the analysis of variance is not usually met by the data obtained in typical JABA and JEAB experiments, and furthermore, that the effects of this failure are not easy to interpret. So, what *can* one do with such data from the point of view of statistical inference?

Keselman presents a formal rationale for the use of an analysis-of-variance model based on repeated measurements. He believes that this model is more appropriate to the type of data being considered and also more sensitive for the detection of significant treatment effects. The adequacy of his formal presentation must be assessed by someone with more statistical background than my own, but it does not seem to me that his model is free from the same assumption of independence of successively obtained observations within treatment conditions that underlies that of Gentile *et al.*

Hartmann suggests limiting the application of analysis-of-variance procedures to single-subject data that do not show sequential nonindependence. This is partially achieved by maintaining each experimental condition long enough to obtain several stable observations for

that condition and then using only those data in the analysis. In addition, some statistical tests of successive dependence can be made and the analysis of variance performed only if the null hypothesis regarding dependence cannot be rejected. At the close of their paper, Thoresen and Elashoff consider Hartmann's suggestions and conclude that it will not be easy to apply them in most current single subject research.

Hartmann, Thoresen and Elashoff, and Kratochwill *et al.* all, with some enthusiasm, refer to recent developments in the area of time-series analysis as the possible solution to the problems of assessing the statistical significance of single subject data, and the latter authors also suggest the relevance of some nonparametric models.

I was left with the general impression that the procedure proposed by Gentile *et al.* is really not applicable to most single-subject data. Several alternatives seem to be available, at least for some limited experimental situations, and there is considerable optimism regarding more generally applicable methods now being developed for time-series analysis—methods that take account of the troublesome autocorrelation found in much single-subject data.

REFERENCE

- Gentile, J. R., Roden, A. H., and Klein, R. D. An analysis-of-variance model for the intrasubject replication design. *Journal of Applied Behavior Analysis*, 1962, 5, 193-198.

Received 20 August 1974.

(Published without revision.)