

SOME PRELIMINARY THOUGHTS

"Facts is facts" may or may not be an accurate Twain-ism, but provides a target that should remain focused in the sights of every practising statistician. For, of course, facts is not facts. Data is not 'given', whatever its Latin root might say. On the contrary, 'facts' and 'data' are products of particular social processes, and must be analyzed as such. This relativism is ignored in many current political debates, and a valuable starting point for the active statistical sceptic can be to emphasise the relativism of facts, and to re-insert this relativism within the political argument. This short paper will outline in a preliminary fashion several questions which may be relevant to such a strategy.

First, of course, we should admit that the title is poorly phrased, since statistics itself (themselves?) can do nothing. If anything is done concerning statistics, it is done by human individuals with motivations and interests of their own. This thought leads to the first question.

1. If data is 'given', who gives it? (Who whom?) This has been most adequately followed up in the context of criminal statistics. Even Durkheim knew how statistics on suicide can mislead. The recent effect of the abolition of capital punishment upon murder figures is another example, as is the impact of the Wolfenden Committee which 'at a stroke' eliminated many offences pertaining to homosexuality. In short, criminal data is 'given', not so much by the criminals, as by the law-makers, police and judiciary. Hence they tell us more about the latter than the former. This theme has been extended by Cicourel, Douglas, and other ethnomethodologists, and it is also relevant in other areas.

Barry Hindess has shown how Indian data on social class can lead to completely new conclusions if analyzed from a novel theoretical perspective. The problematic, one might say, is problematic. One is then led to enquire why one particular conceptualization of data, one particular corpus of 'knowledge' is put forward by the statistics-makers, rather than another. Clearly there is no general answer to this question, except perhaps Blum's assertion that "the constructions of a corpus of knowledge is inextricably linked to the interests of those who produce it". But the fact that the question has no universal answer does not make it a useless question. On the contrary, enquiries concerning the social construction of reality are by their nature critical and potentially subversive. Quoting Blum again, "To find a corpus of knowledge problematic is equivalent to challenging the rules to which members of a collectivity subscribe". The 'corpus of knowledge' considered here is the body of empirical 'facts'. This is regarded as problematic when we refuse to accept them at face value. Facts are not sacred, and this realization challenges the 'rules' of the 'collectivity' which presumably consists of the statisticians and administrators who define the questions to be asked and the methods and categories in terms of which the answer is to be given.

2. How does statistics influence the questions that are asked? Are important issues trivialised or inhibited by the need to use available data or standard techniques? (e.g. scale construction, governmental data). Do questions relate to 'averages' rather than inequalities? (Cost-benefit analysis and overall efficiency). Who poses the questions - the administrator, owner, researcher or those most likely to be affected by the answers? Who is being planned for, and who is being planned against? Is the statistician involved soon enough for an attempt to be made to fit the technique and data to the question, rather than the other way round?

3. How does statistics influence the ways in which questions are 'answered'? Is the importance of 'significance' questioned? Does statistics dominate 'common sense', and if so has this led the layman (with more of the latter) to be dominated by the expert (with possibly more of the former)? Has the existence of stereotyped techniques allowed the statistician to avoid questioning the value judgements incorporated therein? Has this led to an essential conservatism? (Consider the following quotes from pp.87 and 137 of 'Systematic Empiricism' by J. and D. Willer: "The survey researcher is therefore no more than a well-trained spy for those having the economic power to support him", "empiricism is essentially conservative ... its necessary social outcome is totalitarianism".)

4. The reporting of results. Who receives the results? (Commercial or official secrecy). Who understands them? Who could understand them? Who will lose from the results, and who will gain? ("Sometimes people are reluctant to be interviewed because they say 'I don't see the point of it. What good will it do?' It may do none: but we hope to publish a book on the results, which will be read by those interested in the social changes which are occurring in Britain". - from Interviewers' Instructions quoted by P. Willmott and M. Young on p.142 of Family and Class in a London Suburb).

Is there scope for the results to be reworked using an alternative theoretical paradigm? Can disaffected workers publish individually? (Is this freedom guaranteed by his contract of work?) Would the answers to these questions be different if a non-statistical approach to research had been adopted?

5. How does statistics affect policy recommendations? Have non-quantifiable factors been footnoted and forgotten? Did the variables chosen determine the policy outcome? (Beeching rationale for 'non-profitable' railway lines). Have 'objective' factors dominated the 'subjective'?

6. What is the impact of politics on statistics? Does the Official Secrets Act inhibit discussion? A recent paper in a prestigious statistical journal suggests that it might. The authors admit their awareness of an implicit censorship upon statistical issues which may be contentious. "We are not free", they say, "to follow (a political argument advanced by one of the discussants)". And later on the same page "A lot has happened since we wrote about the R*** S***** G**** ... It has become a live political issue and this obviously inhibits our comments."

7. Why are resources devoted to statistical investigations? Are they sufficiently important? After all, as Lloyd George said, you cannot feed the hungry on statistics. But perhaps there are other benefits. Or perhaps statistics obscures the true costs.

8. How has statistics influenced the nature of research? Is funding obtainable for examining important questions in a non-statistical manner? for trivial questions in a statistical manner?

9. Does statistics influence the distribution of power?

the expert	versus the layman
the statistician	versus the non-statistician
the academic	versus the practitioner
the employer	versus the employed
the research funder	versus the researched into
the educated	versus the non-educated
the rich	versus the poor
the writer	versus the reader

10. What does all this imply for teaching statistics? It seems to me that there are only two responsible ways of teaching statistics. One is as a branch of mathematics, the other as part of the philosophy and sociology of science.

11. How should we approach a criticism of applied statistical work? Possibly by a multi-pronged attack, which may be summarised in questions such as the following:

- a). Why is your investigation important?
- b). Why are the questions asked meaningful? Do they rely on any implicit or explicit assumptions?
- c). Has the formulation of the questions been influenced by the research methods open to you? If so, what alternative questions have been thereby ignored?
- d). What concepts are employed in the questions asked? What connotations/assumptions underlie the use of these concepts? What alternative concepts could be employed?

In formulating questions such as these, we must remember that they are addressed to ourselves as well as to others. We are all guilty, at one time or another, of the crudest of crude empiricism. However, in opposing this we need not be led to the opposite extreme, of a nihilistic anti-empiricism. Do not forget that even Uncle Karl was in some sense an empiricist. "The way to proceed is to anchor one's descriptions in 'real' (that is, empirical) descriptions. A 'real' description shows some aspect of man's relation to nature. Thus, we should begin with real premises and deduce our knowledge from them".

Clearly this is incomplete, and the questions listed could probably be rephrased.

I would appreciate feedback and suggestions as to how the ideas presented here can be developed.

John Bibby
The Open University, MK7 6AA
(Dial 0908-63150 or (Home) 0908-312372)