PART 2

how freedom is not just chance but, rather, the result of a subtle interplay between *something almost random or haphazard*, and *something like a restrictive or selective control*—such as an aim or a standard—though certainly not a cast-iron control. For it is clear that the controls which guided Compton back from Italy allowed him plenty of freedom: freedom, say, to choose between an American and a French or Italian boat; or freedom to postpone his lecture, had some more important obligation arisen.

We may say that Compton's postulate of freedom restricts the acceptable solutions of our two problems by demanding that they should conform to *the idea of combining freedom and control*, and also to *the idea of a 'plastic control'*, as I shall call it in contradistinction to a 'cast-iron control'.

Compton's postulate is a restriction which I accept gladly and freely; and my own free and deliberate though not uncritical acceptance of this restriction may be taken as an illustration of that combination of freedom and control which is the very content of Compton's postulate of freedom.

XIII

I HAVE explained our two central *problems*—Compton's problem and Descartes's problem. In order to solve them we need, I believe, a *new theory*; in fact, a new theory of evolution, and a new model of the organism.

This need arises because the existing indeterministic theories are unsatisfactory. They are indeterministic; but we know that indeterminism is not enough, and it is not clear how they escape from Schlick's objection, or whether they conform to Compton's postulate of *freedom plus control*. Again, Compton's problem is quite beyond them: they are hardly relevant to it. And although these theories are attempts to solve Descartes's problem, the solutions they propose do not appear to be satisfactory.

The theories I am alluding to may be called 'master-switch models of control' or, more briefly, 'master-switch theories'. Their underlying idea is that our body is a kind of machine which can be regulated by a lever or switch from one or more *central control points*. Descartes even went so far as to locate the control point precisely: it is in the pineal gland, he said, that mind acts upon body. Some quantum theorists suggested (and

Compton very tentatively accepted the suggestion) that our minds work upon our bodies by influencing or selecting some quantum jumps. These are then amplified by our central nervous system which acts like an electronic amplifier: the amplified quantum jumps operate a cascade of relays or master-switches and ultimately effect muscular contractions.[44] There are, I think, some indications in Compton's books that he did not much like this particular theory or model, and that he used it for one purpose only: to show that human indeterminism (or even 'freedom') does not necessarily contradict quantum physics.[45] I think he was right in all this, including his dislike of master-switch theories.

For these master-switch theories—whether the one of Descartes, or the amplifier theories of the quantum physicists—belong to what I may perhaps call '*tiny baby theories*'. They seem to me to be almost as unattractive as tiny babies.

I am sure you all know the story of the unmarried mother who pleaded: 'But it is only a *very* tiny one.' Descartes's pleading seems to me similar: 'But it is such a tiny one: it is only an unextended mathematical point in which our mind may act upon our body.'

The quantum theorists hold a very similar tiny baby theory: 'But it is only with *one* quantum jump, and just within the Heisenberg uncertainties—and these are very tiny indeed—that a mind can act upon a physical system.' I admit that there is perhaps a slight advance here, in so far as the size of the baby is specified. But I still do not love the baby.

For however tiny the master-switch may be, the master-switch-*cum*-amplifier model strongly suggests that all our decisions are either snap-decisions (as I have called them in section x above) or else composed of snap-decisions. Now I admit that amplifier mechanisms are important characteristics of biological

[44] Compton discussed this theory in some detail, especially in *The Freedom of Man*, pp. 37–65. See especially the reference to Ralph Lillie, op. cit., in *The Freedom of Man*, p. 50. See also *The Human Meaning of Science*, pp. 47–54. Of considerable interest are Compton's remarks, in *The Freedom of Man*, pp. 63 f., and *The Human Meaning of Science*, p. 53, on *the character of individuality of our actions*, and his explanation of why it allows us to avoid what I may call the second horn of the dilemma (whose first horn is pure determinism), that is, the possibility that our actions are due to *pure chance*; cp. note 40.

[45] See especially *The Human Meaning of Science*, pp. viii f., and p. 54, the last statement of the section.

systems (for the energy of the reaction, released or triggered by a biological stimulus, usually exceeds greatly the energy of the triggering stimulus;[46] and I also admit, of course, that snap-decisions do occur. But they differ markedly from the kind of decision which Compton had in mind: they are almost like reflexes, and thus conform neither to the situation of Compton's problem of the influence of the universe of meanings upon our behaviour, nor to Compton's postulate of freedom (nor to the idea of a 'plastic' control). Decisions which conform to all this are as a rule reached almost imperceptibly through lengthy *deliberation*. They are reached by a kind of *maturing* process which is not well represented by the master-switch model.

By considering this process of deliberation, we may get another hint for our new theory. For deliberation always works by *trial and error* or, more precisely, by *the method of trial and of error-elimination*: by tentatively proposing various possibilities, and eliminating those which do not seem adequate. This suggests that we might use in our new theory some mechanism of trial and error-elimination.

I shall now outline how I intend to proceed.

Before formulating my evolutionary theory in general terms I shall first show how it works in a particular case, by applying it to our first problem, that is, to Compton's problem of the *influence of meaning upon behaviour*.

After having in this way solved Compton's problem, I shall formulate the theory in a general way. Then it will be found that it also contains—within the framework of our new theory which creates a new problem-situation—a straightforward and almost trivial answer to Descartes's classical body–mind problem.

[46] This is a point of great importance, so much so that we should hardly describe any process as typically biological unless it involved the release or triggering of stored energy. But the opposite is of course not the case: many non-biological processes are of the same character; and though amplifiers and release processes did not play a great role in classical physics, they are most characteristic of quantum physics and of course of chemistry. (Radioactivity with a triggering energy equal to zero is an extreme case; another interesting case is the—in principle adiabatic—tuning-in to a certain radio frequency, followed by the extreme amplification of the signal or stimulus.) This is one of the reasons why such formulae as 'the cause equals the effect' (and, with it, the traditional criticism of Cartesian interactionism) have long become obsolete, in spite of the continuing validity of the conservation laws. Cp. note 43, and the *stimulating or releasing* function of language, discussed in section XIV below; see also my book, *Conjectures and Refutations*, p. 381.

XIV

LET us now approach our first problem—that is, Compton's problem of the influence of meaning upon behaviour—by way of some comments on *the evolution of languages from animal languages to human languages*.

Animal languages and human languages have many things in common, but there are also differences: as we all know, human languages do somehow transcend animal languages.

Using and extending some ideas of my late teacher Karl Bühler[47] I shall distinguish two functions which animal and human languages share, and two functions which human language alone possesses; or in other words, two lower functions, and two higher ones which evolved on the basis of the lower functions.

The two lower functions of language are these. First, language, like all other forms of behaviour, consists of *symptoms or expressions*; it is symptomatic or expressive of the state of the organism which makes the linguistic signs. Following Bühler, I call this the *symptomatic or expressive function of language*.

Secondly, for language or communication to take place, there must not only be a sign-making organism or a 'sender', but also a reacting one, a 'receiver'. The symptomatic *expression* of the first organism, the sender, releases or evokes or stimulates or triggers a reaction in the second organism, which *responds* to the sender's behaviour, thereby turning it into a *signal*. This function of language to act upon a receiver was called by Bühler the *releasing or signalling function of language*.

To take an example, a bird may be ready to fly away, and may *express* this by exhibiting certain symptoms. These may then *release or trigger* a certain response or reaction in a second bird, and as a consequence it too may get ready to fly away.

[47] The theory of the functions of language is due to Karl Bühler (*The Mental Development of the Child*, 1919, English translation 1930, pp. 55, 56, 57; also *Sprachtheorie*, 1934). I have added to his three functions the argumentative function (and some other functions that play no role here, such as a hortative and a persuasive function). See for example my paper 'Language and the Body–Mind Problem', in *Conjectures and Refutations*, p. 295, note 2 and text. (See also pp. 134 f.) It is not impossible that there exist in animals, especially in bees, transition stages to some descriptive languages; see K. von Frisch, *Bees: their Vision, Chemical Senses, and Language*, 1950; *The Dancing Bees*, 1955; and M. Lindauer, *Communication Among Social Bees*, 1961.

Note that the two functions, the expressive function and the release function, are *distinct*; for it is possible that instances of the first may occur without the second, though not the other way round: a bird may express by its behaviour that it is ready to fly away without thereby influencing another bird. So the first function may occur without the second; which shows that they can be disentangled in spite of the fact that, in any genuine instance of communication by language, they always occur together.

These two lower functions, the symptomatic or expressive function on the one hand, and the releasing or signalling function on the other, are common to the languages of animals *and* men; and these two lower functions are always present when any of the higher functions (which are characteristically human) are present.

For human language is very much richer. It has many functions, and dimensions, which animal languages do not have. Two of these new functions are most important for the evolution of reasoning and rationality: the *descriptive function*, and the *argumentative function*.

As an example of the descriptive function, I might now describe to you how two days ago a magnolia was flowering in my garden, and what happened when snow began to fall. I might thereby express my feelings, and also release or trigger some feeling in you: you may perhaps react by thinking of *your* magnolia trees. So the two lower functions would be present. But *in addition* to all this, I should have described to you some facts; I should have made some *descriptive statements*; and these statements of mine would be factually *true*, or factually *false*.

Whenever I speak I cannot help expressing myself; and if you listen to me you can hardly help reacting. So the lower functions are *always* present. The descriptive function *need not* be present, for I may speak to you without describing any fact. For example, in showing or expressing uneasiness—say, doubt about whether you will survive this long lecture—I need not describe anything. Yet description, including the description of conjectured states of affairs, which we formulate in the form of theories or hypotheses, is clearly an extremely important function of human language; and it is that function which distinguishes human language most clearly from the various animal languages (although there seems

to be something approaching it in the language of the bees[48]). It is, of course, a function which is indispensable for science.

The last and highest of the four functions to be mentioned in this survey is the *argumentative function of language*, as it may be seen at work, in its highest form of development, in a well-disciplined *critical discussion*.

The argumentative function of language is not only the highest of the four functions I am here discussing, but it was also the latest of them to evolve. Its evolution has been closely connected with that of an argumentative, critical, and rational attitude; and since this attitude has led to the evolution of science, we may say that the argumentative function of language has created what is perhaps the most powerful tool for biological adaptation which has ever emerged in the course of organic evolution.

Like the other functions, the art of critical argument has developed by the method of trial and error-elimination, and it has had the most decisive influence on the human ability to think rationally. (Formal logic itself may be described as an 'organon of critical argument'.[49]) Like the descriptive use of language, the argumentative use has led to the evolution of ideal standards of control, or of '*regulative ideas*' (using a Kantian term): the main regulative idea of the descriptive use of language is *truth* (as distinct from *falsity*); and that of the argumentative use of language, in critical discussion, is *validity* (as distinct from *invalidity*).

Arguments, as a rule, are for or against some proposition or descriptive statement; this is why our fourth function—the argumentative function—must have emerged later than the descriptive function. Even if I argue in a committee that the University ought not to authorize a certain expenditure because we cannot afford it, or because some alternative way of using the money would be more beneficial, I am arguing not only for or against a *proposal* but also for and against some *proposition*—*for* the proposition, say, that the proposed use will not be beneficial, and *against* the proposition that the proposed use will be beneficial. So arguments, even arguments about proposals, as

[48] Cp. the books by Frisch, op. cit., and Lindauer, op. cit.

[49] See my book *Conjectures and Refutations*, chapter 1, especially the remark on p. 64 on formal logic as 'the *organon of rational criticism*'; also chapters 8 to 11, and chapter 15.

a rule bear on propositions, and very often on *descriptive* propositions.

Yet the argumentative use of language may be clearly distinguished from its descriptive use, simply because I can describe without arguing: I can describe, that is to say, without giving reasons for or against the truth of my description.

Our analysis of four functions of our language—the expressive, the signalling, the descriptive, and the argumentative functions—may be summed up by saying that, although it must be admitted that the two lower functions—the expressive and signalling functions—are *always* present whenever the higher functions are present, we must nevertheless distinguish the higher functions from the lower ones.

Yet many behaviourists and many philosophers have overlooked the higher functions, apparently because the lower ones are always present, whether or not the higher ones are.

XV

APART from the new functions of language which have evolved and emerged together with man, and with human rationality, we must consider another distinction of almost equal importance, the distinction between the evolution of *organs* and that of *tools or machines*, a distinction to be credited to one of the greatest of English philosophers, Samuel Butler, the author of *Erewhon* (1872).

Animal evolution proceeds largely, though not exclusively, by the modification of organs (or behaviour) or the emergence of new organs (or behaviour). *Human evolution* proceeds, largely, by developing new organs *outside our bodies or persons*: 'exosomatically', as biologists call it, or 'extra-personally'. These new organs are tools, or weapons, or machines, or houses.

The rudimentary beginnings of this exosomatic development can of course be found among animals. The making of lairs, or dens, or nests, is an early achievement. I may also remind you that beavers build very ingenious dams. But man, instead of growing better eyes and ears, grows spectacles, microscopes, telescopes, telephones, and hearing aids. And instead of growing swifter and swifter legs, he grows swifter and swifter motor cars.

Yet the kind of extra-personal or exosomatic evolution which

interests me here is this: instead of growing better memories and brains, we grow paper, pens, pencils, typewriters, dictaphones, the printing press, and libraries.

These add to our language—and especially to its descriptive and argumentative functions—what may be described as new dimensions. The latest development (used mainly in support of our argumentative abilities) is the growth of computers.

XVI

HOW are the higher functions and dimensions related to the lower ones? They do not replace the lower ones, as we have seen, but they establish a kind of *plastic control* over them—a control with feed-back.

Take, for example, a discussion at a scientific conference. It may be exciting and enjoyable, and give rise to expressions and symptoms of its being so; and these expressions in their turn may release similar symptoms in other participants. Yet there is no doubt that up to a point these symptoms and releasing signals will be due to, and controlled by, the scientific *content* of the discussion; and since this will be *of a descriptive and of an argumentative nature*, the lower functions will be controlled by the higher ones. Moreover, though a good joke or a pleasant grin may let the lower functions win in the short run, what counts in the long run is a good argument—a valid argument—and what it establishes or refutes. In other words, our discussion is controlled, though plastically, by the regulative ideas of truth and of validity.

All this is strengthened by the discovery and development of the new dimensions of printing and publishing, especially when these are used for printing and publishing scientific theories and hypotheses, and papers in which these are critically discussed.

I cannot do justice to the importance of critical arguments here: it is a topic on which I have written fairly extensively,[50] and so I shall not raise it again here. I only wish to stress that critical arguments are *a means of control*: they are a means of eliminating errors, a means of selection. We *solve our problems* by

[50] See note 49, and my book *The Open Society and its Enemies*, especially chapter 24 and the *Addendum* to vol. ii (fourth edn., 1962); and *Conjectures and Refutations*, especially the preface and the introduction.

tentatively proposing various competing theories and hypotheses, as trial balloons, as it were; and by submitting them to critical discussion and to empirical tests, for the purpose of error-elimination.

So the evolution of the higher functions of language which I have tried to describe may be characterized as the evolution of new means for problem-solving, by new kinds of trials, and by new methods of error-elimination; that is to say, new methods for *controlling* the trials.

XVII

I CAN now give my solution to our first main problem, that is, Compton's problem of the influence of meaning upon behaviour. It is this.

The higher levels of language have evolved under the pressure of a need for the *better control* of two things: of our lower levels of language, and our adaptation to the environment, by the method of growing not only new tools, but also, for example, new scientific theories, and new standards of selection.

Now in developing its higher functions, our language has also grown abstract meanings and contents; that is to say, we have learned how to abstract from the various modes of formulating or expressing a theory, and how to pay attention to its *invariant content or meaning* (upon which its truth depends). And this holds not only for theories and other descriptive statements, but also for proposals, or aims, or whatever else may be submitted to critical discussion.

What I have called 'Compton's problem' was the problem of explaining and understanding the controlling power of meanings, such as the contents of our theories, or of purposes, or aims; purposes or aims which in some cases we may have adopted after deliberation and discussion. But this is now no longer a problem. Their power of influencing us is part and parcel of these contents and meanings; for part of the function of contents and meanings is to control.

This solution of Compton's problem conforms to Compton's restricting postulate. For the control of ourselves and of our actions by our theories and purposes is a *plastic* control. We are not *forced* to submit ourselves to the control of our theories, for

we can discuss them critically, and we can reject them freely if we think that they fall short of our regulative standards. So the control is far from one-sided. Not only do our theories control us, but we can control our theories (and even our standards): there is a kind of *feed-back* here. And if we submit to our theories, then we do so freely, after deliberation; that is, after the critical discussion of alternatives, and after freely choosing between the competing theories, in the light of that critical discussion.

I submit this as my solution of Compton's problem; and before proceeding to solve Descartes's problem, I shall now briefly outline the more general theory of evolution which I have already used, implicitly, in my solution.

XVIII

I OFFER my general theory with many apologies. It has taken me a long time to think it out fully, and to make it clear to myself. Nevertheless I still feel far from satisfied with it. This is partly due to the fact that it is an *evolutionary* theory, and one which adds only a little, I fear, to existing evolutionary theories, except perhaps a new emphasis.

I blush when I have to make this confession; for when I was younger I used to say very contemptuous things about evolutionary philosophies. When twenty-two years ago Canon Charles E. Raven, in his *Science, Religion and the Future*, described the Darwinian controversy as 'a storm in a Victorian teacup', I agreed, but criticized him[51] for paying too much attention 'to the vapours still emerging from the cup', by which I meant the hot air of the evolutionary philosophies (especially those which told us that there were inexorable laws of evolution). But now I have to confess that this cup of tea has become, after all, *my* cup of tea; and with it I have to eat humble pie.

Quite apart from evolutionary *philosophies*, the trouble about evolutionary *theory* is its tautological, or almost tautological, character: the difficulty is that Darwinism and natural selection, though extremely important, explain evolution by 'the survival of the fittest' (a term due to Herbert Spencer). Yet there does not seem to be much difference, if any, between the assertion

[51] Cp. p. 106, note 1, of my book *The Poverty of Historicism*.

'those that survive are the fittest' and the tautology 'those that survive are those that survive'. For we have, I am afraid, no other criterion of fitness than actual survival, so that we conclude from the fact that some organisms have survived that they were the fittest, or those best adapted to the conditions of life.

This shows that Darwinism, with all its great virtues, is by no means a perfect theory. It is in need of a restatement which makes it less vague. The evolutionary theory which I am going to sketch here is an attempt at such a restatement.

My theory may be described as an attempt to apply to the whole of evolution what we learned when we analysed the evolution from animal language to human language. And it consists of a certain *view of evolution* as a growing hierarchical system of plastic controls, and of a certain *view of organisms* as incorporating—or in the case of man, evolving exosomatically—this growing hierarchical system of plastic controls. The Neo-Darwinist theory of evolution is assumed; but it is restated by pointing out that its 'mutations' may be interpreted as more or less accidental trial-and-error gambits, and 'natural selection' as one way of controlling them by error-elimination.

I shall now state the theory in the form of twelve short theses:

(1) All *organisms* are constantly, day and night, *engaged in problem-solving*; and so are all those evolutionary *sequences of organisms*—the *phyla* which begin with the most primitive forms and of which the now living organisms are the latest members.

(2) These problems are problems in an objective sense: they can be, hypothetically, reconstructed by hindsight, as it were. (I will say more about this later.) Objective problems in this sense need not have their conscious counterpart; and where they have their conscious counterpart, the conscious problem need not coincide with the objective problem.

(3) Problem-solving always proceeds by the method of trial and error: new reactions, new forms, new organs, new modes of behaviour, new hypotheses, are tentatively put forward and controlled by error-elimination.

(4) Error-elimination may proceed either by the complete elimination of unsuccessful forms (the killing-off of unsuccessful forms by natural selection) or by the (tentative) evolution of controls which modify or suppress unsuccessful organs, or forms of behaviour, or hypotheses.

(5) The single organism telescopes[52] into one body, as it were, the controls developed during the evolution of its *phylum*—just as it partly recapitulates, in its ontogenetic development, its phylogenetic evolution.

(6) The single organism is a kind of spearhead of the evolutionary sequence of organisms to which it belongs (its *phylum*): it is itself a tentative solution, probing into new environmental niches, choosing an environment and modifying it. It is thus related to its *phylum* almost exactly as the actions (behaviour) of the individual organism are related to this organism: the individual organism, and its behaviour, are both trials, which may be eliminated by error-elimination.

(7) Using 'P' for problem, 'TS' for tentative solutions, 'EE' for error-elimination, we can describe the fundamental evolutionary sequence of events as follows:

$$P \rightarrow TS \rightarrow EE \rightarrow P.$$

But this sequence is not a cycle: the second problem is, in general, different from the first: it is the result of the new situation which has arisen, in part, because of the tentative solutions which have been tried out, and the error-elimination which controls them. In order to indicate this, the above schema should be rewritten:

$$P_1 \rightarrow TS \rightarrow EE \rightarrow P_2.$$

(8) But even in this form an important element is still missing: the multiplicity of the tentative solutions, the multiplicity of the trials. Thus our final schema becomes something like this:

$$\begin{array}{ccccccc} & \nearrow & TS_1 & \searrow & & & \\ P_1 & \rightarrow & TS_2 & \rightarrow & EE & \rightarrow & P_2 \\ & \searrow & \cdot & \nearrow & & & \\ & & \cdot & & & & \\ & & \cdot & & & & \\ & & \cdot & & & & \\ & & \cdot & & & & \\ & & TS_n & & & & \end{array}$$

Background Knowledge

[52] The idea of 'telescoping' (though not this term which I owe to Alan Musgrave) may perhaps be found in chapter vi of Charles Darwin's *The Origin of Species*, 1859 (I am quoting from the Mentor Book edn., p. 180; italics mine): '. . . every highly developed organism has passed through many changes; and . . . each modified structure tends to be inherited, so that each modification will not . . . be quite lost. . . . *Hence the structure of each part* [of the organism] . . . *is the sum* of many

(9) In this form, our schema can be compared with that of Neo-Darwinism. According to Neo-Darwinism there is in the main *one* problem: the problem of survival. There is, as in our system, a multiplicity of tentative solutions—the variations or mutations. But there is only *one* way of error-elimination—the killing of the organism. And (partly for this reason) the fact that P_1 and P_2 will differ essentially is overlooked, or else its fundamental importance is not sufficiently clearly realized.

(10) In our system, not all problems are survival problems: there are many very specific problems and sub-problems (even though the earliest problems may have been sheer survival problems). For example an early problem P_1 may be reproduction. Its solution may lead to a new problem, P_2: the problem of getting rid of, or of spreading, the offspring—the children which threaten to suffocate not only the parent organism but each other.[53]

It is perhaps of interest to note that *the problem of avoiding suffocation by one's offspring* may be one of those problems which was solved by the evolution of *multicellular organisms*: instead of getting rid of one's offspring, one establishes a *common economy*, with various new methods of living together.

(11) The theory here proposed distinguishes between P_1 and P_2, and it shows that the problems (or the problem situations) which the organism is trying to deal with are often *new*, and arise themselves as products of the evolution. The theory thereby gives implicitly a rational account of what has usually been called by the somewhat dubious names of '*creative evolution*' or '*emergent evolution*'.[54]

(12) Our schema allows for the development of error-eliminating controls (warning organs like the eye; feed-back mechanisms); that is, controls which can eliminate errors without killing the organism; and it makes it possible, ultimately, for our hypotheses to die in our stead.

inherited changes, through which the species has passed. . . .' See also E. Baldwin in the book, *Perspectives in Biochemistry*, pp. 99 ff., and the literature there quoted.

[53] The emergence of a new problem-situation could be described as a change or a differentiation of the 'ecological niche', or the significant environment, of the organism. (It may perhaps be called a 'habitat selection'; cp. B. Lutz, *Evolution*, 2, 1948, pp. 29 ff.) The fact that *any* change in the organism *or* its habits *or* its habitat produces new problems accounts for the incredible wealth of the (always tentative) solutions.

[54] See note 23 for reference to Compton's remarks on 'emergent evolution'.

XIX

EACH organism can be regarded as a hierarchical system of *plastic controls*—as a system of clouds controlled by clouds. The controlled subsystems make trial-and-error movements which are partly suppressed and partly restrained by the controlling system.

We have already met an example of this in the relation between the lower and higher functions of language. The lower ones continue to exist and to play their part; but they are constrained and controlled by the higher ones.

Another characteristic example is this. If I am standing quietly, without making any movement, then (according to the physiologists) my muscles are constantly at work, contracting and relaxing in an almost random fashion (see TS_1 to TS_n in thesis (8) of the preceding section), but controlled, without my being aware of it, by error-elimination (*EE*) so that every little deviation from my posture is almost at once corrected. So I am kept standing, quietly, by more or less the same method by which an automatic pilot keeps an aircraft steadily on its course.

This example also illustrates the thesis (1) of the preceding section—that each organism is all the time engaged in problem-solving by trial and error; that it reacts to new and old problems by more or less chance-like,[55] or cloud-like, trials which are eliminated if unsuccessful. (If successful, they increase the probability of the survival of mutations which 'simulate' the solutions so reached, and tend to make the solution hereditary,[56] by incorporating it into the spatial structure or form of the new organism.)

[55] The method of trial and error-elimination *does not operate with completely chance-like or random trials* (as has been sometimes suggested), even though the trials may look pretty random; there must be at least an 'after-effect' (in the sense of my *The Logic of Scientific Discovery*, pp. 162 ff.). For the organism is constantly learning from its mistakes, that is, it establishes *controls* which suppress or eliminate, or at least reduce the frequency of, certain *possible* trials (which were perhaps *actual* ones in its evolutionary past).

[56] This is now sometimes called the 'Baldwin Effect'; see for example, G. G. Simpson, 'The Baldwin Effect', *Evolution*, 7, 1953, pp. 110 ff., and C. H. Waddington, the same volume, pp. 118 ff. (see especially p. 124), and pp. 386 f. See also J. Mark Baldwin, *Development and Evolution*, 1902, pp. 174 ff. and H. S. Jennings, *The Behaviour of the Lower Organisms*, 1906, pp. 321 ff.

XX

THIS is a very brief outline of the theory. It needs, of course, much elaboration. But I wish to explain *one* point a little more fully—the use I have made (in theses (1) to (3) of section XVIII) of the terms '*problem*' and '*problem-solving*' and, more particularly, my assertion that *we can speak of problems in an objective, or non-psychological sense.*

The point is important, for evolution is clearly not a conscious process. Many biologists say that the evolution of certain organs solves certain problems; for example, that the evolution of the eye solves the problem of giving a moving animal a timely warning to change its direction before bumping into something hard. Nobody suggests that this kind of solution to this kind of problem is consciously sought. Is it not, then, just a metaphor if we speak of problem-solving?

I do not think so; rather, the situation is this: when we speak of a problem, we do so almost always from hindsight. A man who works on a problem can seldom say clearly what his problem is (unless he has found a solution); and even if he can explain his problem, he may mistake it. And this may even hold of scientists—though scientists are among those few who consciously try to be fully aware of their problems. For example, Kepler's conscious problem was to discover the harmony of the world order; but we may say that the problem he solved was the mathematical description of motion in a set of two-body planetary systems. Similarly, Schrödinger was mistaken about the problem he had solved by finding the (time-independent) Schrödinger equation: he thought his waves were charge-density waves, of a changing continuous field of electric charge. Later Max Born gave a statistical interpretation of the Schrödinger wave amplitude; an interpretation which shocked Schrödinger and which he disliked as long as he lived. He had solved a problem—but it was not the one he thought he had solved. This we know now, by hindsight.

Yet clearly it is in science that we are most conscious of the problems we try to solve. So it should not be inappropriate to use hindsight in other cases, and to say that the amoeba solves some problems (though we need not assume that it is in any sense aware of its problems): from the amoeba to Einstein is just one step.

XXI

BUT Compton tells us that the amoeba's actions are not rational,[57] while we may assume that Einstein's actions are. So there should be some difference, after all.

I admit that there is a difference: even though their methods of almost random or cloud-like trial and error movements are fundamentally not very different,[58, 55] there is a great difference in their attitudes towards error. Einstein, unlike the amoeba, consciously tried his best, whenever a new solution occurred to him, to fault it and detect an error in it: he approached his own solutions *critically*.

I believe that this consciously critical attitude towards his own ideas is the one really important difference between the method of Einstein and that of the amoeba. It made it possible for Einstein to reject, quickly, hundreds of hypotheses as inadequate before examining one or another hypothesis more carefully, if it appeared to be able to stand up to more serious criticism.

As the physicist John Archibald Wheeler said recently, 'Our whole problem is to make the mistakes as fast as possible'.[59] This problem of Wheeler's is solved by consciously adopting the critical attitude. This, I believe, is the highest form so far of the rational attitude, or of rationality.

The scientist's trials and errors consist of hypotheses. He formulates them in words, and often in writing. He can then try to find flaws in any one of these hypotheses, by criticizing it, and by testing it experimentally, helped by his fellow scientists who will be delighted if they can find a flaw in it. If the hypothesis does not stand up to these criticisms and to these tests at least as well as its competitors,[60] it will be eliminated.

It is different with primitive man, and with the amoeba. Here

[57] See *The Freedom of Man*, p. 91, and *The Human Meaning of Science*, p. 73.

[58] Cp. H. S. Jennings, op. cit., pp. 334 f., 349 f. A beautiful example of a problem-solving fish is described by K. Z. Lorenz, *King Solomon's Ring*, 1952, pp. 37 f.

[59] John A. Wheeler, *American Scientist*, 44, 1956, p. 360.

[60] That we can only choose the 'best' of a set of competing hypotheses—the 'best' in the light of a critical discussion devoted to the search for truth—means that we choose the one which appears, in the light of the discussion, to come 'nearest to the truth'; see my *Conjectures and Refutations*, chapter 10. See also *The Freedom of Man*, pp. vii f., and especially p. 74 (on the principle of conservation of energy).

there is no critical attitude, and so it happens more often than not that natural selection eliminates a mistaken hypothesis or expectation by eliminating those organisms which hold it, or believe in it. So we can say that the critical or rational method consists in letting our hypotheses die in our stead: it is a case of exosomatic evolution.

XXII

HERE I may perhaps turn to a question which has given me much trouble although in the end I arrived at a very simple solution.

The question is: Can we show that plastic controls exist? Are there inorganic physical systems in nature which may be taken as examples or as physical models of plastic controls?

It seems that this question was implicitly answered in the negative by many physicists who, like Descartes or Compton, operate with master-switch models, and by many philosophers who, like Hume or Schlick, deny that anything intermediate between complete determinism and pure chance can exist. Admittedly, cyberneticists and computer engineers have more recently succeeded in constructing computers made of hardware but incorporating highly plastic controls; for example, computers with built-in mechanism for chance-like trials, checked or evaluated by feed-back (in the manner of an automatic pilot or a self-homing device) and eliminated if erroneous. But these systems, although incorporating what I have called plastic controls, consist essentially of complex relays of master-switches. What I was seeking, however, was a simple physical model of Peircean indeterminism; a purely physical system resembling a very cloudy cloud in heat motion, controlled by some other cloudy clouds—though by somewhat less cloudy ones.

If we return to our old arrangement of clouds and clocks, with a cloud on the left and a clock on the right, then we could say that what we are looking for is something intermediate, like an organism or like our cloud of gnats, but not alive: a pure physical system, controlled plastically and 'softly', as it were.

Let us assume that the cloud to be controlled is a gas. Then we can put on the extreme left an uncontrolled gas which will soon diffuse and so cease to constitute a physical *system*. We put on the extreme right an iron cylinder filled with gas: this is our example of a 'hard' control, a 'cast-iron' control. In between, but far to the left, are many more or less 'softly' controlled systems, such as our cluster of gnats, and huge balls of particles, such as a gas kept together by gravity, somewhat like the sun. (We do not mind if the control is far from perfect, and many particles escape.) The planets may perhaps be said to be cast-iron controlled in their movements—comparatively speaking, of course, for even the planetary system is a cloud, and so are all the milky ways, star clusters, and clusters of clusters. But are there, apart from organic systems and those huge systems of particles, examples of any 'softly' controlled small physical systems?

I think there are, and I propose to put in the middle of our diagram a child's balloon or, perhaps better, a soap bubble; and this, indeed, turns out to be a very primitive and in many respects an excellent example or model of a Peircean system *and* of a 'soft' kind of plastic control.

The soap bubble consists of two subsystems which are both clouds and which control each other: without the air, the soapy film would collapse, and we should have only a drop of soapy water. Without the soapy film, the air would be uncontrolled: it would diffuse, ceasing to exist as a system. Thus the control is mutual; it is plastic, and of a feed-back character. Yet it is possible to make a distinction between the controlled system (the air) and the controlling systems (the film): the enclosed air is not only more cloudy than the enclosing film, but it also ceases to be a physical (self-interacting) system if the film is removed. As against this, the film, after removal of the air, will form a droplet which, though of a different shape, may still be said to be a physical system.

Comparing the bubble with a 'hardware' system like a precision clock or a computer, we should of course say (in accordance with Peirce's point of view) that even these hardware systems are clouds controlled by clouds. But these 'hard' systems are built with the purpose of minimizing, so far as it is possible, the cloud-like effects of molecular heat motions and fluctuations: though they are clouds, the controlling mechanisms are designed to suppress, or compensate for, all cloud-like effects as far as possible. This holds even for computers with mechanisms simulating chance-like trial-and-error mechanisms.

Our soap bubble is different in this respect and, it seems, more similar to an organism: the molecular effects are not eliminated but contribute essentially to the working of the system which is enclosed by a skin—a permeable wall[61] that leaves the system 'open', and able to 'react' to environmental influences in a manner which is built, as it were, into its 'organization': the soap bubble, when struck by a heat ray, absorbs the heat (much like a hot-house), and so the enclosed air will expand, keeping the bubble floating.

As in all uses of similarity or analogy we should, however, look out for limitations; and here we might point out that, at least in some organisms, molecular fluctuations are apparently amplified and so used to release trial-and-error movements. At any rate, amplifiers seem to play important roles in all organisms (which in this respect resemble some computers with their master-switches and cascades of amplifiers and relays). Yet there are no amplifiers in the soap bubble.

However this may be, our bubble shows that natural physical cloud-like systems which are plastically and softly controlled by other cloud-like systems do exist. (Incidentally, the film of the bubble need not, of course, be derived from organic matter, though it will have to contain large molecules.)

XXIII

THE evolutionary theory here proposed yields an immediate solution to our second main problem—the classical Cartesian body–mind problem. It does so (without saying *what* 'mind' or 'consciousness' is) by saying something about the evolution, and thereby about the functions, of mind or consciousness.

We must assume that consciousness grows from small beginnings; perhaps its first form is a vague feeling of irritation, experienced when the organism has a problem to solve such as getting away from an irritant substance. However this may be,

61 Permeable walls or membranes seem to be characteristic of all biological systems. (This may be connected with the phenomenon of biological individuation.) For the pre-history of the idea that membranes and bubbles are primitive organisms, see C. H. Kahn, *Anaximander*, 1960, pp. 111 ff.

consciousness will assume evolutionary significance—and increasing significance—when it begins to *anticipate* possible ways of reacting: possible trial-and-error movements, and their possible outcomes.

We can say now that conscious states, or sequences of conscious states, may function as systems of control, of error-elimination: the elimination, as a rule, of (incipient) behaviour, that is (incipient) movement. Consciousness, from this point of view, appears as just one of many interacting kinds of control; and if we remember the control systems incorporated for example in books—theories, systems of law, and all that constitutes the 'universe of meanings'—then consciousness can hardly be said to be the highest control system in the hierarchy. For it is to a considerable extent controlled by these exosomatic linguistic systems—even though they may be said to be *produced* by consciousness. Consciousness in turn is, we may conjecture, *produced* by physical states; yet it controls them to a considerable extent. Just as a legal or social system is produced by us, yet controls us, and is in no reasonable sense 'identical' to or 'parallel' with us, but *interacts* with us, so states of consciousness (the 'mind') control the body, and *interact* with it.

Thus there is a whole set of analogous relationships. As our exosomatic world of meanings is related to consciousness, so consciousness is related to the behaviour of the acting individual organism. And the behaviour of the individual organism is similarly related to its body, to the individual organism taken as a physiological system. The latter is similarly related to the evolutionary sequence of organisms—the *phylum* of which it forms the latest spearhead, as it were: as the individual organism is thrown up experimentally as a probe by the *phylum* and yet largely controls the fate of the *phylum*, so the behaviour of the organism is thrown up experimentally as a probe by the physiological system and yet controls, largely, the fate of this system. Our conscious states are similarly related to our behaviour. They anticipate our behaviour, working out, by trial and error, its likely consequences; thus they not only control but they try out, *deliberate*.

We now see that this theory offers us an almost trivial answer to Descartes's problem. Without saying *what 'the mind' is*, it leads immediately to the conclusion that our *mental states control* (*some* *of*) *our physical movements*, and that there is some give-and-take, some feed-back, and so some *interaction*, between mental activity and the other functions of the organism.[62]

The control will again be of the 'plastic' kind; in fact all of us—especially those who play a musical instrument such as the piano or the violin—know that the body does not always do what we want it to do; and that we have to learn, from our ill-success, how to modify our aims, making allowances for those limitations which beset our control: though we are free, to some considerable extent, there are always conditions—physical or otherwise—which set limits to what we can do. (Of course, before giving in, we are free to try to transcend these limits.)

Thus, like Descartes, I propose the adoption of a dualistic outlook, though I do *not* of course recommend talking of *two kinds of interacting substances*. But I think it is helpful and legitimate to distinguish *two kinds of interacting states* (or events), physio-chemical and mental ones. Moreover, I suggest that if we distinguish only these two kinds of states we still take too narrow a view of our world: at the very least we should also distinguish those artifacts which are products of organisms, and especially the products of our minds, and which can interact with our minds and thus with the state of our physical environment. Although these artifacts are often 'mere bits of matter', 'mere tools' perhaps, they are even on the animal level sometimes consummate works of art; and on the human level, the products of our minds are often very much more than 'bits of matter'—marked bits of paper, say; for these bits of paper may represent states of a discussion, states of the growth of knowledge, which may transcend (sometimes with serious consequences) the grasp of most or even all of the minds that helped to produce them. Thus we have to be not merely dualists, but pluralists; and we have to recognize that the great changes which we have brought about, often unconsciously, in our

[62] As hinted in several places, I conjecture that the acceptance of an '*interaction*' of mental and physical states offers the only satisfactory solution of Descartes's problem; see also note 43. I wish to add here that I think that we have good reason to assume that there exist mental states, or conscious states (for example in dreams) in which the consciousness of the ego (or of one's spatio-temporal position and identity) is very weak, or absent. It seems therefore reasonable to assume that full consciousness of the ego is a late development, and that it is a mistake to formulate the body-mind problem in such a way that this form of consciousness (or conscious 'will') is treated as if it were the only one.

physical universe show that abstract rules and abstract ideas, some of which are perhaps only partially grasped by human minds, may move mountains.

XXIV

As an afterthought, I should like to add one last point.

It would be a mistake to think that, because of natural selection, evolution can only lead to what may be called 'utilitarian' results: to adaptations which are useful in helping us to survive.

Just as, in a system with plastic controls, the controlling and controlled subsystems interact, so our tentative solutions interact with our *problems* and also with our *aims*. This means that our aims can change and that *the choice of an aim may become a problem*; different aims may compete, and new aims may be invented and controlled by the method of trial and error-elimination.

Admittedly, if a new aim clashes with the aim of surviving, then this new aim may be eliminated by natural selection. It is well known that many mutations are lethal and thus suicidal; and there are many examples of suicidal aims. Others are perhaps neutral with respect to survival.

Many aims that at first are subsidiary to survival may later become autonomous, and even opposed to survival; for example, the ambition to excel in courage, to climb Mount Everest, to discover a new continent, or to be the first on the Moon; or the ambition to discover some new truth.

Other aims may from the very beginning be autonomous departures, independent of the aim to survive. Artistic aims are perhaps of this kind, or some religious aims, and to those who cherish them they may become much more important than survival.

All this is part of the superabundance of life—the almost excessive abundance of trials and errors upon which the method of trial and error-elimination depends.[63]

It is perhaps not uninteresting to see that artists, like scientists, actually use this trial-and-error method. A painter may put down, tentatively, a speck of colour, and step back for a

[63] Cp. for example my *Conjectures and Refutations*, especially p. 312.

critical assessment of its effect[64] in order to alter it if it does not solve the problem he wants to solve. And it may happen that an unexpected or accidental effect of his tentative trial—a colour speck or brush stroke—may change his problem, or create a new subproblem, or a new aim: the evolution of artistic aims and of artistic standards (which, like the rules of logic, may become exosomatic systems of control) proceeds also by the trial-and-error method.

We may perhaps here look back for a moment to the problem of physical determinism, and to our example of the deaf physicist who had never experienced music but would be able to 'compose' a Mozart opera or a Beethoven symphony, simply by studying Mozart's or Beethoven's bodies and their environments as physical systems, and predicting where their pens would put down black marks on lined paper. I presented these as unacceptable consequences of physical determinism. Mozart and Beethoven are, partly, controlled by their 'taste', their system of musical evaluation. Yet this system is not cast iron but rather plastic. It responds to new ideas, and it can be modified by new trials and errors—perhaps even by an accidental mistake, an unintended discord.[65]

In conclusion, let me sum up the situation.

We have seen that it is unsatisfactory to look upon the world as a closed physical system—whether a strictly deterministic system or a system in which whatever is not strictly determined is simply due to chance: on such a view of the world human creativeness and human freedom can only be illusions. The attempt to make use of quantum-theoretical indeterminacy is also unsatisfactory, because it leads to chance rather than freedom, and to snap-decisions rather than deliberate decisions.

[64] See, for example, Ernst H. Gombrich, *Meditations on a Hobby Horse*, 1963, especially p. 10; and the same author's *Art and Illusion*, 1960, 1962 (see the Index under 'trial and error'). Cp. also note 65.

[65] For the close similarity of scientific and artistic production see *The Freedom of Man*, Preface, pp. vii f., and the remark in *The Freedom of Man*, p. 74, referred to in note 60 above; further E. Mach, *Wärmelehre*, 1896, pp. 440 f., where he writes: 'The history of art . . . teaches us how shapes which arise accidentally may be used in works of art. Leonardo da Vinci advises the artist to look for shapes of clouds or patches on dirty or smoky walls, which might suggest to him ideas that fit in with his plans and his moods. . . . Again, a musician may sometimes get new ideas from random noises; and we may hear on occasion from a famous composer that he has been led to find valuable melodic or harmonic motifs by accidentally touching a wrong key while playing the piano.'

I have therefore offered here a different view of the world—one in which the physical world is an open system. This is compatible with the view of the evolution of life as a process of trial and error-elimination; and it allows us to understand rationally, though far from fully, the emergence of biological novelty and the growth of human knowledge and human freedom.

I have tried to outline an evolutionary theory which takes account of all this and which offers solutions to Compton's and Descartes's problems. It is, I am afraid, a theory which manages to be too humdrum *and* too speculative at the same time; and even though I think that testable consequences can be derived from it, I am far from suggesting that my proposed solution is what philosophers have been looking for. But I feel that Compton might have said that it presents, in spite of its faults, a possible answer to his problem—and one which might lead to further advance.