NATURE Volume 245 OCTOBER 5 1973

What to Say About Scientific Evidence

NOT all the letters we receive each week convey hard scientific facts ripe for publication. A few chide us for omissions and commissions and some question our ability to publish a hymn sheet, let alone a learned journal. But one received recently set off an interesting train of thought. The writer had singular views on low temperature physics, and was outraged that the journal in the past had declined his contributions on this subject. "Mr Brimble, one of your predecessors," he wrote, "refused to see the Truth of this proposition; he died shortly thereafter". The implication was palpably *post hoc ergo propter hoc* and we tempt providence by turning our backs again on this Truth. The question that seemed fascinating was; granted that a scientist could see flaws in the implication of this indelicately worded letter, how well could he see flaws in the logic of his own work when it had all the trappings of a scientific paper?

The most dangerous pitfall in scientific reasoning is buried little deeper than the *post hoc* proposition. It is the acceptance of the statement 'if A implies B, then B implies A'. To take an example of such reasoning which does not these days point an accusing finger at anyone, 'if the Earth being at the centre of the Universe with the Sun rotating daily around it implies a succession of night and day, then our observation of night and day implies an Earth-centred Universe'.

Mathematicians are particularly aware of the snags in this because the issue is that of sufficient and necessary conditions. The truth of B is, of course, a necessary but not sufficient condition for the truth of A.

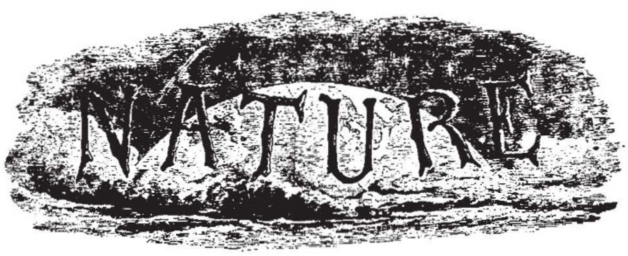
In observational science the chances are relatively slight that we can measure enough of the implications of A to rule out any other hypothesis than A. A common problem is of failure to identify enough alternative hypotheses. A coin is tossed 1,000 times—it shows 'heads' in 600 cases. It would be totally wrong, for instance, to say that 600 'heads' makes it highly credible that the coin is weighted. Instead, it may be bent, the experimental procedures may be at fault, or the experimenter may even be lying; all of these alternative hypotheses have to be assessed carefully. It is the mark of the successful scientist that he has rich enough an imagination to look for these alternative hypotheses, particularly when the conventional one is popular.

An unseemly phenomenon familiar to all scientists is the rolling bandwagon, on which our logical fallacy frequently rides. A first-rate new idea is introduced into a subject. It is instantly attractive since it explains a variety of previously mystifying observations. Whole laboratories swing into action and within months, papers are emerging which run thus—the new idea implies the following predictable observations; we have made these observations and therefore we have helped verify the idea. A frequent phrase is "we consider that our observations strongly support the notion that . . ." This may be pure nonsense, and history takes its toll of such papers. Even if the observations are indeed concordant with the hypothesis and hindsight shows them to have been of use in establishing the validity of the hypothesis, they cannot support the new idea unless they clearly also refute all other hypotheses or at least render them all highly incredible. Such crucial experiments are rarely possible. When the experiment is not a crucial one, the most that the experimenter should say is that he is bringing forward new evidence to bear on the issue. The validity of new concepts in science is usually established ultimately by processes not dissimilar to those involved in the determination of guilt or innocence in a court of law, and if the witnesses each bring their own opinion as well as their evidence this rapidly discredits both witness and evidence.

Moreover, the desire to vote fashionably can seriously inhibit science by requiring any imaginative thinker, with an alternative hypothesis which he wishes to test, to sift out evidence and decide how much of the published data has not been tarnished. Preconceptions not only creep into 'interpretation' but often even control the way in which the experiment is performed. Scientists are sometimes guilty not only of accepting, suitably camouflaged, the proposition that B implies A, but are also capable of designing experiments that will only yield B as an answer, or designing a rationalisation to remove all not-B from their data.

The solution to the problem of evaluating scientific evidence lies with the practitioner himself. There is no point in attaching to research departments experts in scientific method alone. Scientists must indulge in more healthy scepticism of their own and other people's treatment of data; there is a continuing need of strong and informed criticism of the way science handles it evidence. When properly done this is the greatest benefit of the refereeing system.

100 Years Ago



Dynamometers, by R. S. Ball, LL.D., F.R.S.

If we adopt that force which acting on one gramme for one second will impart the velocity of one centimetre per second as the unit, then one million of such units is a convenient magnitude for practical purposes. The large figures on the dynamometers represent these million units, for which it is hoped that ere long a suitable name will be adopted. The dynamometers are intended for educational purposes. They are exhibited to the Association with the desire of aiding the present movement in favour of an improved system of fundamental unity.

A report from the British Association

From *Nature*, **8**, 497, October 9, 1873.