HEN A STRANGER, hearing that I am a physicist, asks me in what area of physics I work, I generally reply that I work on the theory of elementary particles. Giving this answer always makes me nervous. Suppose that the stranger should ask, “What is an elementary particle?” I would have to admit that no one really knows.

Let me declare first of all that there is no difficulty in saying what is meant by a particle. A particle is simply a physical system that has no continuous degrees of freedom except for its total momentum. For instance, we can give a complete description of an electron by specifying its momentum, as well as its spin around any given axis, a quantity that in quantum mechanics is discrete rather than continuous. On the other hand, a system consisting of a free electron and a free proton is not a particle, because to describe it one has to specify the momenta of both the electron and the proton—not just their sum. But a bound state of an electron and a proton, such as a hydrogen atom in its state of lowest energy, is a particle. Everyone would agree that a hydrogen atom is not an elementary particle, but it is not always so easy to make this distinction, or even to say what it means.
For the first few decades of this century there did not seem to be any trouble in saying what is meant by an elementary particle. J. J. Thomson could use the electric field in a cathode-ray tube to pull electrons out of atoms, so atoms were not elementary. Nothing could be pulled or knocked out of electrons, so it seemed that electrons were elementary. When atomic nuclei were discovered in Ernest Rutherford’s laboratory in 1911, it was assumed that they were not elementary, partly because it was known that some radioactive nuclei emit electrons and other particles, and also because nuclear charges and masses could be explained by assuming that nuclei are composed of two types of elementary particles: light, negatively charged electrons and heavy, positively charged protons.

Even without a definite idea of what is meant by an elementary particle, the idea that all matter consists of just two types of elementary particle was pervasive and resilient in a way that is difficult to understand today. For instance, when neutrons were discovered by James Chadwick in 1932, it was generally assumed that they were bound states of protons and electrons. In his paper announcing the discovery, Chadwick offered the opinion: “It is, of course, possible to suppose that the neutron is an elementary particle. This view has little to recommend it at present, except the possibility of explaining the statistics of such nuclei as $N^{14}$.” (One might have thought this was a pretty good reason: molecular spectra had revealed that the $N^{14}$ nucleus is a boson, which is not possible if it is a bound state of protons and electrons.) It was the 1936 discovery of the charge independence of nuclear forces by Merle Tuve et al. that showed clearly that neutrons and protons have to be treated in the same way; if protons are elementary, then neutrons must be elementary too. Today, in speaking of protons and neutrons, we often lump them together as nucleons.

This was just the beginning of a great increase in the roster of so-called elementary particles. Muons were added to the list in 1937 (though their nature was not understood until later), and pions and strange particles in the 1940s. Neutrinos had been proposed by Wolfgang Pauli in 1930, and made part of beta-decay theory by Enrico Fermi in 1933, but were not detected until the Reines-Cowan experiment of 1955. Then in the late 1950s the use of particle accelerators and bubble chambers revealed a great number of new particles, including mesons of spin higher than 0 and baryons of spin higher than 1/2, with various values for charge and strangeness.

On the principle that—even if there are more than two types of elementary particles—there really should not be a great number of types, theorists speculated that most of these particles are composites of a few types of elementary particles. But such bound states would have to be bound very deeply, quite unlike atoms or atomic nuclei. For instance, pions are much lighter than nucleons and antinucleons, so if the pion were a bound state of a nucleon and an antinucleon, as proposed by Fermi and Chen-Ning Yang, then its binding energy would have to be large enough to cancel almost all of...
the mass of its constituents. The composite nature of such a particle would be far from obvious.

How could one tell which of these particles is elementary and which composite? As soon as this question was asked, it was clear that the old answer—that particles are elementary if you can’t knock anything out of them—was inadequate. Mesons come out when protons collide with each other, and protons and antiprotons come out when mesons collide with each other, so which is a composite of which? Geoffrey Chew and others in the 1950s turned this dilemma into a point of principle, known as “nuclear democracy,” which held that every particle may be considered to be a bound state of any other particles that have the appropriate quantum numbers. This view was reflected decades later in a 1975 talk to the German Physical Society by Werner Heisenberg, who reminisced that:

In the experiments of the fifties and sixties . . . many new particles were discovered with long and short lives, and no unambiguous answer could be given any longer to the question about what these particles consisted of, since this question no longer has a rational meaning. A proton, for example, could be made up of neutron and pion, or Lambda-hyperon and kaon, or out of two nucleons and an antinucleon; it would be simplest of all to say that a proton just consists of continuous matter, and all these statements are equally correct or equally false. The difference between elementary and composite particles has thus basically disappeared. And that is no doubt the most important experimental discovery of the last fifty years.

LONG BEFORE Heisenberg reached this rather exaggerated conclusion, a different sort of definition of elementary particle had become widespread. From the perspective of quantum field theory, as developed by Heisenberg, Pauli, and others in the period 1926–34, the basic ingredients of Nature are not particles but fields; particles such as the electron and photon are bundles of energy of the electron and the electromagnetic fields. It is natural to define an elementary particle as one whose field appears in the fundamental field equations—or, as theorists usually formulate these theories, in the Lagrangian of the theory. It doesn’t matter if the particle is heavy or light, stable or unstable—if its field appears in the Lagrangian, it is elementary; if not, not.

This is a fine definition if one knows the field equations or the Lagrangian, but for a long while physicists didn’t. A fair amount of theoretical work in the 1950s and 1960s went into trying to find some objective way of telling whether a given particle type is elementary or composite when the underlying theory is
We will not be able to say which particles are elementary until we have a final theory of force and matter.

not known. This turned out to be possible in certain circumstances in nonrelativistic quantum mechanics, where an elementary particle might be defined as one whose coordinates appear in the Hamiltonian of the system. For instance, a theorem due to the mathematician Norman Levinson shows how to count the numbers of stable non-elementary particles minus the number of unstable elementary particles in terms of changes in phase shifts as the kinetic energy rises from zero to infinity. The trouble with using this theorem is that it involves the phase shifts at infinite energy, where the approximation of nonrelativistic potential scattering clearly breaks down.

I worried about this a good deal in the 1960s, but all I could come up with was a demonstration that the deuteron is a bound state of a proton and neutron. This was not exactly a thrilling achievement—everyone had always assumed that the deuteron is a bound state—but the demonstration had the virtue of relying only on nonrelativistic quantum mechanics and low-energy neutron-proton scattering data, without any specific assumptions about the Hamiltonian or about what happens at high energy. There is a classic formula that gives the spin triplet s-wave neutron-proton scattering length in terms of the nucleon mass and the deuteron binding energy, but the derivation of this formula actually relies on the assumption that the deuteron is a bound state. If we assume instead that the free-particle part of the Hamiltonian contains an elementary deuteron state, then this formula for the scattering length becomes incorrect, and instead we get a formula for the scattering length in terms of the nucleon mass, the deuteron binding energy, and the fraction of the time that the deuteron spends as an elementary particle (that is, the absolute value squared of the matrix element between the physical deuteron state and the elementary free-deuteron state). Comparing this formula with experiment showed that the deuteron spends most of its time as a composite particle. Unfortunately, arguments of this sort cannot be extended to deeply bound states, such as those encountered in elementary particle physics.

The lack of any purely empirical way of distinguishing composite and elementary particles does not mean that this distinction is not useful. In the 1970s the distinction between elementary and composite particles seemed to become much clearer, with the general acceptance of a quantum field theory of elementary particles known as the Standard Model. It describes quark, lepton, and gauge fields, so these are the elementary particles: six varieties or “flavors” of quarks, each coming in three colors; six flavors of leptons, including the electron; and twelve gauge bosons, including the photon, eight gluons, and the $W^+$, $W^-$, and $Z^0$ particles. The proton and neutron and all of the hundreds of mesons and baryons discovered after World War II are not elementary after all; they are composites of quarks and gluons, not because we can knock quarks and gluons out of them, which is believed to be impossible, but because that is the way they appear in the theory.

The one uncertain aspect of the Standard Model is the mechanism that breaks the electroweak gauge symmetry and gives the W and Z particles their masses, thereby adding an extra helicity state to what would have been the two helicities of a massless W or Z particle of spin 1. Theories of electroweak symmetry breakdown fall into two categories, according to whether these extra helicity states are elementary, as in the original form of the Standard Model, or composite, as in so-called technicolor theories. In a sense, the prime task driving the design of both the Large Hadron Collider and the ill-fated SSC was to settle the question of whether the extra helicity states of the W and Z particles are elementary or composite particles.

This might have been the end of the story, but since the late 1970s our understanding of quantum field theory has taken another turn. We have come to understand that particles may be described at sufficiently low energies by fields appearing in so-called effective quantum field theories, whether or not these particles are truly elementary. For instance, even though nucleon and pion fields do not appear in the Standard Model, we can calculate the rates for processes involving low-energy pions and nucleons by using an effective quantum field theory that involves pion and nucleon fields rather than quark and
gluon fields. In this field theory pions and nucleons are elementary, though nuclei are not. When we use a field theory in this way, we are simply invoking the general principles of relativistic quantum theories, together with any relevant symmetries; we are not really making any assumption about the fundamental structures of physics.

From this point of view, we are entitled only to say that the quarks and gluons are more elementary than nucleons and pions, because their fields appear in a theory, the Standard Model, that applies over a much wider range of energies than the effective field theory that describes nucleons and pions at low energy. We cannot reach any final conclusion about the elementarity of the quarks and gluons themselves. The Standard Model itself is probably only an effective quantum field theory, which serves as an approximation to some more fundamental theory whose details would be revealed at energies much higher than those available in modern accelerators, and which may not involve quark, lepton, or gauge fields at all.

One possibility is that the quarks and leptons and other particles of the Standard Model are themselves composites of more elementary particles. The fact that we see no structure in the quarks and leptons only tells us that the energies involved in their binding must be quite large—larger than several trillion electron volts. But so far no one has worked out a convincing theory of this sort.

We will not be able to give a final answer to the question of which particles are elementary until we have a final theory of force and matter. When we have such a theory, we may find that the elementary structures of physics are not particles at all. Many theorists think that the fundamental theory is something like a superstring theory, in which quarks, leptons, etc. are just different modes of vibration of the strings. It seems impossible in principle to identify one set of strings as truly elementary, because, as recently realized, different string theories with different types of strings are often equivalent.

There is a lesson in all this. The task of physics is not to answer a set of fixed questions about Nature, such as deciding which particles are elementary. We do not know in advance what are the right questions to ask, and we often do not find out until we are close to an answer.