ON WRITING STATISTICAL ARTICLES FOR THE INTERNATIONAL ENCYCLOPEDIA OF THE SOCIAL SCIENCES

F. J. Anscombe, Yale University

1. Statistics was originally a social science. Today the statistical method finds many applications in the social sciences. (For example, a college senior undertaking a research project as part of an honors program is, I believe, much more likely to find himself unexpectedly faced with a statistical problem if his field is one of the social sciences than if it is in the natural sciences or humanities.) But most of the impetus for development of statistical methodology in this century seems to have come from outside of the social sciences. Statistical methodology is basically not oriented towards any particular subject matter; it is as neutral as mathematics. Although most of the methods and concepts are fully as appropriate in the social sciences as anywhere else, the examples of statistical phenomena that come to mind most readily are not (for me) drawn from the social sciences. I have found it a bit embarrassing, in writing for the Encyclopedia, to try to impart a social science flavor, and no doubt some other contributors have had the same experience. When working on an article on outliers, I found that all the good examples I could recall concerning outliers were from the natural sciences. The last example of an outlier phenomenon that I have seen in a social science (it was in political science) was remarkable because nearly every observation seemed to be an outlier! All this is a matter of the accidents of one's personal experience.

2. What should the articles on statistics be like? Presumably they are addressed, not to professional statisticians, but primarily to social scientists who already have some acquaintance with statistics. There is surely no point in trying to do what is already done well in many textbooks, namely (i) explain how to make statistical calculations (of correlation coefficients, standard errors, analyses of variance, factor analyses, etc.) and (ii) give the mathematical theory underlying these methods. There is probably little point in trying to interest novices; it would be hard to compete with, say, Wallis and Roberts's paperback "The Nature of Statistics". Our best target would presumably be to try to do what is not well done in most books -- to address an adult reader, impart wisdom and insight, try to bridge the gap between the too glib textbooks and reality. Some target, yes!

3. Harry Roberts's very interesting talk raises the problem of controversy within statistical science. How much should appear in material addressed to non-professional statisticians and general readers? Surely it is unwise to suppress controversy so that a united front should be falsely presented to the outside world. But there are more appropriate places than the Encyclopedia for an author to participate actively in such controversy. We do not expect a physician to be as argumentative with his patients about fundamentals of medicine as he might be with some of his colleagues. Most of the controversies in statistics seem to have had a by-product that is both valuable and not in itself controversial, namely, an increased understanding of the diversity of statistical problems and of the many factors that enter into them. Ideally, the Encyclopedia articles on statistics should not hide the existence of controversy, but as far as possible divert attention to this by-product rather than attempt to persuade the reader to any one side. For example, the controversy round about 1950 over a randomized test for association in contingency tables brought out a distinction between decision making and inference, and that distinction can be expressed in dispassionate terms unlikely to raise anyone's ire. The present controversy over Bayesian methods has made us think more about the various kinds of uncertainty and vagueness in statistical problems, and these can usefully be described and discussed without partisanship.

4. Whether there is any point in making the above remarks now, I do not know. They would have been more appropriate several years ago in a discussion of policy, instead of now when most of us contributors have done our bit. But only now, having blundered through an assignment, can I (for one) see such policy questions clearly enough to venture an opinion on them.