

## CHAPTER 1

# *Background*

Towards the end of the Second World War, in the Admiralty's radar establishment, I found myself trying to follow the behaviour of electrical pulses over extremely short intervals of time. Inevitably, one came up against fundamental physical limits to the accuracy of measurement. Typically, these limits seemed to be related in a complementary way, so that one of them could be widened only at the expense of a narrowing of another. An increase in time-resolving power, for example, seemed always to be bought at the expense of a reduction in frequency-resolving power; an improvement in signal-to-noise ratio was often inseparable from a reduction in time-resolving power, and so on. The art of physical measurement seemed to be ultimately a matter of compromise, of choosing between reciprocally related uncertainties.

I was struck by a possible analogy between this situation and the one in atomic physics expressed by Heisenberg's well-known 'Principle of Uncertainty'. This states that the momentum ( $p$ ) and position ( $q$ ) of a particle can never be exactly determined at the same instant. The smaller the imprecision ( $\Delta p$ ) in  $p$ , the larger must be the imprecision ( $\Delta q$ ) in  $q$ , and vice versa. In fact, the product  $\Delta p \times \Delta q$  can never be less than Planck's Constant  $h$ , the 'quantum of action'. Action (energy  $\times$  time) is thus the fundamental physical quantity whose 'atomicity' underlies the compromise-relation expressed in Heisenberg's Principle.

It seemed natural to ask what would happen if one multiplied together the reciprocally varying quantities in various compromise-relations of the sort I had encountered in experimental measurement. Perhaps these also might reflect the invariance of some fundamental 'quantum'. But a quantum of what? Tentatively, I called it a quantum of *information*. Multiplying together the conjugate pairs of uncertainty-limits mentioned, however, I found that they formed invariant products of not one but two distinct kinds. In one group, the product was a dimensionless number of order 1. In the other, it was a small quantity with the dimensions of physical entropy. Each of these represented an irreducible limiting factor of experimentation, a 'quantum' more fundamental than either of the two 'uncertainties' of which it was the product.

What then distinguished the two groups? It turned out to be something very simple. The first group of limits, which formed a dimensionless product, were calculable *a priori* (without reference to the measurement actually made), from a specification of the instrument. The second group, which formed products with the dimensions of entropy, could be calculated only *a posteriori*, from a specification of what was *done* with the instrument.

This was in 1945/46. In Autumn 1946 I moved to King's College, London, to teach physics, and a colleague on whom I tried out these notions suggested that there might be a link between my 'quanta of information' and the 'atomic propositions' postulated in Wittgenstein's famous *Tractatus Logico-Philosophicus*. Although in the analysis of ordinary language Wittgenstein himself had now repudiated the atomistic approach of the *Tractatus*,<sup>40</sup> the field of scientific measurement seemed well suited to logical treatment on these lines; for it builds everything ideally upon simple assertions of what Eddington called 'coincidence-relations' and 'occupance-relations' between pointers and scales. At all events, it seemed worth while to explore the possibility

that the 'quantization of information' I had stumbled upon in an experimental context was connected with the logical 'atomicity' of the statements that would ideally represent the outcome of the experiment. Was it possible that by analysing the logical requirements for the making of a scientific statement one might find a rational connection between the two kinds of 'quanta'?

As will appear later in this volume, it seemed that one might. A measurement could be thought of as a process in which elementary physical events, each of some prescribed minimal significance, are grouped into conceptually distinguishable categories so as to delineate a certain form (for example, the image on a photographic plate or the wave form on an oscilloscope) with a given degree of precision. In principle, the notion was that each elementary event could justify the addition of one 'atomic fact' to the logical framework representing what took place in ultimate physical terms. For example, any observation pressed to its physical limits could be regarded as an occasion on which the thermodynamic balance of an instrument was disturbed to an extent that could be measured in units of physical entropy. Each of these units, then — suitably defined — could be considered to provide one elementary building-block for an abstract representation of what occurred, the total of such building-blocks being distributed among the various categories or 'degrees of freedom' made available by the instrument, in much the same way as unit events are distributed among the columns of a histogram. Thus the 'informational efficiency' of a given measurement could be estimated by the proportion of those elementary events involved that find themselves represented by atomic facts in a logically adequate statement of the result.

Here then was a possible clue to the meaning of my two kinds of 'quanta'. A scientific representation, viewed in this light, could have its 'size' specified in two complementary ways: (*a*) by enumerating its degrees of freedom, and (*b*)

by enumerating its atomic facts. Correspondingly, a representation could be *augmented* in two complementary ways: (a) by adding to its degrees of freedom or 'logical dimensionality' (number of logically distinguishable categories); (b) by adding to its total of atomic facts or 'weight of evidence' (number of elementary events represented). On inspection, it turned out that the two kinds of quantal limits found in physical measurement represented simply the minimal physical costs of augmenting a representation by *one unit* of the corresponding kind. One could thus speak of a 'unit of information' in either sense as 'that which validates one elementary addition to a logical form representing the result'; but in the first case each unit would add one additional *dimension* (conceptual category), whereas in the second each unit would add one additional *atomic fact*. In order to avoid confusion it seemed appropriate to distinguish the first as 'structural units' and the second as 'metrical units' of information.

It soon became clear that the idea of measuring information was not new. In 1946 Dennis Gabor, then with the B.T.H. Company in Rugby (England), published his classic paper<sup>6</sup> entitled 'Theory of Communication', in which the Fourier-transform theory used in wave-mechanics was applied to the frequency-time ( $f \cdot t$ ) domain of the communication engineer, with the suggestion that a signal occupying an elementary area of  $\Delta f \cdot \Delta t = \frac{1}{2}$  could be regarded as a 'unit of information', which he termed a 'logon'. Much earlier, in 1935, the statistician R. A. Fisher<sup>4</sup> had proposed a measure of the 'Information' in a statistical sample, which in the simplest case amounted to the reciprocal of the variance (see Appendix).

It was far from obvious, on first encountering these disparate concepts, whether they could be fitted in any sensible way into the same framework; but on reflection it became apparent that they were in fact examples of 'structural' and 'metrical' measures, respectively. Gabor's logons, each occupying an area  $\Delta f \cdot \Delta t = \frac{1}{2}$  in the  $f \cdot t$  plane, represented the logical dimensions of his communication signals. They belonged to the

same family as the 'structural units' that occupy an analogous elementary area (the Airy disc) in the focal plane of a microscope, or of a radar aerial. It thus seemed appropriate, with Gabor's blessing, to give the term 'logon-content' a more general definition, as the measure of the logical dimensionality of representations of any form, whether spatial or temporal.

Fisher's measure, which is additive for averaged samples, invited an equally straightforward interpretation as an index of 'weight of evidence'. If we define (arbitrarily but reasonably) a unit or quantum of metrical information (termed a 'metron') as the weight of evidence that gives a probability of  $\frac{1}{2}$  to the corresponding proposition, Fisher's 'amount of information' becomes simply proportional to the number of such units supplied by the evidence in question.

Gabor's and Fisher's measures of information were not, however, the only candidates in the field. In 1948, C. E. Shannon published his 'Mathematical Theory of Communication',<sup>31</sup> which proposed for communication signals a measure based on the statistical *improbability* of the signal. Since the logarithm of improbability is additive for independent signals, this obviously had attractive properties. The question was how it related to the 'structural' and 'metrical' measures already in being. Not unnaturally, there was a strong tendency in the late forties to regard all these measures as rivals for a single title, or as suggesting rival concepts of 'information'. A further question was where, if anywhere, the notion of 'meaning' fitted into the scheme of things. Shannon's analysis of the 'amount of information' in a signal, which disclaimed explicitly any concern with its meaning, was widely misinterpreted to imply that the engineers had defined a concept of information *per se* that was totally divorced from that of meaning.

By the time I had plucked up enough courage to publish something on the subject,<sup>12</sup> the outlines of a possible synthesis had begun to emerge, and this was given a trial airing in papers<sup>13</sup> prepared for our first London Symposium on

Information Theory in 1950. Although some aspects of these early explorations may be only of historical interest, a sample is reprinted as an Appendix at the end of the present collection.

Resurrecting and sorting over these twenty-year-old speculations, what strikes one most forcibly is the number of vaguely perceived lines of development, not all of them unpromising, that one has never followed up. The main deflecting factor, it must be confessed, was the lure of a new trail that originated in the conjunction of my interests in information theory and electronic computation. If the concept of information had these dual aspects, the structural and the metrical, what kind of computing mechanism, one wondered, would be best adapted to handle the most general possible transformations of information? In particular, what sort of mechanism must the human brain be, in order to deal as it does with the sort of thing that information is?

Some attempts to wrestle with these problems will appear in a companion volume to the present collection. I mention them now because by the end of 1950 the challenge of the most complex of all computing mechanisms had become my focal interest. A year among neurophysiologists in the United States (1951) completed the transition process; and, for good or ill, most of my remaining half-baked ideas in the field of 'pure' information theory were left to grow cold.

The bulk of the papers in the present volume were written after this change of research emphasis, and they reflect a corresponding preoccupation with information as represented and utilized in the brain and exchanged between human beings, rather than as formalized in logical patterns of elementary propositions. I have not lost hope of a fruitful link between physical and logical atomism in fundamental physics, and it should perhaps be emphasized that the discrediting of logical atomism in the domain of ordinary language has done nothing to diminish its relevance to theoretical cosmology; but the pleasures of matchmaking in this area must be deferred, if not now left to others.