

Can evidence ever be inconclusive?

The belief that stark conclusions can be made less definite by suggesting that inconclusive evidence is relevant, but in some manner undefined, can lead to trouble.

EVERY so often, readers of the journals come across articles whose titles begin with the weasel words "Evidence of..." or "Evidence for..."; they should be appropriately cautious of the text that follows. In the interests of plain speaking, this journal does its best (not always successfully) to ban these forms of words, which have been devalued by ambiguous and often tendentious usage. Their precise meaning has most recently arisen in connection with Fermilab's announcement a few weeks ago (see *Nature* **368**, 805; 1994) of what has widely been described as the "discovery of the top quark": both the laboratory's announcement and, it is said, the article submitted for publication have titles that begin in just this way.

In the case of Fermilab, mercifully, there are ample (if unpublished) indications of what the words are meant to mean. For at least a week before the announcement, the laboratory had agonized about the significance of the data it had gathered. An announcement was supposed imminent at the spring meeting of the American Physical Society, but was then put off. In the end, the laboratory went out of its way to emphasize that the data available for analysis were less conclusive than it would have wished. It is not merely that the estimated error of the estimated mass of the top quark is uncomfortably large (roughly 10 per cent), but that it is conceivable that the measurements may be differently interpreted.

What seems to have decided Fermilab in favour of publication is that its data are more persuasive than others have so far produced, and that it is likely to be some time before anybody will be able to do substantially better. So in this case "Evidence for..." means something like "The best evidence yet...". Nobody who has followed the laboratory's statements on just this point can doubt that it has acted responsibly, seeking not to seem to make a false claim. So why did it not use the phrase "The best evidence yet..." or, if persuaded that only latinate constructions appeal to the journals with which it deals, "Further evidence for..."? One version of a press announcement uses the equally proper "New evidence for..."

None of this implies that "evidence for" should never be used, but merely that it must be sufficiently well qualified for its meaning to be clear. Thus "evidence from seismic observations for a molten core" makes sense, but "seismic evidence for a molten core" is better. The improper usage is that in which the words are used to refer to a conclusion that is only partly

supported by the self-same evidence.

Because people these days read so little, the rhetorical stratagem succeeds once the unsupported conclusion is embodied in a bibliographically retrievable title, and there are of course other ways of doing that, as in "Possible theories of cold fusion" (Fleischmann, M., Pons, S. & Preparata, G. *Il Nuovo Cimento* **107A**, 143; 1994), against which journals should also be on their guard.

Yet "Evidence for..." persists. Conversations (or, more accurately, arguments) with authors reveal a profound misunderstanding on this point. If an article with a fine declarative title should run into trouble with its referees on the grounds that the text does not fully support the conclusion, but the work itself is inherently interesting, there is every likelihood that the author will seek to keep the original title, prefacing it with "Evidence for...". The misunderstanding seems to be that the weasel words betoken evidence that falls short of proof. The most effective way of concluding these conversations is to offer the alternative, "Inconclusive evidence for..."

The same misunderstanding seems to underlie a dispute that has arisen over an important article published last year under the title, "Evidence for large upward trends of ultraviolet-B radiation linked to ozone depletion" (*Science* **262**, 1032–1034; 1993). This is an important article, due to J. B. Kerr and C. T. McElroy of Environment Canada, based at the Atmospheric Environment Service, Toronto, who have developed and operated for the past four years a ground-based instrument for the measurement of ultraviolet radiation in the 300-nm region (from 290 nm to 325 nm).

In the argument about the consequences for living things of ozone concentrations in the stratosphere, one obvious logical gap has been the absence of information about the intensity of ultraviolet-B radiation at the surface of the Earth. While the concentration of stratospheric ozone (which absorbs at 300 nm) must be one of the chief determinants of the ultraviolet intensity, absorption in the troposphere (by particulate pollution, for example) and even the weather can affect day-to-day intensity on the ground.

Kerr and McElroy have designed an instrument that is excellently suited to the accumulation of an understanding of ultraviolet intensity on the ground. First, it measures intensity at intervals of 0.5 nm throughout the spectral range, routinely scanning the defined range twice each hour. The same station at Toronto also operates a

system for the direct measurement of stratospheric ozone, so that the effect and its supposed cause can be directly correlated.

So what can have gone wrong? By "large upward trend", Kerr and McElroy meant what they say in their abstract and their text, that there is an upward trend of ultraviolet radiation to match the measured downward trend of stratospheric ozone in the past four years. But do four years constitute a sufficient interval of time to prove a trend, especially when ozone concentrations in one of the years concerned were exceptionally reduced by debris from the Pinatubo volcano, and when ozone data for the winter of 1991–92 are missing (because the ozone instrument was away for calibration)?

The issue has been seized upon by Dr S. Fred Singer, the long-standing thorn in the flesh of the environmentalists who is now director of the Science and Environmental Project in Washington, DC, but otherwise a professor at the University of Virginia. His complaint is not so much that four years (of which two are exceptional) are not enough on which to base claims of a trend, but that the appearance of a statistical trend is almost entirely accounted for by the exceptionally high ultraviolet measurements towards the end of the period.

The argument, apparently about to be rehearsed by a published criticism of Kerr and McElroy and a rebuttal, is in reality an irrelevance. What matters about their work is that there is now an instrument of proven reliability with which, it is hoped, ultraviolet intensities elsewhere than at Toronto will be tracked. Moreover, their data do indeed show a good correlation between ozone concentration and ultraviolet intensity at the surface, irrespective of season or time of day; analysis will in due course reveal how important are the confounding factors (particulate pollution and the like).

But Singer is surely right to protest that the trend now claimed cannot, on the evidence, be a proof that ultraviolet intensities are increasing in the global and secular fashion implied. (When the article appeared, there was a great deal of publicity for the estimate that winter-ultraviolet intensities appeared to be increasing at 0.4 per cent a year, even though they are negligible compared with summer intensities at 300 nm). But the authors say that their rebuttal will in part consist of saying that they did not claim a "large upward trend", but only "evidence for" one. The disputed words have all the appearance of having been added to satisfy a referee, but so what? **John Maddox**