CHAPTER 10

THE NUSAP CATEGORIES: ASSESSMENT AND PEDIGREE

The final pair of categories, assessment and pedigree, are the most obviously innovative features of the NUSAP scheme. They represent levels of uncertainty that go beyond those which can be managed by purely technical means. As such, they are the most qualitative categories in the scheme, and also the most flexible in their possible interpretations and associated notations.

They express conceptual operations of criticisms and evaluation on the entries in previous categories. They are reflective, and they foster self awareness among users. The operations related to assessment and pedigree cannot be performed by automatic means, neither by a formal calculus nor by a computer package. Nor can they be accomplished in isolation from the whole body of relevant scientific knowledge and craft skills. In this sense, they introduce philosophical considerations into the practice of research; but this is done in a constructive rather than in a sceptical way. The result of these reflective activities can be operationalized in special notations expressing the distinctions relevant to any particular case. These notations enable the development of conventions for a useful "arithmetic of uncertainty", for the combination of assessment ratings, which are gauges (with a very coarse topology) in the sense discussed in Chapter 7.

We can exhibit the structure of the NUSAP approach by a slight modification of the original string of five places. We separate it into three sections, consisting first of numeral, then unit and spread, and finally assessment and pedigree. The first section corresponds to pure mathematics; that is, an uninterpreted calculus, as common arithmetic standing alone, or algebra. The second section describes the application of the mathematics, as in measurement or estimation. The unit entry relates number to operations through topology and scale, and spread conveys the ever-present technical uncertainty of the operation. Thus, this middle section modifies the first, and conceptually it may be said to include it, since it provides the extension of the arithmetic to a larger system relating to empirical operations.

Finally, the third section of the NUSAP categories is the most general and inclusive, introducing the contexts of production and of use. The former includes the state-of-the-art of the relevant field of scientific production, and thereby conveys the border with ignorance, the deepest sort of uncertainty. The latter, the context of use, includes the policy process in which the quantity is used as an input. It extends even to the general culture in which it is received, and which in the last resort shapes our concepts of science and of knowledge.
10.1. ASSESSMENT

Assessment comes after spread in the string of places in NUSAP; and in its simplest interpretations is closely related to it. In statistical practice, confidence limits will be naturally expressed as an assessment entry. Indeed, in many technical contexts, the assessment category may be considered as nearly an alternative form of spread, and hence redundant. Thus if there is a normal distribution involved in a test, then the confidence limits relate in an entirely straightforward way to the variance; and little extra information may be conveyed by having two entries instead of one.

This close association in an extreme case tends to conceal the radical difference between the two categories. By permitting an impression of similarity between them, it inhibits the understanding of either. Those experienced in statistics are the most prone to make this easy and misleading identification. Even within statistical practice, any but the least demanding examples, illustrate the independence of the two categories. For instance, we may consider a statistical process depending on a number of trials, as a coin-toss or a modelling exercise. Suppose we are interested in the value corresponding to the 95th percentile of the resulting distribution; this will appear in numeral, while assessment will contain the entry %95 (to distinguish the entry from the traditional 95% confidence). The spread entry is then a measure of inexactness, a function of the number of trials. In this way, the meaning of the two terms are independent of each other.

The distinction between spread and assessment is more easily appreciated in terms of the different sorts of “error” encountered in experimental research. We have already mentioned random error, the spread of values obtained through measurement. Contrasted to this is the systematic error, which is estimated in terms of the historic experience of that class of experiments. We have discussed this distinction in the case of the fine-structure constant $\alpha$. Judgements of a still more qualitative character may be expressed concisely in assessment. Particularly in the context of policy-related research, the data may be so sparse and inexact that no meaningful assessment entry can be derived mathematically. Then qualitative judgements are employed on a scale, from “strong” to “weak” or “good” to “poor”. For convenience, these may have a numerical code associated with them, understood to be a gauge and not a measure or even an estimate.

The judgements about quality, or reliability, may refer to a broader context than the purely cognitive one of production. If a NUSAP expression is being prepared for a particular use in the policy context, then a pragmatically-oriented interpretation of assessment may be appropriate. The “same” quantity may then be given with two (or more) alternative forms, thus displaying the different meanings which can be conveyed for different functions, by the NUSAP scheme.