Statistical Inference in Behavior Analysis: Friend or Foe?

Alan Baron

University of Wisconsin–Milwaukee

Behavior analysts are undecided about the proper role to be played by inferential statistics in behavioral research. The traditional view, as expressed in Sidman's *Tactics of Scientific Research* (1960), was that inferential statistics has no place within a science that focuses on the steady-state behavior of individual organisms. Despite this admonition, there have been steady inroads of statistical techniques into behavior analysis since then, as evidenced by publications in the *Journal of the Experimental Analysis of Behavior*. The issues raised by these developments were considered at a panel held at the 24th annual convention of the Association for Behavior Analysis, Orlando, Florida (May, 1998). The proceedings are reported in this and the following articles.

Key words: research design, statistical inference, experimental control, single-subject design

This and the following five articles report the proceedings of a panel held at the 24th annual convention of the Association for Behavior Analysis in Orlando, Florida. The panel members, besides myself, were Nancy Ator, Marc Branch, John Crosbie, Michael Davison, and Michael Perone. Subsequent to the meeting, Richard Shull prepared a commentary based on written versions of the panelists' presentations. His comments are included as a sixth article. As chair of the panel, I introduced the topic:

Behavior analysis has had a stormy relationship with statistical methods. Note that our concern is with those statistical procedures referred to as "inferential," in other words, procedures whose goal is to infer characteristics of a hypothetical population from a limited set of sample observations. There is no quarrel about the value of descriptive statistics. Experimental data

must be organized and summarized, and the methods of behavioral analysts are similar to those of workers in other research traditions. Thus, we too use the mean as a measure of central tendency, the standard deviation as a measure of variation, and regression analysis as a measure of association.

The issues first appear to have been broached in *The Behavior of Organisms*, published in 1938, where B. F. Skinner laid out the principles with which we are familiar. His research had a number of unique features, not the least of which was the unconventional way that he designed his experiments and treated the results.

Unlike much of the research of that time, Skinner's experiments focused on the behavior of individual subjects. He studied a small number of rats under a series of conditions, and within each condition observations were prolonged until the animal's performance was stable. This was quite different from the usual procedure in which groups of subjects were observed briefly, one experimental condition per group, and the results were reported as the average performance for the group. Also different was the way in which Skinner analyzed the results. Each subject's performance was reported and conclusions were reached from data summarized in graphs, usually cumulative records. He saw no need to verify his re-

This and the following papers were presented at the 24th annual convention of the Association for Behavior Analysis, Orlando, Florida, May, 1998. I thank the panel members, Nancy Ator, Marc Branch, John Crosbie, Michael Davison, and Michael Perone, and the commentator, Richard Shull, for helping me to understand better the place of inferential statistics in behavior analysis.

Please address correspondence about this article to Alan Baron, Department of Psychology, University of Wisconsin–Milwaukee, Milwaukee, Wisconsin 53201 (E-mail: ab@uwm.edu).

TABLE 1
Questions to address

Concerning your own research:
Do you use inferential statistical methods to analyze the results?
If so, how do you decide when statistical verification of findings is needed?
If not, how do you establish the reliability of your findings?
Concerning research done by others:
Do you regard findings from steady-state and group-statistical methods as interchangeable?
If not, how do you reconcile experiments that use different methods?
Concerning graduate education and publication policies:
Should training in statistics be regarded as essential preparation for a career in the experimental analysis of behavior?
What should the stance of the <i>Journal of the Experimental Analysis of Behavior</i> be about using statistics to clarify experimental results?

sults through the statistical tests that were coming into vogue at that time.

In *The Behavior of Organisms*, Skinner did not explain or defend his approach until the very last few pages when almost as an afterthought he briefly gave his reasons for deviating from more conventional methods. Later on, Murray Sidman discussed these reasons in considerably more detail in what has become the definitive methodological treatise for behavior analysis, *Tactics of Scientific Research* (1960).

According to Skinner and Sidman, group experiments and the statistical methods used to analyze the results are seriously flawed. The limitations are familiar to students of behavior analysis. First, although group differences may be statistically significant, the group averages conceal exceptions at the individual level. Laws that apply only to groups do not allow prediction of what the single organism will do. Second, when functional relationships are derived from groups exposed to different treatments, the functions do not originate in the behavior of any particular organism. In other words,

the average performances from different subjects do not have a clear counterpart in nature. Finally, and most troubling, is that the group-statistical method diverts the researcher from a full experimental analysis. The hallmark of the scientific experiment is the manipulation and control of variables. But the group-statistical researcher must be prepared to tolerate uncontrolled differences among subjects as an inherent feature of the research. This concession appears to undermine the reason for conducting experiments in the first place.

By comparison, Skinner's single-subject designs do not have these problems. The inquiry focuses on the ultimate concern: the behavior of the individual organism. Extraneous variables are controlled within the experiment rather than averaged out statistically. Effects of variables—functional relationships—are examined as they naturally occur, within the same organism rather than as a construction from the performance of a group. And there is no need for inferential statistics because behavior is observed as a steady state.

To judge from the research literature, the methods of Skinner and Sidman were taken quite seriously by the behavior-analytic researchers of that time. Several years ago, I surveyed the designs of experiments published in issues of the *Journal of the Experimental Analysis of Behavior* during the late 1950s (Baron, 1990). Not surprisingly, inferential statistics were almost never used. They were employed in fewer than 5% of the reports, and it was the rare experiment that followed a between-group design (the type of design most closely linked to inferential tests). But over the years, something happened. By 1990, close to one third of the experiments employed inferential statistics, and between-group designs also had increased in frequency. Although I do not have any new data to present today, my impression is that this trend is continuing.

So it appears that a significant number of us are violating the precepts of

our founding fathers. The following six articles address this puzzling discontinuity between contemporary behavior-analytic research practices and traditional views of the way that research should be conducted. The individuals I enlisted in this effort all have distinguished research records, as well as considerable experience as editors and reviewers of behavior-analytic research. To organize our discussion, I provided them with a series of questions (see Table 1). However, knowing

that they all are free spirits, I will not be offended if they take the discussion in other directions; our goal is to explore the issues as broadly as we can.

REFERENCES

- Baron, A. (1990). Experimental designs. *The Behavior Analyst*, 13, 167–171.
- Sidman, M. (1960). *Tactics of scientific research: Evaluating experimental data in psychology*. New York: Basic Books.
- Skinner, B. F. (1938). *The behavior of organisms: An experimental analysis*. New York: Appleton-Century-Crofts.