BIOLOGICAL AND MENTAL EVOLUTION: AN EXERCISE IN ANALOGY*

By ARTHUR KOESTLER

ALLOW me to take you on a ride on the treacherous wings of analogy, starting with an excursion into genetics. Creativity is a concept notoriously difficult to define; and it is sometimes useful to approach a difficult subject by way of contrast. The opposite of the creative individual is the pedant, the slave of habit, whose thinking and behaviour move in rigid grooves. His biological equivalent is the over-specialized animal. Take, for example, that charming and pathetic creature, the koala bear, which specializes in feeding on the leaves of a particular variety of eucalyptus tree and on nothing else; and which, in lieu of fingers, has hook-like claws, ideally suited for clinging to the bark of the tree—and for nothing else. Some of our departments of higher learning seem expressly designed for breeding koala bears.

Sir Julian Huxley has described over-specialization as the principal cause why evolution in all branches of the animal kingdom—except man's—seems to have ended either in stagnation or in extinction. But, having made his point, he drew a conclusion which you may find less convincing. "Evolution," he concluded, "is thus seen as an enormous number of blind alleys with a very occasional path to progress. It is like a maze in which almost all turnings are wrong turnings[1]." With due respect, I think this metaphor is suspiciously close to the old-fashioned behaviourist's views of the rat in the maze as a paradigm of human learning. In both cases the explicit or tacit assumption is that progress results from a kind of blind man's buff—random mutations preserved by natural selection, or random tries preserved by reinforcement—and that that is all there is to it. However, it is possible to dissent from this view without invoking a *deus ex machina*, or a Socratic *daimon*, by making the simple assumption that, while random events no doubt play an important part in the picture, that is not all there is to it.

One line of escape from the maze is indicated by a phenomenon known to students of evolution by the ugly name of paedomorphism, a term coined by Garstang[2] some forty years ago. The existence of the phenomenon is well established; but there is little mention of it in the text-books, perhaps because it runs against the *Zeitgeist*. It indicates that in certain circumstances evolution can re-trace its steps, as it were, along the path which led to the dead-end and make a fresh start in a more promising direction. To put it simply, paedomorphism means the appearance of some evolutionary novelty in the larval or embryonic stage of the ancestral animal, a novelty which may disappear before the adult stage is reached, but which reappears in the adult descendant. This bit of evolutionary magic is made possible by the well-known mechanism of neoteny, that is to say, the gradual retardation of bodily development beyond the age of sexual maturity, with the result that breeding takes place while the animal still displays larval or juvenile features. Hardy[3], de Beer[4] and others have pointed out that if this tendency toward 'prolonged childhood' were accompanied by a corresponding squeezing out of the later adult stages of ontogeny, the result would be a rejuvenation and de-specialization of the race which would thus regain some of its lost adaptive plasticity. But of even greater importance than this re-winding of the biological clock is the fact that in the paedomorphic type of evolution selective pressure operates on the early, malleable stages of ontogeny. In contrast to this, gerontomorphism—the appearance of novel characters in the late-adult stages—can only modify structures which are already highly specialized. One is accordingly led to expect that the major evolutionary advances were due to paedomorphism and not to gerontomorphism—to changes in the larval or embryonic, and not in the adult, stage.

Let me give an example, which will make clearer what I am driving at. There is now strong evidence in favour of the theory, proposed by Garstang[2] in 1922, that the chordates, and thus we, the vertebrates, descended from the larval state of some primitive echinoderm, perhaps rather like the sea-urchin or sea-cucumber. Now an adult sea-cucumber would not be a very inspiring ancestor—it is a sluggish creature which looks like an ill-stuffed sausage, lying on the sea-bottom. But its free-floating larva is a much more promising proposition: unlike the adult, it has bilateral symmetry, a ciliary band presumed to be the forerunner of the neural fold, and other sophisticated features not found in the adult animal. We must assume that the sedentary adult residing on the sea-bottom had to rely on mobile larvae to spread the species far and wide in the ocean, as plants scatter their seeds in the wind; and that the larvae, which had to fend for themselves, exposed to much stronger selective pressures than the adults, gradually became more fish-like; and lastly became sexually mature while still in the free-swimming, larval state—thus giving rise to a new type of animal which never settled on the bottom at all and altogether eliminated the senile, sessile cucumber stage from its life-history.

It seems that the same re-tracing of steps to escape the dead-ends of the maze was repeated at each decisive evolutionary turning-point—the last time, so far as we know, when the line which bore our own species branched off from some ancestral primate. It is now generally recognized that the human adult resembles more the embryo of an ape than an adult ape. In both, the ratio of brain-weight to body-weight is disproportionately high; in both, the closing of the sutures of the skull is retarded to allow for further brain growth. The back to front axis through man's head—the direction of his line of sight—forms an angle of ninety degrees with his spinal column; a condition which, in apes and other mammals, is only found in the embryonic stage. The same applies to the angle between the uro-genital canal and the backbone, which accounts for the singularity of the human way of mating. Other embryonic—or, to use Bolk's[5] term, foetalized—features are the absence of brow-ridges, scantness of body-hair, retarded development of the teeth, and so on. As Haldane[6] has said: "If human evolution is to continue along the same lines as in the past, it will probably involve a still greater prolongation of childhood and retardation of maturity. Some of the characters distinguishing adult man will be lost." But there is a reverse to the medal, which Aldous Huxley gleefully showed us in *After Many A Summer*: artificial prolongation of the absolute life-span of man might provide an opportunity for features of the adult ape to re-appear in Methuselah. But this only by the way.

The essence of the process which I have described is a retreat from highly specialized adult forms of bodily structure and behaviour to an earlier, more plastic and less committed stage—followed by a sudden advance in a new direction. It is as if the stream of life had momentarily reversed its course, flowing uphill for a while, then opened up a new stream-bed—leaving the koala bear stranded on its tree like a discarded hypothesis. We have now reached the crucial point in our excursion,

* Substance of an address delivered on September 18 at the Bicentennia Celebration commemorating the birth of James Smithson, held in Washington during September 16–18 (see *Nature*, **208**, 320; 1965).

because it seems to me that this process of *reculer pour mieux sauter*—of drawing back to leap, of undoing and re-doing—is a basic feature of all significant progress, both in biological and mental evolution.

It can be shown, I think, that these two types of progress —the emergence of biological novelties and the creation of mental novelties—are analogous processes on different levels of the developmental hierarchy. But to demonstrate the connexion we must proceed stepwise from lower to higher organisms. One of the fundamental properties of living organisms is their power of self-repair, and the most dramatic manifestations of this power are the phenomena of regeneration (which Needham[7] called "one of the more spectacular pieces of magic in the repertoire of living organisms"). Primitive creatures, like flatworms, when cut into slices, can regenerate a whole animal from a tiny fragment; Amphibia can regenerate limbs and organs; and once more the 'magic' is performed by *reculer pour mieux sauter*—the regression of specialized tissues to a genetically less committed, quasi-embryonic stage, a de-differentiation or de-specialization followed by a re-differentiation.

Now the replacement of a lost limb or lost eye is a phenomenon of a quite different order from the adaptive processes in a normal environment. Regeneration could be called a meta-adaptation to traumatizing challenges. The power to perform such meta-adaptations manifests itself only when the challenge exceeds a critical limit and can only be met by having recourse to the genetic plasticity of the embryonic stage. We have just seen that the major phylogenetic changes were brought about by a similar retreat from adult to embryonic forms. Indeed, the main line of development which led up to our species could be described as a series of operations of phylogenetic self-repair: of escapes from blind alleys by the undoing and re-moulding of maladapted structures.

Evidently, self-repair by the individual produces no evolutionary novelty, it merely restores the *status quo ante*. But that is all the individual needs in order to regain its normal adaptive balance in a static environment (assuming that the traumatizing disturbance was only a momentary one). Phylogenetic 'self-repair', on the other hand, implies changes in the genotype to restore the adaptive balance in a changing environment.

As we move toward the higher animals, the power of regenerating physical structures is superseded by the equally remarkable power of the nervous system to re-organize its mode of function. (Ultimately, of course, these reorganizations must also involve structural changes of a fine-grained nature in terms of circuitry, molecular chemistry or both, and so we are still moving along a continuous line.) Lashley[8] taught his rats certain visual discrimination skills; when he removed their optical cortex, the learning was gone, as one would expect; but, contrary to what one would expect, the mutilated rats were able to learn the same tasks again. Some other brain area, not normally specializing in visual learning, must have taken over this function, deputizing for the lost area.

Similar feats of meta-adaptation have been reported in insects, birds, chimpanzees and so on. But let us get on to man, and to those lofty forms of self-repair which we call self-realization, and which include creativity in its broadest sense. Psycho-therapy, ancient and modern, from shamanism down to contemporary forms of abreaction therapy, has always relied on what Ernst Kris[9] has called "regression in the service of the ego". The neurotic with his compulsions, phobias and elaborate defence-mechanisms is a victim of maladaptive specialization—a koala bear hanging on for dear life to a barren telegraph pole. The therapist's aim is to regress the patient to an infantile or primitive condition; to make him retrace his steps to the point where they went wrong, and to come up again, metamorphosed, re-born. Goethe's *Stirb und Werde*, the inexhaustible variations of the archetype of death and resurrection, dark night and spiritual rebirth, all revolve around this basic paradigm—Joseph in the well, Jesus in the tomb, Buddha in the desert, Jonah in the belly of the whale.

There is no sharp dividing line between self-repair and self-realization. All creative activity is a kind of do-it-yourself therapy, an attempt to come to terms with traumatizing experiences. In the scientist's case the trauma is some apparent paradox of Nature, some anomaly in the motion of the planets, the sting of data which contradict each other, disrupt an established theory, and make nonsense of his cherished beliefs. In the artist's case, challenge and response are manifested in his tantalizing struggle to express the inexpressible, to conquer the resistance of his medium, to escape from the distortions and restraints imposed by the conventional styles and techniques of his time.

In other words, the so-called revolutions in the history of both science and art are successful escapes from blind alleys. The evolution of science is neither continuous nor strictly cumulative except for those periods of consolidation and elaboration which follow immediately after a major breakthrough. Sooner or later, however, the process of consolidation leads to increasing rigidity and orthodoxy, and so into the dead-end of over-specialization. The proliferation of esoteric jargons which seems to characterize this phase reminds one sometimes of the monstrous antlers of the Irish elk, and sometimes of the neurotic's elaborate defence-mechanisms against the threats of reality. Eventually, the process leads to a crisis, and thus to a new revolutionary break-through—followed by another period of consolidation, a new orthodoxy, and so the cycle starts again.

In the history of art, this cyclic process is even more obvious: periods of cumulative progress within a given school and technique end inevitably in stagnation, mannerism or decadence, until the crisis is resolved by a revolutionary shift in sensibility, emphasis, style.

Every revolution has a destructive and a constructive aspect. In science the destruction is wrought by jettisoning previously unassailable doctrines, including some seemingly self-evident axioms of thought. In art, it involves an equally agonizing re-appraisal of accepted values, criteria of relevance, frames of perception. When we discuss the evolution of art and science from the historian's detached point of view, this un-doing and re-doing process appears as a normal and inevitable part of the whole story. But when we focus our attention on any concrete individual who initiated a revolutionary change, we are immediately made to realize the immense intellectual and emotional obstacles he had to overcome. I mean not only the inertial forces of society; the primary locus of resistance against heretical novelty is inside the skull of the individual who conceives of it. It reverberates in Kepler's agonized cry when he discovered that the planets move in elliptical pathways: "who am I, Johannes Kepler, to destroy the divine symmetry of the circular orbits!". On a more down-to-earth level the same agony is reflected in Jerome Bruner's[10] experimental subjects who, when shown for a split second a playing card with a black queen of hearts, saw it as red, as it should be; and when the card was shown again, reacted with nausea at such a perversion of the laws of Nature. To unlearn is more difficult than to learn; and it seems that the task of breaking up rigid cognitive structures and reassembling them into a new synthesis cannot, as a rule, be performed in the full daylight of the conscious, rational mind. It can only be done by reverting to those more fluid, less committed and specialized forms of ideation which normally operate in the twilight below the level of focal awareness. Such intervention of unconscious processes in the creative act is now generally, if sometimes reluctantly, accepted even by behaviourists with a strong positivist bias. Allow me, therefore, to take it for granted that in the period of incubation—to use Graham Wallas's[11]

term—the creative individual experiences a temporary regression to patterns of thinking which are normally inhibited in the rational adult.

But it would be a gross over-simplification to identify—as is sometimes done—these patterns with Freud's so-called "Primary Process". The primary process is supposedly devoid of logic, governed by the pleasure principle, apt to confuse perception and hallucination, expressed in spontaneous action, and accompanied by massive affective discharge. I believe that between this very primary process, and the so-called secondary process governed by the reality principle, we must interpolate a whole hierarchy of cognitive structures which are not simply mixtures of primary and secondary processes, but are autonomous systems in their own right, each governed by a distinct set of rules. The paranoid delusion, the dream, the daydream, free association, the mentality of children at various ages and of primitives at various stages, should not be lumped together, for each has its own logic or rules of the game. But while clearly different in many respects, all these forms of ideation have certain features in common, since they are ontogenetically, and perhaps phylogenetically, older than those of the civilized adult. I have elsewhere[12] called them "games of the underground", because if not kept under restraint they would play havoc with the routines of disciplined thinking. But under exceptional conditions, when disciplined thinking is at the end of its tether, a temporary indulgence in these underground games may suddenly produce a solution which was beyond the reach of the conscious, rational mind—that new synthesis which Poincare[13] called the happy combination of ideas, and which I like to call 'bisociation' (as distinct from associative routine). I have discussed this process in some detail in a recent book[12] and shall not dwell on its intricate details. The point I want to make here is that the creation of novelty in mental evolution follows the same pattern of *reculer pour mieux sauter*, of a temporary regression to a naïve or juvenile level, followed by a forward leap, which we have found in biological evolution. We can carry the analogy further and interpret the Aha reaction, or 'Eureka!' cry, as the signal of a happy escape from a blind alley—an act of mental self-repair, achieved by the de-differentiation of cognitive structures to a more plastic state, and the resulting liberation of creative potentials—the equivalent of the release of genetic growth-potentials in regenerating tissues.

It is a truism to say that in mental evolution social inheritance replaces genetic inheritance. But there is a less trivial parallel between phylogenesis and the evolution of ideas: neither of them proceeds along a continuous curve in a strictly cumulative manner. Newton said that if he saw farther than others it was because he stood on the shoulders of giants. But did he really stand on their shoulders or some other part of their anatomy? He adopted Galileo's laws of free fall, but rejected Galileo's astronomy. He adopted Kepler's planetary laws, but demolished the rest of the Keplerian edifice. He did not take as his point of departure their completed 'adult' theories, but retraced their development to the point where it had gone wrong. Nor was the Keplerian edifice built on top of the Copernican structure. That ramshackle structure of epicycles he tore down and kept only its foundations. Nor did Copernicus continue to build where Ptolemy had left off. He went back two thousand years to Aristarchus. The great revolutionary turns in the evolution of ideas have a decidedly paedomorphic character. The new paradigm, to use Thomas Kuhn's[14] term, which emerges from the revolution is not derived from a previous adult paradigm; not from the aged sea-urchin but from its mobile larva, floating in the currents of the ocean. Only in the relatively peaceful periods of consolidation and elaboration do we find gerontomorphism—small improvements to a fully mature body of knowledge. In the history of art the process is again all too obvious; there is no need to elaborate on it.

I began with a wistful remark about the treacherous wings of analogy, aware of the fact that those who trust these waxen wings usually share the fate of Icarus. But it is one thing to argue from analogy, and quite another to point to an apparent similarity which has perhaps not been paid sufficient attention, and then to ask whether that similarity has some significance or whether it is trivial and deceptive. I believe that the parallel between certain processes underlying biological and mental evolution has some significance. Biological evolution could be described as a history of escapes from over-specialization, the evolution of ideas as a series of escapes from the bondage of mental habit; and the escape-mechanism in both cases is based on the same principles. We get an inkling of them through the phenomena of regeneration—the remoulding of structures and reorganization of functions—which only enter into action when the challenge exceeds a critical limit. They point to the existence of unsuspected 'meta-adaptive' potentials which are inhibited or dormant in the normal routines of existence, and, when revealed, make us sometimes feel that we move like sleepwalkers in a world of untapped resources and unexplored possibilities.

It could be objected that I have presented a reductionist view; that it is sacrilegious to call the creation of a Brahms symphony or of Maxwell's field equations an act of self-repair, and to compare it with the mutation of a sea-squirt larva, the regeneration of a newt-tail, the re-learning process in the rat or the rehabilitation of patients by psycho-therapy. But I think that such a view is the opposite of sacrilegious. It points, however tentatively, at a common denominator, a factor of purposiveness, without invoking a *deus ex machina*. It does not deny that trial and error are inherent in all progressive development. But there is a world of difference between the random tries of the monkey at the typewriter, and the process which I called, for lack of a better name, *reculer pour mieux sauter*. The first means reeling off all possible responses in the organism's repertory until the correct one is hit on by chance and stamped in by reinforcement. The second may still be called trial and error, but of a purposive kind, using more complex, sophisticated methods: a groping and searching, retreating and advancing towards a goal. "Purpose," to quote Herbert J. Muller[15], "is not imported into Nature and need not be puzzled over as a strange or divine something. . . . It is simply implicit in the fact or organisation." This directiveness of vital processes is present all along the line, from conscious behaviour down to what Needham[7] called "the striving of the blastula to grow into a chicken". How tenacious and resourceful that striving is has been demonstrated by experimental embryology, from Speeman to Paul Weiss—though its lessons have not yet been fully digested.

Thus to talk of goal-directedness or purpose in ontogeny has become respectable again. In phylogeny the monkey still seems to be hammering away at the typewriter, perhaps because the crude alternatives that had been offered—amorphous entelechies, or the Lysenko brand of Lamarckism—were even more repellent to the scientific mind. On the other hand, some evolutionary geneticists are beginning to discover that the typewriter is structured and organized in such a way as to defeat the monkey, because it will print only meaningful words and sentences. In recent years the rigid, atomistic concepts of Mendelian genetics have undergone a softening process and have been supplemented by a whole series of new terms with an almost holistic ring. Thus we learn that the genetic system represents a "micro-hierarchy" which exercises its selective and regulative control on the molecular, chromosomal and cellular level; that development is "canalized", stabilized by "developmental homeostasis" or "evolutionary homoestasis"[16] so that mutations affect not a single unit character but a "whole organ in a harmonious way"[17], and, finally, that these various forms of "internal

selection" create a restricted "mutation spectrum"[18] or may even have a "direct, moulding influence guiding evolutionary change along certain avenues"[19]—and all this happens long before external, Darwinian selection gets to work. But if this is the case, then the part played by a lucky chance mutation is reduced to that of the trigger which releases the co-ordinated action of the system; and to maintain that evolution is the product of blind chance means to confuse the simple action of the trigger, governed by the laws of statistics, with the complex, purposive processes which it sets off. Their purposiveness is manifested in different ways on different levels of the hierarchy, from the self-regulating properties of the genetic system through internal and external selection, culminating perhaps in the phenomena of phylogenetic self-repair: escapes from blind alleys and departures in new directions. On each level there is trial and error, but on each level it takes a more sophisticated form. Some twenty years ago, Tolman and Krechevsky[20] created a stir by proclaiming that the rat learns to run a maze by forming hypotheses; soon it may be permissible to extend the metaphor and to say that evolution progresses by making and discarding hypotheses.

Any directive process, whether you call it selective, adaptive or expectative, implies a reference to the future. The equifinality of developmental processes, the striving of the blastula to grow into an embryo, regardless of the obstacles and hazards to which it is exposed, might lead the unprejudiced observer to the conclusion that the pull of the future is as real and sometimes more important than the pressure of the past. The pressure may be compared to the action of a compressed spring, the pull to that of an extended spring, threaded on the axis of time. Neither of them is more or less mechanistic than the other. If the future is completely determined in the Laplacian sense, then there is nothing to choose between the actions of the two springs. If it is indeterminate in the Heisenbergian sense, then indeterminacy works in both directions, and the distant past is as blurred and unknowable as the future; and if there is something like a free choice operating within the air-bubbles in the stream of causality, then it must be directed towards the future and oriented by feed-back from the past.

[1] Huxley, J., *Man in the Modern World*, 13 (New York, 1948).
[2] Garstang, W., *J. Linnean Soc. Lond. (Zoology)*, **35**, 81 (1922).
[3] Hardy, A. C., in *Evolution as a Process* (New York, 1954).
[4] de Beer, G. R., *Embryos and Ancestors* (Oxford, 1940).
[5] Bolk, L., *Das Problem der Menschwerdung* (Jena, 1926).
[6] Haldane, J. B. S., *The Causes of Evolution*, 150 (London, 1932).
[7] Needham, A. E., *New Scientist*, London, November 2, 1961.
[8] Lashley, K. S., *Brain Mechanisms and Intelligence* (Chicago, 1929).
[9] Kris, E., *Psychoanalytic Explorations in Art* (New York, 1952).
[10] Bruner, J. S., and Postman, L., *J. of Personality*, XVIII (1949).
[11] Wallas, G., *The Art of Thought* (London, 1954).
[12] Koestler, A., *The Act of Creation* (New York, 1964).
[13] Poincare, H., in *The Creative Process* (Berkeley, Calif., 1952).
[14] Kuhn, T. H., *The Structure of Scientific Revolutions* (Chicago, 1962).
[15] Muller, H. J., *Science and Criticism* (New Haven, Conn., 1943).
[16] Cannon, H. G., *The Evolution of Living Things* (Manchester, 1958).
[17] Waddington, C. H., *The Listener*, London (Nov. 13, 1952).
[18] Spurway, H., in *Supplemento. La Ricerca Scientifica* (Pallanza Symp.), **18** (Cons. Naz. delle Richerche, Rome, 1949).
[19] For a survey of literature in this field see Whyte, L. L., *Internal Factors in Evolution* (London, 1965).
[20] Krechevsky, I., *Psychol. Rev.*, **39** (1932).

HISTORY AS THE ORGANIZATION OF MAN'S MEMORY*

By Prof. HERBERT BUTTERFIELD
Master of Peterhouse, University of Cambridge

WE often think of Western civilization as scientific, and we do not always remember that it is equally remarkable for being so historically minded. In both respects the only parallel to it is ancient China—so wonderful in its science and technology, but possessing also a historical literature of almost incredible vastness. Even in China there did not develop those modern techniques which, in our section of the globe, led to the scientific revolution of the seventeenth century and the somewhat parallel historiographical revolution in the nineteenth. In both fields the developments in Europe were unique; and the Chinese have had to become pupils of the West.

Some civilizations, like that of ancient India, seem to have been governed by religions or philosophies which deny significance to the facts of history as such, and the sequence of events in time. Yet, between a culture which has soaked itself in historical memories and one for which the past is only chance and change—only bubble and froth—there must develop great differences in general mentality, in intellectual habits, and in the degree of control that can be acquired over the course of events: and the differences must extend to still deeper things that affect the very nature of the human consciousness.

Our interest in the past—our very sense for the past—(like our prowess in the natural sciences) is therefore a thing that requires to be explained. Even the case that we to-day might make for the study of history would have no meaning for those earlier generations of mankind that gave the start to the whole endeavour. We of the twentieth century might say that a society is going to be very constricted in its development unless it looks behind itself, organizes its memory, reflects on its larger and longer experiences, learns to measure the direction in which it is moving, and gets some notion of long-term tendencies which have been observed. But this kind of diagnosis—this way of learning where we stand in the processes of time—is a thing that comes only late in the day, when civilization and scholarship itself have progressed very far. Nobody could have known in advance that by the study of the past we should be able to examine the processes of things in time. Indeed, until the world was fairly mature, nobody could even have guessed that there existed such things as historical processes which might call for analysis.

It is possible that, in every age and society, children will love to listen to the tales of a grandfather and will look back at least to the time just before they were born. All the world seems to love a story, and, even if there were no inferences to be drawn from it, we are all likely to be interested in the account by Arthur M. Schlesinger, jun., of the expedition to the Bay of Pigs—interested in it purely as the narrative of something that actually happened. I suspect that, however scientific and analytical and statistical historical scholarship may become in the latter half of the twentieth century, a great mass of people will go on loving this narrative history—the history that tells of men and their vicissitudes. Perhaps these are the people who will keep the subject sane—keep history an important humanistic factor in our civilization. But, though the telling of stories may awaken an interest in the past, it is not likely in itself to alter the structure of our mentality. Also, it is scarcely enough in itself to drive the mind to research and criticism and the passionate quest for truth. The *raconteur* knows

* Substance of an address delivered at the Bicentennial Celebration commemorating the birth of James Smithson, held in Washington during September 16–18 (see *Nature*, **208**, 320; 1965).