

What is the scientific literature?

Professional people have won a poor reputation for their skill at communicating with each other. The complaint may unfortunately be justified.

By what test are the scientific journals counted as literature? The bare minimum of an answer is that they are collectively referred to in this way by their contributors. Collectively, they also have the quality of permanence; they sit on library shelves for decades on end, and are referred to with reverence by those who contribute to later issues. So much is the bread and butter, or at least the bread and water, of scholarship; there are few complaints at the value of the accumulation of journals as the means by which the record of discovery may be reconstructed. But the same cannot be said of the value of the scientific literature as a means by which information is conveyed between people unable to communicate directly, by dialogue. Why should the general admiration of the way that discoveries roll off the laboratory benches coincide with a general belief that the scientific literature is impenetrable, often to its users?

There are some obvious excuses that do not deserve the dignity of being called explanations. Thus, some hold that the process of discovery inevitably splinters previously coherent fields of study, giving sub-fields an existence and a language of their own and creating barriers over which the practitioners no longer wish to jump. Others say that the content of science steadily becomes conceptually more difficult than can be handled by language as it is, which is a way of saying that the problem is not the fault of professional scientists but of other unspecified professionals. Still others protest that there would be no difficulty if only journals and their editors were more generous with the space they allow to the practitioners who are the creators of this growing volume of literature. All three defences are implausible.

That seems to be a legitimate inference from the opinions of regular readers of the journals, even journals such as this which aim to cater for readers with a variety of interests, both in particular but different fields of research and in more general matters. It is sobering how often those who kindly answer questionnaires about this weekly partial diet of science reply with the opinion that it must all be worthy, but that there are only a few parts of it that they could hope, and even wish, to understand. That there are conceptual difficulties that many people do not have the time to surmount is forgiveable. It is even proper that people working in some chosen field should choose to stay there, at least

for the time being. Why should everybody be required to be a generalist? But the sheer clumsiness of the scientific literature is a needless impediment not merely to wider understanding but even to the understanding of those at whom a specialized paper may be directed.

Although the common faults are not easily categorized (and people are forever inventing more), the general drift of error is generally understood even by the transgressors. There is, for example, a disinclination to believe that rules of grammar are particularly relevant to the understanding of scientific prose, whatever their significance in other connections. To judge from the literature as a whole, scientists care less for the accuracy with which the number of a pronoun agrees with that of its noun than they do for the numerical data they report, which is strange to say the least. It is also strange that people who fuss endlessly about the quality of reproduction of a diagram can so fecklessly scatter a word such as "only" in a sentence, forgetting that it belongs only where it most properly qualifies its antecedent.

To judge from much of what is published, most contributors must hope that inventiveness will see them (and others) through. Stringing long lists of words together was the strategem of the 1970s (as in "language-inventing-capability" or "DNA-binding-protein-sequence-homology"). The strategem serves no useful purpose because, until the string of words is so familiar within a sub-discipline that it can be represented simply by initials, it remains for the reader to unpackage the concatenation each time he comes across it. To how many is RFLP for "restriction-fragment-length-polymorphism" (or is it FLRP for "fragment-length-restriction-polymorphism"?) a concept to hold in one's head as clearly as that of, say, "rabbit"? Even so some journals have now gone so far as to grade the length of the hyphens by which the words are separated in such a string, as if it were always possible to tell which pairs of consecutive words are the more closely related.

Meanwhile, new forms of inventiveness have already begun to appear. Ribose is a sugar which, when combined appropriately with a purine or pyrimidine phosphate, will make a nucleotide; everybody knows that. So why not coin the verb "to ribosylate"? There can be no objection. But what is the object of this ribosylation? ATP, or GTP, or something quite differ-

ent? No problem. The statement suggests the solution. Why not "ATP-ribosylate", in whatever tense may be appropriate? The difficulty that unfamiliarity may often impede the understanding of even those who know what is intended will surely melt away in time, as the whole world takes the trouble to learn this special language?

The difficulty of the literature derives in part from the neglect of many common rules and the accompanying invention of others which are thought preferable, perhaps because they are better suited to the circumstances. But there are faults of a more structural character. The conventions of the research business, for example, require that authors' delight in what they have accomplished should be restricted to the use of conventional phrases, among which "for the first time" is unaccountably one. The result is that an author is prevented from giving an account of why his or her paper is important (which would help towards its understanding) but is allowed baldly to state that an observation has been made for "the first time in a closed vessel under water". Similarly, the convention that a person describing the results of a series of biochemical experiments should give such a full account of the details that a novice starting from scratch could repeat the work unaided is a pleasing concession to the need that work should be reproducible, but no substitute for an intelligent variation on some phrase such as "we followed so-and-so's method with this difference". Yet nobody seems to care that readers must each time unpackage each account of an experiment in the literature before they can properly understand its novelty.

Thus the underlying fault with the scientific literature is its intricacy, its requirement of readers that they should not merely read and understand, but read, pull to pieces and then put together again in ways they understand. Given the requirement, the process is plainly more prone to ambiguity than the ordinary person would think credible. And there is nothing in the demands of specialization, or in those made by editors on authors, to suggest that the characteristics of the scientific literature that give it a bad name are externally imposed. This, it seems, is what people want. But is not much of the rest of scholarship much the same? Of course.

John Maddox