have special slants or blind spots that make it difficult for them to do a good index. An individual may be the authority on the content of his particular book, but it does *not* follow that he has a good enough over-all, objective view of the whole field to which his book is a contribution to do a competent piece of indexing. This is particularly true in those fields where, for various reasons, terminology is not stabilized.

The point overlooked by both Clark and the gadgeteers is that research, writing, publishing, indexing, literature searching, and criticism are all parts of an extremely complex, ever-changing *social* process in which more judgments and decisions, involving values as well as subject-matter technicalities, must be made than any one person can handle.

HENRY BLACK
Bibliographical Services,
New York, New York

Statistical Evidence

Warren Weaver's statement ["The disparagement of statistical evidence," *Science* **132**, 1859 (1960)] that "statistical evidence is, in essentially all nontrivial cases, the only sort of evidence we can possibly have" seems to me to be as "wholly unwarranted" as is the practice he very properly condemns—namely, the "automatic discarding of evidence because it is statistical." Statistical evidence is usually very useful, sometimes essential, but there certainly are many kinds of useful evidence that are not statistical in the usual meaning of that term. Our belief that the earth is not flat, that it revolves once in 24 hours, and that it completes an orbit around the sun in a year is not based on statistical evidence. Neither does our acceptance of the theory of evolution, or of a dozen other theories that might be mentioned, depend to any great degree on statistical evidence, although these theories may be and often are supported by such evidence.

The difficulty is partly semantic. Some writers seem to regard as statistical practically any method of dealing with quantitative data, but usually the term implies frequency distributions, standard errors, analysis of variance, correlation coefficients, and so on. Presumably it is these latter that Weaver had in mind. If so, he certainly must realize, on second thought, that failure to use these techniques does not automatically negate the usefulness of quantitative data. Mendel, for example, did not use them, and yet he revolutionized our ideas of heredity.

Weaver is not the first to imply or state that statistical methods are essen-

tial for all really worth-while research. And if we enlarge the subject to include mathematics, of which statistics is a branch, several noted scientists could be added to the list. Yet Darwin had little use for mathematics, and Faraday's experience with mathematics is said to have been limited to turning the crank of a calculating machine. G. N. Lewis has deplored the tendency of many to overrate mathematics.

Both mathematics and statistics have made, and will no doubt continue to make, revolutionary contributions. But we might well ask whether the occasional reluctance to use them or to accept the conclusion derived from them may not be due in part to the tendency to overemphasize their value. Pasteur pointed out many years ago that overemphasis of any kind leads to reaction, which, again overshooting the mark, makes the search for truth ever more difficult.

S. C. SALMON
4103 Roanoke Road,
Hyattsville, Maryland

Warren Weaver's editorial has stirred up some discussion here. As a statistician in experimental work, I should like to present a statistician's view.

The simple phrase *statistical evidence* is, unfortunately, widely used to describe two very different kinds of evidence: (*a*) that provided by data gathered from uncontrolled events just as they happened to occur "in life," and (*b*) that provided by data gathered from a planned experiment in which every effort is made to prevent the effects being studied from being confounded with effects of irrelevant factors.

Weaver mentions this dichotomy, but his ensuing remarks seem to imply that he regards the difference as one of degree. I would go further and say that the two are entirely different, and that a major reason for the slow improvement of the quality of scientific inference has been the confusion of *a* with *b*. Scientists who call both *a* and *b* "statistical evidence" have properly rejected *a* but have then gone on to reject *b* merely because they have given it the same name; the result is that they do not know or use as much statistics as they should. Some of them might even be surprised to know that the statisticians are very much on their side and have even invented a distinguishing phrase, "historical evidence," for evidence of the kind described under *a*.

Evidence of type *a* is exemplified by the sports column I once read in which the writer decried the firing of football coaches and produced "statistical evidence" of the folly of this action by displaying data showing that colleges which keep their coaches for long periods of time have much better records on the field. Much of the "statistical evidence" relating to smoking and lung cancer may be in a similar category; we do not know.

Type *a* evidence is good only for suggesting experiments which may produce type *b* evidence. Unfortunately, there are fields in which type *b* evidence cannot be obtained. This may force us to use type *a* evidence, but it does not force us to assign to it the degree of confidence that is associated with type *b*. In fact, those fields wherein type *b* evidence cannot be produced might be better described by some word other than *science*.

ROBERT HOOKE
Westinghouse Electric Corporation
Research Laboratories,
Pittsburgh, Pennsylvania

Warren Weaver's provocative comments on the power of statistical evidence merit repeating. They merit repeating in particular to scientists who are not mathematicians.

The disparagement of statistics on the part of the tobacco industry and of political parties, on which Weaver comments, is understandable (although not commendable) in view of the vested interests of these groups in the results. But the disparagement of statistics on the part of scientists—of men, presumably, with a vested interest only in seeking truth—is deplorable.

In a field of science, such as geology, where there are so many variables, the application of statistical methods is most appropriate. The usefulness of experimental design should be particularly obvious to any working geologist. Yet it is true that only the most prescient geologists are receptive not only to statistical evidence but to statistical methods as well.

I think the reasons for this reluctance are many. In part it reflects the general opposition people (and scientists among them) display toward new ideas that they cannot fully understand. Certainly this is true of geologists, who, as a group, know too little mathematics to comprehend even the general problem, and who are reluctant to accept new ideas that originate outside the ranks. There is also this factor: statistical methods have been taken up by a few zealots who, through a combination of ignorance and enthusiasm, have succeeded in overselling them to the scientific community. There is no doubt that the proper application of statistical theory requires a mathematical maturity not readily come by. Then, certainly, the wild use of statistics in the hands of the advertising agencies has not raised the reputation of statistical

methods with thoughtful segments of the public.

How to make the general scientific fraternity more receptive to these practical statistical methods poses a problem. It would seem to me that one effort which might help this situation would be to emphasize in education the importance of combining enough statistical theory with other scientific education to make scientists better informed with respect to its potential. And vice versa: urge the would-be statisticians to study enough of one other science so that they can comprehend the problem to which they may subsequently apply their methods. Perhaps, also, there should be more general articles written by qualified statistical experts, who could write for the essentially nonmathematical scientist in *his* technical journals. These articles should describe specific case studies where statistical methods apply.

In any event, I greet Weaver's editorial as welcome prose in *Science*.

M. MATHEZ
Standard Oil Company (New Jersey), New York, New York

Warren Weaver's editorial is trenchant and timely and deals effectively with a most urgent problem. In directing attention to the growing role of probabilistic processes, Weaver has done all of us a great service.

In my opinion, the effectiveness of Weaver's message is marred by an unfortunate usage of the phrase *statistical evidence*. It is important to distinguish between evidence from both experimental and nonexperimental observations, and between the statistical methods with which this evidence is analyzed. For example, the relationship between cigarette smoking and lung cancer, on the basis of all evidence and its analysis, is amply confirmed. The basis for this conviction, however, is not so much statistical as it is the convergence of multiple studies to support the same conclusions, and the finding that diminution in cigarette smoking leads to diminished risk of lung cancer and other diseases. The nub of this protest is that the "evidence" ought to be described by an adjective referring to the nature of the problem and to methods of collecting data—for example, genetic evidence, mortality evidence, and radiological evidence. The proper adjectival use of the word *statistical* seems to me to be restricted to the method by which the data are analyzed and to conclusions drawn concerning whether or not the data conform to random distributions or to other prescribed distributions.

This objection would be trivial were it not for the suspicion that the use of the phrase *statistical evidence* tends to obscure a more fundamental property of many of the sets of evidence which are cited in the editorial—namely, that the evidence is obtained from the observation of naturally occurring phenomena rather than from experiments. It is precisely in this realm of evidence that statistical analysis has provided some powerful insights; nevertheless, the nature of the evidence is often of greater importance to the interpretation of any conclusions than the fact that it was subjected to statistical analysis. As indicated by three out of the four examples given in the editorial, there is special likelihood that man or his reactions will be studied from sets of nonexperimental rather than experimental observations. It is a matter of some concern that, by placing the emphasis on statistical analysis or even upon the probabilistic character of data, the inherent problems associated with the collection and proper interpretation of nonexperimental data about man may readily be overlooked.

These problems are particularly serious in areas of chronic-disease epidemiology, in sociology, and in human genetics. One of these inherent problems is the recognition, measurement, and selection for study of important variables which may influence the as-

sociation which is being studied. A second problem is the collection of data under circumstances which prevent bias and which provide full freedom for the data collected to refute the hypothesis being tested. A third problem has to do with the relationship of evidence to conclusions. While statistical criteria are often important, the convergence of evidence from several independent studies, the demonstration of joint gradients in the magnitude of the supposedly related variables, and the demonstration of reversibility of the dependent variable by the selection and study of suitable sets of data are important guides for drawing conclusions from nonexperimental data. Despite the important contributions which statisticians have made to the management of the inherent problems of nonexperimental data, these are not solely statistical problems but problems of logic and method.

That there are other areas of science where comparable problems exist is shown in the same issue of *Science* by the article "Recent statistical studies in astronomy" [**132**, 1870 (1960)]. Here, also, the basic data are nonexperimental, but the use of systematic data collection and skilled statistical analysis permits some striking and very important conclusions. It is of interest that the use of statistics in the article about astronomy has not led the author to describe his data as "statistical" data. It is proper to describe them as "astronomical" data.

The word *evidence* is also used to designate certain facts presented in a trial; following this analogy, one thinks of statistics as a sort of jury or judge, which helps to decide, on the basis of law (accepted scientific criteria), whether the evidence presented supports the allegations of counsel (the scientists) with sufficient certainty to lead to a verdict. With naturally occurring data, particularly concerning man, we need to focus more attention on the adequacy of the criteria for drawing conclusions (the law, in the analogy of the trial).

In my opinion, use of the phrase *statistical evidence* should be discouraged, since it tends to obscure recognition of this problem.

JOHN R. GOLDSMITH
State of California Department of Public Health, Berkeley

I am delighted that my editorial aroused some interest, and I appreciate a chance to comment. I will not deal in detail with all four letters, for that would be tedious and repetitive. I think a few remarks may clarify my position.

1) An editorial in *Science* contains about 500 words. It is not feasible very fully to develop a subject—even a restricted and minor one—in so brief a statement, nor is there space to give qualifying refinements.

2) The phrase *statistical evidence* appears to have been assumed synonymous with statements made by professional statisticians and making explicit use of "frequency distributions, . . . analysis of variance, correlation coefficients, and so on." There even was some assumption (in other letters which I received) that I had argued in favor of *all* "statistical" reasoning, whether good or bad!

Not being a statistician, I used the phrase less professionally and much more broadly. By statistical evidence I meant evidence in the gathering *or* analysis of which probabilistic considerations enter. Since the entire universe is, as far as science now knows, made up of elementary particles all of whose ultimate laws are probabilistic in nature, it would, at least to me, seem rather difficult to produce any evidence which does not, at some stage of refinement, involve such considerations.

If Salmon will approach the subject from this point of view, he will, I am sure, recognize why I think his "nonstatistical" conclusions (about the nonflatness of the earth, and so on) are so inescapably statistical. His remark that

the earth "revolves once in 24 hours" is particularly vulnerable, partly because it is strictly untrue, but chiefly because one cannot possibly discuss the degree of accuracy of the statement or the reliability of the evidence without a detailed and explicitly statistical analysis.

3) One can, of course do a great deal of very useful and quantitative science without utilizing the procedures of the professional statistician, and certainly without bothering to remember that all evidence is (in the sense stated above) statistical. I had no slightest intention of implying otherwise.

4) Hooke states that the phrase *statistical evidence* is applied to data gathered by observation of uncontrolled events, and also to data gathered from planned experiments. Since these two together include, as far as I can see, all data, Hooke would appear to agree with me that all evidence is statistical evidence. But Hooke has a low opinion of type *a* evidence, this naturally resulting from the fact that by his type *a* he really means (as revealed by his next-to-last paragraph) badly argued conclusions from poorly observed data. Darwin used type *a* evidence and revolutionized man's thinking in the process.

5) I do not advocate turning all of science over to the statisticians, nor do I think that a small boy, when he is counting his marbles, need be reminded that the counting of electrons is a very queer and slippery business. But I do object to the snide implication that evidence which is "merely statistical" is, by virtue of that fact, silly and unreliable.

WARREN WEAVER
Alfred P. Sloan Foundation, New York, New York

Nature and Nurture

Questions about the effect of environment on intellectual potential [G. Allen, *Science* **133**, 378 (1961)] should be considered in light of Spitz's studies on "hospitalism" [R. A. Spitz, in *Psychoanalytic Study of the Child* (International Universities Press, New York, 1946), vol. 1, pp. 53–74]. Controlled studies of institutionalized infants showed a drop in developmental quotient from 124 to 72 in institution X during a given period; in institution Y there was no change, and infants at home in comparable socioeconomic areas showed no change. The significant variable was the presence of one mother or mother substitute for each infant in institution Y and of one mother-substitute for each eight infants in institution X.

Follow-up studies 2 years later