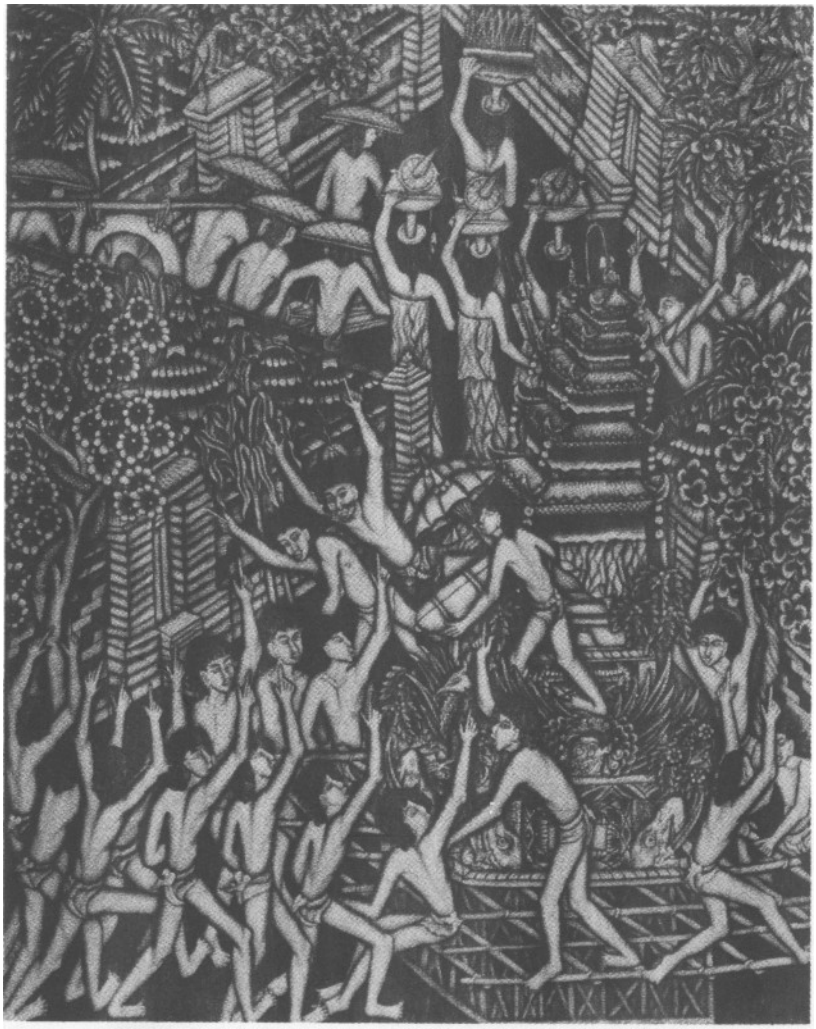
STEPS TO AN ECOLOGY OF MIND

COLLECTED ESSAYS IN ANTHROPOLOGY, PSYCHIATRY, EVOLUTION, AND EPISTEMOLOGY

Gregory Bateson

Jason Aronson Inc.
Northvale, New Jersey
London



Balinese Painting (Ida Bagus Djati Sura; Batuan, 1937)

[Analysis, p. 147]

10 9 8 7 6 5 4 3 2 1

Library of Congress Cataloging-in-Publication Data

Bateson, Gregory.
Steps to an ecology of mind.

(Chandler publications for health sciences)
Reprint. Originally published: San Franciso :
Chandler Pub. Co., 1972. With new pref.
Collection of articles reprinted from various
periodicals.
"The published work of Gregory Bateson": p.
Includes index.
1. Anthropology, Cultural—Collected works.
2. Knowledge, Theory of—Collected works. 3. Psychiatry — Collected works. 4. Evolution — Collected works.
I. Title. II. Series. [DNLM: 1. Anthropology, Cultural—collected works. 2. Ecology — collected works.
Evolution — collected works. 4. Schizophrenic
Psychology—collected works. 5. Thinking—collected
works. GN 6 B329s 1972a]
GN6.B3 1987 306 87-24133
ISBN 0-87668-950-0

CONTENTS

Part V: Epistemology and Ecology

Cybernetic Explanation[*]

It may be useful to describe some of the peculiarities of cybernetic explanation.

Causal explanation is usually positive. We say that billiard ball B moved in such and such a direction because billiard ball A hit it at such and such an angle. In contrast to this, cybernetic explanation is always negative. We consider what alternative possibilities could conceivably have occurred and then ask why many of the alternatives were not followed, so that the particular event was one of those few which could, in fact, occur. The classical example of this type of explanation is the theory of evolution under natural selection. Ac-cording to this theory, those organisms which were not both physiologically and environmentally viable could not possibly have lived to reproduce. Therefore, evolution always followed the pathways of viability. As Lewis Carroll has pointed out, the theory explains quite satisfactorily why there are no bread-and-butter-flies today.

In cybernetic language, the course of events is said to be subject to *restraints,* and it is assumed that, apart from such restraints, the pathways of change would be governed only by equality of probability. In fact, the "restraints" upon which cybernetic explanation depends can in all cases be regarded as factors which determine inequality of probability. If we find a monkey striking a typewriter apparently at random but in fact writing meaningful prose, we shall look for restraints, either inside the monkey or inside the typewriter. Perhaps the monkey could not strike inappropriate letters; perhaps the type bars could not move if improperly struck; perhaps incorrect letters could not survive on the paper. Somewhere there must have been a circuit which could identify error and eliminate it.

---

[*] This article is reprinted from the *American Behavioral Scientist, Vol. 10,* No. *8,* April 1967, pp. *29-32,* by per-mission of the publisher, Sage Publications, Inc.

Ideally—and commonly—the actual event in any sequence or aggregate is uniquely determined within the terms of the cybernetic explanation. Restraints of many different kinds may combine to generate this unique determination. For example, the selection of a piece for a given position in a jigsaw puzzle is "restrained" by many factors. Its shape must conform to that of its several neighbors and possibly that of the boundary of the puzzle; its color must conform to the color pattern of its region; the orientation of its edges must obey the topological regularities set by the cutting machine in which the puzzle was made; and so *on.* From the point of view of the man who is trying to solve the puzzle, these are all clues, i.e., sources of information which will guide him in his selection. From the point of view of the cybernetic observer, they are *restraints.*

Similarly, from the cybernetic point of view, a word in a sentence, or a letter within the word, or the anatomy of some part within an organism, or the role of a species in an ecosystem, or the behavior of a member within a family—these are all to be (negatively) explained by an analysis of restraints.

The negative form of these explanations is precisely comparable to the form of logical proof by *reductio ad absurdum.* In this species of proof, a sufficient set of mutually exclusive alternative propositions is enumerated, e.g., "P" and "not P," and the process of proof procedes by demonstrating that all but one of this set are untenable or "absurd." It follows that the surviving member of the set must be tenable within the terms of the logical system. This is a form of proof which the nonmathematical sometimes find unconvincing and, no doubt, the theory of natural selection sometimes seems unconvincing to nonmathematical persons for similar reasons—whatever those reasons may be.

Another tactic of mathematical proof which has its counterpart in the construction of cybernetic explanations is the use of "mapping" or rigorous metaphor. An algebraic proposition may, for example, be mapped onto a system of geometric coordinates and there proven by geometric methods. In cybernetics, mapping appears as a technique of explanation whenever a conceptual "model" is invoked or, more concretely, when a computer is used to simulate a complex communicational process. But this is not the only appearance of mapping in this science. Formal processes of mapping, translation,

or transformation are, in principle, imputed to *every* step of any sequence of phenomena which the cyberneticist is attempting to explain. These *mappings* or trans-formations may be very complex, e.g., where the output of some machine is regarded as a transform of the input; or they may be very simple, e.g., where the rotation of a shaft at a given point along its length is regarded as a transform (albeit identical) of its rotation at some previous point.

The relations which remain constant under such transformation may be of any conceivable kind.

This parallel, between cybernetic explanation and the tactics of logical or mathematical proof, is of more than trivial interest. Outside of cybernetics, we look for explanation, but not for anything which would simulate logical proof. This simulation of proof is something new. We can say, however, with hindsight wisdom, that explanation by simulation of logical or mathematical proof was expectable. After all, the subject matter of cybernetics is not events and objects but the *information* "carried" by events and objects. We consider the objects or events only as proposing facts, propositions, messages, percepts, and the like. The subject matter being propositional, it is expectable that explanation would simulate the logical.

Cyberneticians have specialized in those explanations which simulate *reductio ad absurdum* and "mapping." There are perhaps whole realms of explanation awaiting discovery by some mathematician who will recognize, in the informational aspects of nature, sequences which simulate other types of proof.

Because the subject matter of cybernetics is the propositional or informational aspect of the events and objects in the natural world, this science is forced to procedures rather different from those of the other sciences. The differentiation, for example, between map and territory, which the semanticists insist that scientists shall respect in their writings must, in cybernetics, be watched for in the very phenomena about which the scientist writes. Expectably, communicating organisms and badly programmed computers will mistake map for territory; and the language of the scientist must be able to cope with such anomalies. In human behavioral systems, especially in religion and ritual and wherever primary process

dominates the scene, the name often is the thing named. The bread is the Body, and the wine is the Blood.

Similarly, the whole matter of induction and deduction —and our doctrinaire preferences for one or the other—will take on a new significance when we recognize inductive and deductive steps not only in our own argument but in the relationships among data.

Of especial interest in this connection is the relationship between *context* and its content. A phoneme exists as such only in combination with other phonemes which make up a word. The word is the *context* of the phoneme. But the word only exists as such—only has "meaning"—in the larger context of the utterance, which again has meaning only in a relationship.

This hierarchy of contexts within contexts is universal for the communicational (or "emic") aspect of phenomena and drives the scientist always to seek for explanation in the ever larger units. It may (perhaps) be true in physics that the explanation of the macroscopic is to be sought in the microscopic. The opposite is usually true in cybernetics: without context, there is no communication.

In accord with the negative character of cybernetic ex-planation, "information" is quantified in negative terms. An event or-object such as the letter K in a given position in the text of a message *might* have been any other of the limited set of twenty-six letters in the English language. The actual letter excludes (i.e., eliminates by restraint) twenty-five alternatives. In comparison with an English letter, a Chinese ideograph would have excluded several thousand alternatives. We say, therefore, that the Chinese ideograph carries more information than the letter. The quantity of information is conventionally expressed as the log to base 2 of the improbability of the actual event or object.

Probability, being a ratio between quantities which have similar dimensions, is itself of zero dimensions. That is, the central explanatory quantity, information, is of zero dimensions. Quantities of real dimensions (mass, length, time) and their derivatives (force, energy, etc.) have no place in cybernetic explanation.

The status of energy is of special interest. In general in communicational systems, we deal with sequences which resemble

stimulus-and-response rather than cause-and-effect. When one billiard ball strikes another, there is an energy transfer such that the motion of the second ball is energized by the impact of the first. In communicational systems, on the other hand, the energy of the response is usually provided by the respondent. If I kick a dog, his immediately sequential behavior is energized by his metabolism, not by *my* kick. Similarly, when one neuron fires another, or an impulse from a microphone activates a circuit, the sequent event has its own energy sources.

Of course, everything that happens is still within the limits defined by the law of energy conservation. The dog's metabolism might in the end limit his response, but, in general, in the systems with which we deal, the energy supplies are large compared with the demands upon them; and, long before the supplies are exhausted, "economic" limitations are imposed by the finite number of available alternatives, i.e., there is an economics of probability. This economics differs from an economics of energy or money in that probability—being a ratio—is not subject to addition or subtraction but only to multiplicative processes, such as fractionation. A telephone exchange at a time of emergency may be "jammed" when a large fraction of its alternative pathways are busy. There is, then, a low probability of any given message getting through.

In addition to the restraints due to the limited economics of alternatives, two other categories of restraint must be discussed: restraints related to "feedback" and restraints related to "redundancy."

We consider first the concept of feedback:

When the phenomena of the universe are seen as linked together by cause-and-effect and energy transfer, the resulting picture is of complexly branching and interconnecting chains of causation. In certain regions of this universe (notably organisms in environments, ecosystems, thermostats, steam engines with governors, societies, computers, and the like), these chains of causation form circuits which are *closed* in the sense that causal interconnection can be traced around the circuit and back through whatever position was (arbitarily) chosen as the starting point of the description. In such a circuit, evidently, events at any position in the circuit may be expected to have effect at *all* positions on the circuit at later times.

Such systems are, however, always *open: (a)* in the sense that the circuit is energized from some external source and loses energy usually in the form of heat to the outside; and (b) in the sense that events within the circuit may be influenced from the outside or may influence outside events.

A very large and important part of cybernetic theory is concerned with the formal characteristics of such causal circuits, and the conditions of their stability. Here I shall consider such systems only as sources of *restraint.*

Consider a variable in the circuit at any position and sup-pose this variable subject to random change in value (the change perhaps being imposed by impact of some event external to the circuit). We now ask how this change will affect the value of this variable at that later time when the sequence of effects has come around the circuit. Clearly the answer to this last question will depend upon the characteristics of the circuit and will, therefore, be *not random.*

In principle, then, a causal circuit will generate a non-random response to a random event *at that position in the circuit at which the random event occurred.*

This is the general requisite for the creation of cybernetic restraint in any variable at any given position. The particular restraint created in any given instance will, of course, depend upon the characteristics of the particular circuit—whether its overall gain be positive or negative, its time characteristics, its thresholds of activity, etc. These will together determine the restraints which it will exert at any given position.

For purposes of cybernetic explanation, when a machine is observed to be (improbably) moving at a constant rate, even under varying load, we shall look for restraints—e.g., for a circuit which will be activated by changes in rate and which, when activated, will operate upon some variable (e.g., the fuel supply) in such a way as to diminish the change in rate.

When the monkey is observed to be (improbably) typing prose, we shall look for some circuit which is activated whenever he makes a "mistake" and which, when activated, will delete the evidence of that mistake at the position where it occurred.

The cybernetic method of negative explanation raises the question: Is there a difference between "being right" and "not being

wrong"? Should we say of the rat in a maze that he has "learned the right path" or should we say only that he has learned "to avoid the wrong paths"?

Subjectively, I feel that I know how to spell a number of English words, and I am certainly not aware of discarding as unrewarding the letter K when I have to spell the word "many." Yet, in the first level cybernetic explanation, I should be viewed as actively discarding the alternative K when I spell "many."

The question is not trivial and the answer is both subtle and fundamental: *choices are not all at the same level. I* may have to avoid error in my choice of the word "many" in a given context, discarding the alternatives, "few," "several," "frequent," etc. But if I can achieve this higher level choice on a negative base, it follows that the word "many" and its alternatives somehow must be conceivable to me—must exist as distinguishable and possibly labeled or coded patterns in my neural processes. If they do, in some sense, exist, then it follows that, after making the higher level choice of what word to use, I shall not necessarily be faced with alternatives at the lower level. It may become unnecessary for me to exclude the letter K from the word "many." It will be correct to say that I know positively how to spell "many"; not merely that I know how to avoid making mistakes in spelling that word.

It follows that Lewis Carroll's joke about the theory of natural selection is not entirely cogent. If, in the communicational and organizational processes of biological evolution, there be something like levels—items, patterns, and possibly patterns of patterns—then it is logically possible for the evolutionary system to make something like positive choices. Such levels and patterning might conceivably be in or among genes or elsewhere.

The circuitry of the above mentioned monkey would be required to recognize deviations from "prose," and prose is characterized by pattern or—as the engineers call it—by redundancy.

The occurrence of the letter K in a given location in an English prose message is not a purely random event in the sense that there was ever an equal probability that any other of the twenty-five letters might have occurred in that location. Some letters are more common in English than others, and certain combinations of letters are more common than others. There is, thus, a species of patterning

which partly determines which letters shall occur in which slots. As a result: if the receiver of the message had received the entire rest of the message but had not received the particular letter K which we are discussing, he might have been able, with better than random success, to guess that the missing letter was, in fact, K. To the extent that this was so, the let-ter K did not, for that receiver, exclude the other twenty-five letters because these were already partly excluded by information which the recipient received from the rest of the message. This patterning or predictability of particular events within a larger aggregate of events is technically called "redundancy."

The concept of redundancy is usually derived, as I have derived it, by considering first the maximum of information which might be carried by the given item and then considering how this total might be reduced by knowledge of the surrounding patterns of which the given item is a component part. There is, however, a case for looking at the whole matter the other way round. We might regard patterning or predictability as the very essence and *raison d'etre* of communication, and see the single letter unaccompanied by collateral clues as a peculiar and special case.

The idea that communication is the creation of redundancy or patterning can be applied to the simplest engineering examples. Let us consider an *observer* who is watching A send a message to B. The purpose of the transaction (from the point of view of A and B) is to create in B's message pad a sequence of letters identical with the sequence which formerly occurred in A's pad. But from the point of view of the observer this is the creation of redundancy. If he has seen what A had on his pad, he will not get any new information about the message itself from inspecting B's pad.

Evidently, the nature of "meaning," pattern, redundancy, information and the like. depends upon where we sit. In the usual engineers' discussion of a message sent from A to B, it is customary to omit the observer and to say that B received information from A which was measurable in terms of the number of letters transmitted, reduced by such redundancy in the text as might have permitted B to do some guessing. But in a wider universe, i.e., that defined by the point of view of the observer, this no longer appears as a "transmission" of information but rather as a spreading of redundancy. The activities of A and B have combined to make the

universe of the observer more predictable, more ordered, and more redundant. We may say that the rules of the "game" played by A and B explain (as "restraints") what would otherwise be a puzzling and improbable coincidence in the observer's universe, namely the conformity between what is written on the two message pads.

To guess, in essence, is to face a cut or slash in the sequence of items and to predict across that slash what items might be on the other side. The slash may be spatial or temporal (or both) and the guessing may be either predictive or retrospective. A pattern, in fact, is definable as an aggregate of events or objects which will permit in some degree such guesses when the entire aggregate is not available for inspection.

But this sort of patterning is also a very, general phenomenon, outside the realm of communication *between* organisms. The reception of message material by *one* organism is not fundamentally different from any other case of perception. If I see the top part of a tree standing up, I can predict —with better than random success— that the tree has roots in the ground. The percept of the tree top is redundant with (i.e., contains "information" about) parts of the system which I cannot perceive owing to the slash provided by the opacity of the ground.

If then we say that a message has "meaning" or is "about" some referent, what we mean is that there is a larger universe of relevance consisting of message-plus-referent, and that redundancy or pattern or predictability is introduced into this universe by the message.

If I say to you "It is raining," this message introduces redundancy into the universe, message-plus-raindrops, so that from the message alone you could have guessed—with better than random success—something of what you would see if you looked out of the window. The universe, message-plus-referent, is given pattern or form—in the Shakespearean sense, the universe is *informed* by the message; and the "form" of which we are speaking is not in the message nor is it in the referent. It is a correspondence between message and referent.

In loose talk, it seems simple to locate information. The letter K in a given slot proposes that the letter in that particular slot is a K. And, so long as all information is of this very direct kind, the

information can be "located": the information about the letter K is seemingly in that slot.

The matter is not quite so simple if the text of the message is redundant but, if we are lucky and the redundancy is of low order, we may still be able to point to parts of the text which indicate (carry some of the information) that the letter K is expectable in that particular slot.

But if we are asked: Where are such items of information as that: (a) "This message is in English"; and (b) "In English, a letter K often follows a letter C, except when the C begins a word"; we can only say that such information is *not* localized in any part of the text but is rather a statistical induction from the text as a whole (or perhaps from an aggregate of "similar" texts). This, after all, is metainformation and is of a basically different order—of different logical type—from the information that "the letter in this slot is K."

This matter of the localization of information has be-deviled communication theory and especially neurophysiology for many years and it is, therefore, interesting to consider how the matter looks if we start from redundancy, pattern or form as the basic concept.

It is flatly obvious that no variable of zero dimensions can be truly located. "Information" and "form" resemble contrast, frequency, symmetry, correspondence, congruence, conformity, and the like in being of zero dimensions and, therefore, are not to be located. The contrast between this white paper and that black coffee is not somewhere between the paper and the coffee and, even if we bring the paper and coffee into close juxtaposition, the contrast between them is not thereby located or pinched between them. Nor is that contrast located between the two objects and my eye. It is not even in my head; or, if it be, then it must also be in your head. But you, the reader, have not seen the paper and the coffee to which I was referring. I have in my head an image or transform or name of the contrast between them; and you have in your head a transform of what I have in mine. But the conformity between us is not localizable. In fact, information and form are not items which can be localized.

It is, however, possible to begin (but perhaps not complete) a sort of mapping of formal relations within a system containing

redundancy. Consider a finite aggregate of objects or events (say a sequence of letters, or a tree) and an observer who is already informed about all the redundancy rules which are recognizable (i.e., which have statistical significance) within the aggregate. It is then possible to delimit regions of the aggregate within which the observer can achieve better than random guessing. A further step toward localization is accomplished by cutting across these regions with slash marks, such that it is across these that the educated observer can guess, from what is on one side of the slash, something of what is on the other side.

Such a mapping of the distribution of patterns is, how-ever, in principle, incomplete because we have not considered the sources of the observer's prior knowledge of the redundancy rules. If, now, we consider an observer with *no* prior knowledge, it is clear that he might discover some of the relevant rules from his perception of *less* than the whole aggregate. He could then use his discovery in predicting *rules* for the remainder—rules which would be correct even though not exemplified. He might discover that "H often follows T" even though the remainder of the aggregate contained no example of this combination. For this order of phenomenon a different order of slash mark—metaslashes —will be necessary.

It is interesting to note that metaslashes which demarcate what is necessary for the naive observer to discover a rule are, in principle, displaced relative to the slashes which would have appeared on the map prepared by an observer totally informed as to the rules of redundancy for that aggregate. (This principle is of some importance in aesthetics.

To the aesthetic eye, the form of a crab with one claw bigger than the other is not simply asymmetrical. It first pro-poses a rule of symmetry and then subtly denies the rule by proposing a more complex combination of rules.)

When we exclude all things and all real dimensions from our explanatory system, we are left regarding each step in a communicational sequence as *a transform* of the previous step. If we consider the passage of an impulse along an axon, we shall regard the events at each point along the pathway as a transform (albeit identical or similar) of events at any previous point. Or if we consider a series of neurons, each firing the next, then the firing of

each neuron is a transform of the firing of its predecessor. We deal with event sequences which do not necessarily imply a passing on of the same energy.

Similarly, we can consider any network of neurons, and arbitrarily transect the whole network at a series of different positions, then we shall regard the events at each transection as a transform of events at some previous transection.

In considering perception, we shall not say, for example, "I see a tree," because the tree is not within our explanatory system. At best, it is only possible to see an image which is a complex but systematic transform of the tree. This image, of course, is energized *by my* metabolism and the nature of the transform is, in part, determined by factors within my neural circuits: "I" make the image, under various restraints, some of which are imposed by *my* neural circuits, while others are imposed by the external tree. An hallucination or dream would be more truly "mine" insofar as it is produced without immediate external restraints.

All that is not information, not redundancy, not form and not restraints—is noise, the only possible source of *new* patterns.

Form, Substance, and Difference[*]

Let me say that it is an extraordinary honor to be here tonight, and a pleasure. I am a little frightened of you all, because I am sure there are people here who know every field of knowledge that I have touched much better than I know it. It is true that I have touched a number of fields, and I probably can face any one of you and say I have touched a field that you have not touched. But I am sure that for every field I have touched, there are people here who are much more expert than I. I am not a well-read philosopher, and philosophy is not my business. I am not a very well-read anthropologist, and anthropology is not exactly my business.

But I have tried to do something which Korzybski was very much concerned with doing, and with which the whole semantic movement has been concerned, namely, I have studied the area of impact between very abstract and formal philosophic thought on the one hand and the natural history of man and other creatures on the other. This overlap between formal premises and actual behavior is, I assert, of quite dreadful importance today. We face a world which is threatened not only with disorganization of many kinds, but also with the destruction of its environment, and we, today, are still unable to think clearly about the relations between an organism and its environment. What sort of a thing is this, which we call "organism plus environment"?

Let us go back to the original statement for which Korzybski is most famous—the statement that *the map is not the territory.* This statement came out of a very wide range of philosophic thinking, going back to Greece, and wriggling through the history of European thought over the last 2000 years. In this history, there has been a sort of rough dichotomy and often deep controversy. There has been a violent enmity and bloodshed. It all starts, I suppose,

[*] This was the Nineteenth Annual Korzybski Memorial Lecture, delivered January 9, 1970, under the auspices of the Institute of General Semantics. It is here re-printed from the *General Semantics* Bulletin, No. 37, 1970, by permission of the Institute of General Semantics.

with the Pythagoreans versus their predecessors, and the argument took the shape of "Do you ask what it's made of—earth, fire, water, etc?" Or do you ask, "What is its pattern?" Pythagoras stood for inquiry into pattern rather than inquiry into substance.1 That controversy has gone through the ages, and the Pythagorean half of it has, until recently, been on the whole the submerged half. The Gnostics follow the Pythagoreans, and the alchemists follow the Gnostics, and so on. The argument reached a sort of climax at the end of the eighteenth century when a Pythagorean evolutionary theory was built and then discarded—a theory which involved Mind.[5]

The evolutionary theory of the late eighteenth century, the Lamarckian theory, which was the first organized transformist theory of evolution, was built out of a curious historical background which has been described by Lovejoy in *The Great Chain of Being.* Before Lamarck, the organic world, the living world, was believed to be hierarchic in structure, with Mind at the top. The chain, or ladder, went down through the angels, through men, through the apes, down to the infusoria or protozoa, and below that to the plants and stones.

What Lamarck did was to turn that chain upside down. He observed that animals changed under environmental pressure. He was incorrect, of course, in believing that those changes were inherited, but in any case, these changes were for him the evidence of evolution. When he turned the ladder upside down, what had been the explanation, namely, the Mind at the top, now became that which had to be explained. His problem was to explain Mind. He was convinced about evolution, and there his interest in it stopped. So that if you read the *Philosophic Zoologique* (1809), *you* will find that the first third of it is devoted to solving the problem of evolution and the turning upside down of the taxonomy, and the rest of the book is really devoted to comparative psychology, a science which he founded. *Mind* was what he was really interested in. He had used habit as one of the axiomatic phenomena in his

[5] R. G. Collingwood has given a clear account of the Pythagorean position in *The Idea of Nature,* Oxford, 1945.

theory of evolution, and this of course also took him into the problem of comparative psychology.

Now mind and pattern as the explanatory principles which, above all, required investigation were pushed out of biological thinking in the later evolutionary theories which were developed in the mid-nineteenth century by Darwin, Huxley, etc. There were still some naughty boys, like Samuel Butler, who said that mind could not be ignored in this way—but they were weak voices, and incidentally, they never looked at organisms. I don't think Butler ever looked at anything except his own cat, but he still knew more about evolution than some of the more conventional thinkers.

Now, at last, with the discovery of cybernetics, systems theory, information theory, and so on, we begin to have a formal base enabling us to think about mind and enabling us to think about all these problems in a way which was totally heterodox from about *1850* through to World War II. What I have to talk about is how the great dichotomy of epistemology has shifted under the impact of cybernetics and information theory.

We can now say—or at any rate, can begin to say—what we think a mind is. In the next twenty years there will be other ways of saying it and, because the discoveries are new, I can only give you my personal version. The old versions are surely wrong, but which of the revised pictures will survive, we do not know.

Let us start from the evolutionary side. It is now empirically clear that Darwinian evolutionary theory contained a very great error in its identification of the unit of survival under natural selection. The unit which was believed to be crucial and around which the theory was set up was either the breeding individual or the family line or the sub-species or some similar homogeneous set of conspecifics. Now I suggest that the last hundred years have demonstrated empirically that if an organism or aggregate of organisms sets to work with a focus on its own survival and thinks that that is the way to select its adaptive moves, its "progress" ends up with a destroyed environment. If the organism ends up destroying its environment, it has in fact destroyed itself. And we may very easily see this process carried to its ultimate *reductio ad absurdum* in the next twenty years. The unit of survival is not the breeding organism, or the family line, or the society.

The old unit has already been partly corrected by the population geneticists. They have insisted that the evolutionary unit is, in fact, not homogeneous. A wild population of any species consists always of individuals whose genetic constitution varies widely. In other words, potentiality and readiness for change is already built into the survival unit. The heterogeneity of the wild population is already one-half of that trial-and-error system which is necessary for dealing with environment.

The artificially homogenized populations of man's domestic animals and plants are scarcely fit for survival.

And today a further correction of the unit is necessary. The flexible environment must also be included along with the flexible organism because, as I have already said, the organism which destroys its environment destroys itself. The unit of survival is a flexible organism-in-its-environment.

Now, let me leave evolution for a moment to consider what is the unit of mind. Let us go back to the map and the territory and ask: "What is it in the territory that gets onto the map?" We know the territory does not get onto the map. That is the central point about which we here are all agreed. Now, if the territory were uniform, nothing would get onto the map except its boundaries, which are the points at which it ceases to be uniform against some larger matrix. What gets onto the map, in fact, is *difference,* be it a difference in altitude, a difference in vegetation, a difference in population structure, difference in surface, or what-ever. Differences are the things that get onto a map.

But what is a difference? A difference is a very peculiar and obscure concept. It is certainly not a thing or an event. This piece of paper is different from the wood of this lectern. There are many differences between them—of color, texture, shape, etc. But if we start to ask about the localization of those differences, we get into trouble. Obviously the difference between the paper and the wood is not in the paper; it is obviously not in the wood; it is obviously not in the space between them, and it is obviously not in the time between them. (Difference which occurs across time is what we call "change.")

A difference, then, is an abstract matter.

In the hard sciences, effects are, in general, caused by rather concrete conditions or events—impacts, forces, and so forth. But when you enter the world of communication, organization, etc., you leave behind that whole world in which effects are brought about by forces and impacts and energy exchange. You enter a world in which "effects"—and I am not sure one should still use the same word—are brought about by *differences.* That is, they are brought about by the sort of "thing" that gets onto the map from the territory. This is difference.

Difference travels from the wood and paper into my retina. It then gets picked up and worked on by this fancy piece of computing machinery in my head.

The whole energy relation is different. In the world of mind, nothing—that which is *not—can* be a cause. In the hard sciences, we ask for causes and we expect them to exist and be "real." But remember that zero is different from one, and because zero is different from one, zero can be a cause in the psychological world, the world of communication. The letter which you do not write can get an angry reply; and the income tax form which you do not fill in can trigger the Internal Revenue boys into energetic action, because they, too, have their breakfast, lunch, tea, and dinner and can react with energy which they derive from their metabolism. The letter which never existed is no source of energy.

It follows, of course, that we must change our whole way of thinking about mental and communicational process. The ordinary analogies of energy theory which people borrow from the hard sciences to provide a conceptual frame upon which they try to build theories about psychology and behavior—that entire Procrustean structure—is non-sense. It is in error.

. I suggest to you, now, that the word "idea," in its most elementary sense, is synonymous with "difference." Kant, in the *Critique* of *Judgment—if* I understand him correctly—asserts that the most elementary aesthetic act is the selection of a fact. He argues that in a piece of chalk there are an infinite number of potential facts. The *Ding an sich,* the piece of chalk, can never enter into communication or mental process because of this infinitude. The sensory receptors cannot accept it; they filter it out.

What they do is to select certain *facts* out of the piece of chalk, which then become, in mod-ern terminology, information.

I suggest that Kant's statement can be modified to say that there is an infinite number of *differences* around and within the piece of chalk. There are differences between the chalk and the rest of the universe, between the chalk and the sun or the moon. And within the piece of chalk, there is for every molecule an infinite number of differences between its location and the locations in which it *might* have been. Of this infinitude, we select a very limited number, which be-come information. In fact, what we mean by information—the elementary unit of information—is *a difference which makes a difference,* and it is able to make a difference because the neural pathways along which it travels and is continually transformed are themselves provided with energy. The path-ways are ready to be triggered. We may even say that the question is already implicit in them.

There is, however, an important contrast between most of the pathways of information inside the body and most of the pathways outside it. The differences between the paper and the wood are first transformed into differences in the propagation of light or sound, and travel in this form to my sensory end organs. The first part of their journey is energized in the ordinary hard-science way, from "behind." But when the differences enter my body by triggering an end. organ, this type of travel is replaced by travel which is energized at every step by the metabolic energy latent in the protoplasm which *receives* the difference, recreates or transforms it, and passes it on.

When I strike the head of a nail with a hammer, an impulse is transmitted to its point. But it is a semantic error, a misleading metaphor, to say that what travels in an axon is an "impulse." It could correctly be called "news of a difference."

Be that as it may, this contrast between internal and external pathways is not absolute. Exceptions occur on both sides of the line. Some external chains of events are energized by relays, and some chains of events internal to the body are energized from "behind."

Notably, the mechanical interaction of muscles can be used as a computational model.[6]

In spite of these exceptions, it is still broadly true that the coding and transmission of differences outside the body is very different from the coding and transmission inside, and this difference must be mentioned because it can lead us into error. We commonly think of the external "physical world" as somehow separate from an internal "mental world." I believe that this division is based on the contrast in coding and transmission inside and outside the body.

The mental world—the mind—the world of information processing—is not limited by the skin.

Let us now go back to the notion that the transform of a difference traveling in a circuit is an elementary idea. If this be correct, let us ask what a mind is. We say the map is different from the territory. But what is the territory? Operationally, somebody went out with a retina or a measuring stick and made representations which were then put upon paper. What is on the paper map is a representation of what was in the retinal representation of the man who made the map; and as you push the question back, what you find is an infinite regress, an infinite series of maps. The territory never gets in at all. The territory is *Ding an sich* and you can't do anything with it. Always the process of representation will filter it out so that the mental world is only maps of maps of maps, ad infinitum.[7] All "phenomena" are literally appearances.

Or we can follow the chain forward. I receive various sorts of mappings which I call data or information. Upon receipt of these I act. But my actions, my muscular con-tractions, are transforms of differences in the input material. And I receive again data which are transforms of my actions. We get thus a picture of the mental world

[6] It is interesting to note that digital computers depend upon transmission of energy "from behind" to send "news" along wire from one relay to the next. But each relay has its own energy source. Analogic computers, e.g., tide machines and the like, are commonly entirely driven by energy "from behind." Either type of energization can be used for computational purposes.

[7] Or we may spell the matter out and say that at every step, as a difference is transformed and propagated along its pathway, the embodiment of the difference before the step is a "territory" of which the embodiment after the step is a "map." The map-territory relation obtains at every step.

which has some-how jumped loose from our conventional picture of the physical world.

This is not new, and for historic background we go again to the alchemists and Gnostics. Carl Jung once wrote a very curious little book, which I recommend to all of you. It is called *Septem Sermones ad Mortuos,* Seven Sermons to the Dead.[8] In his *Memoirs, Dreams and Reflections,* Jung tells us that his house was full of ghosts, and they were noisy. They bothered him, they bothered his wife, and they bothered the children. In the vulgar jargon of psychiatry, we might say that everybody in the house was as psychotic as hooty owls, and for quite good reason. If you get your epistemology confused, you go psychotic, and Jung was going through an epistemological crisis. So he sat down at his desk and picked up a pen and started to write. When he started to write the ghosts all disappeared, and he wrote this little book. From this he dates all his later insight. He signed it "Basilides," who was a famous Gnostic in Alexandria in the second century.

He points out that there are two worlds. We might call them two worlds of explanation. He names them the *pleroma* and the *creatura,* these being Gnostic terms. The pleroma is the world in which events are caused by forces and impacts and in which there are no "distinctions." Or, as I would say, no "differences." In the creatura, effects are brought about precisely by difference. In fact, this is the same old dichotomy between mind and substance.

We can study and describe the pleroma, but always the distinctions which we draw are attributed *by* us to the pleroma. The pleroma knows nothing of difference and distinction; it contains no "ideas" in the sense in which I am using the word. When we study and describe the creatura, we must correctly identify those differences which are effective within it.

I suggest that "pleroma" and "creatura" are words which we could usefully adopt, and it is therefore worthwhile to look at the

[8] Written in 1916, translated by H. G. Baynes and privately circulated in 1925. Republished by Stuart & Watkins, London, and by Random House, 1961. In later work, Jung seems to have lost the clarity of the Seven Sermons. In his "Answer to Job," the archetypes are said to be "pleromatic." It is surely true, however, that constellations of ideas may seem subjectively to resemble "forces" when their ideational character is unrecognized.

bridges which exist between these two "worlds." It is an oversimplification to say that the "hard sciences" deal only with the pleroma and that the sciences of the mind deal only with the creatura. There is more to it than that.

First, consider the relation between energy and negative entropy. The classical Carnot heat engine consists of a cylinder of gas with a piston. This cylinder is alternately placed in contact with a container of hot gas and with a container of cold gas. The gas in the cylinder alternately expands and contracts as it is heated or cooled by the hot and cold sources. The piston is thus driven up and down.

But with each cycle of the engine, the *difference* between the temperature of the hot source and that of the cold source is reduced. When this difference becomes zero, the engine will stop.

The physicist, describing the pleroma, will write equations to translate the temperature difference into "available energy," which he will call "negative entropy," and will go on from there.

The analyst of the creatura will note that the whole system is a sense organ which is triggered by temperature difference. He will call this difference which makes a difference "information" or "negative entropy." For him, this is only a special case in which the effective difference happens to be a matter of energetics. He is equally interested in all differences which can activate some sense organ. For him, any such difference is "negative entropy."

Or consider the phenomenon which the neurophysiologists call "synaptic summation." What is observed is that in certain cases, when two neurons, A and B, have synaptic connection to a third neuron, C, the firing of neither neuron by it-self is sufficient to fire C; but that when both A and B fire simultaneously (or nearly so), their combined "impulses" will cause C to fire.

In pleromatic language, this combining of events to surmount a threshold is called "summation."

But from the point of view of the student of creatura (and the neurophysiologist must surely have one foot in the pleroma and the other in creatura), this is not summation at all. What happens is that the system operates to create differences. There are two differentiated *classes* of firings by A: those firings which are accompanied by B and those which are unaccompanied. Similarly there are two classes of firings by B.

The so-called "summation," when both fire, is not an additive process from this point of view. It is the formation of a logical product—a process of fractionation rather than summation.

The creatura is thus the world seen as mind, wherever such a view is appropriate. And wherever this view is appropriate, there arises a species of complexity which is absent from pleromatic description: creatural description is al-ways hierarchic.

I have said that what gets from territory to map is trans-forms of difference and that these (somehow selected) differences are elementary ideas.

But there are differences between differences. Every effective difference denotes a demarcation, a line of classification, and all classification is hierarchic. In other words, differences are themselves to be differentiated and classified. In this context I will only touch lightly on the matter of classes of difference, because to carry the matter further would land us in problems of *Principia Mathematica.*

Let me invite you to a psychological experience, if only to demonstrate the frailty of the human computer. First note that differences in texture are *different* (a) from differences in color. Now note that differences in size are *different* (b) from differences in shape. Similarly ratios are different (c) from subtractive differences.

Now let me invite you, as disciples of Korzybski, to define the differences between "different *(a) ,"* "different (b)," and "different (c) " in the above paragraph. The computer in the human head boggles at the task. But not all classes of difference are as awkward to handle.

One such class you are all familiar with. Namely, the class of differences which are created by the process of trans-formation whereby the differences immanent in the territory become differences immanent in the map. In the corner of every serious map you will find these rules of transformation spelled out—usually in words. Within the human mind, it is absolutely essential to recognize the differences of this class, and, indeed, it is these that form the central subject matter of "Science and Sanity."

An hallucination or a dream image is surely a transformation of something. But of what? And by what rules of trans-formation?

Lastly there is that hierarchy of differences which biologists call "levels." I mean such differences as that between a cell and a tissue, between tissue and organ, organ and organism, and organism and society.

These are the hierarchies of units or *Gestalten,* in which each subunit is a part of the unit of next larger scope. And, always in biology, this difference or relationship which I call "part of" is such that certain differences in the part have informational effect upon the larger unit, and vice versa.

Having stated this relationship between biological part and whole, I can now go on from the notion of creatura as Mind in general to the question of what is a mind.

What do I mean by "my" mind?

I suggest that the delimitation of an individual mind must always depend upon what phenomena we wish to under-stand or explain. Obviously there are lots of message path-ways outside the skin, and these and the messages which they carry must be included as part of the mental system whenever they are relevant.

Consider a tree and a man and an axe. We observe that the axe flies through the air and makes certain sorts of gashes in a pre-existing cut in the side of the tree. If now we want to explain this set of phenomena, we shall be concerned with differences in the cut face of the tree, differences in the retina of the man, differences in his central nervous system, differences in his efferent neural messages, differences in the behavior of his muscles, differences in how the axe flies, to the differences which the axe then makes on the face of the tree. Our explanation (for certain purposes) will go round and round that circuit. In principle, if you want to explain or understand anything in human behavior, you are always dealing with total circuits, completed circuits. This is the elementary cybernetic thought.

The elementary cybernetic system with its messages in circuit is, in fact, the simplest unit of mind; and the trans-form of a difference traveling in a circuit is the elementary idea. More complicated systems are perhaps more worthy to be called mental systems, but essentially this is what we are talking about. The unit which shows the characteristic of trial and error will be legitimately called a mental system.

But what about "me"? Suppose I am a blind man, and I use a stick. I go tap, tap, tap. Where do I start? Is my mental system bounded at the handle of the stick? Is it bounded by my skin? Does it start halfway up the stick? Does it start at the tip of the stick? But these are nonsense questions. The stick is a pathway along which transforms of difference are being transmitted. The way to delineate the system is to draw the limiting line in such a way that you do not cut any of these pathways in ways which leave things inexplicable. If what you are trying to explain is a given piece of behavior, such as the locomotion of the blind man, then, for this purpose, you will need the street, the stick, the man; the street, the stick, and so on, round and round.

But when the blind man sits down to eat his lunch, his stick and its messages will no longer be relevant—if it is his eating that you want to understand.

And in addition to what I have said to define the individual mind, I think it necessary to include the relevant parts of memory and data "banks." After all, the simplest cybernetic circuit can be said to have memory of a dynamic kind—not based upon static storage but upon the travel of information around the circuit. The behavior of the governor of a steam engine at Time 2 is partly determined by what it did at Time 1—where the interval between Time 1 and Time 2 is that time necessary for the information to complete the circuit.

We get a picture, then, of mind as synonymous with cybernetic system—the relevant total information-processing, trial-and-error completing unit. And we know that within Mind in the widest sense there will be a hierarchy of sub-systems, any one of which we can call an individual mind.

But this picture is precisely the same as the picture which I arrived at in discussing *the unit of evolution.* I believe that this identity is the most important generalization which I have to offer you tonight.

In considering units of evolution, I argued that you have at each step to include the completed pathways outside the protoplasmic aggregate, be it DNA-in-the-cell, or cell-in-the-body, or body-in-the-environment. The hierarchic structure is not new. Formerly we talked about the breeding individual or the family line or the taxon, and so on. Now each step of the hierarchy is to be thought of as *a*

system, instead of a chunk cut off and visualized as *against* the surrounding matrix.

This identity between the unit of mind and the unit of evolutionary survival is of very great importance, not only theoretical, but also ethical.

It means, you see, that I now localize something which I am calling "Mind" immanent in the large biological system—the ecosystem. Or, if I draw the system boundaries at a different level, then mind is immanent in the total evolutionary structure. If this identity between mental and evolutionary units is broadly right, then we face a number of shifts in our thinking.

First, let us consider ecology. Ecology has currently two faces to it: the face which is called bioenergetics—the economics of energy and materials within a coral reef, a red-wood forest, or a city—and, second, an economics of information, of entropy, negentropy, etc. These two do not fit together very well precisely because the units are differently bounded in the two sorts of ecology. In bioenergetics it is natural and appropriate to think of units bounded at the cell membrane, or at the skin; or of units composed of sets of conspecific individuals. These boundaries are then the frontiers at which measurements can be made to determine the additive-subtractive budget of energy for the given unit. In contrast, informational or entropic ecology deals with the budgeting of pathways and of probability. The resulting bud-gets are fractionating (not subtractive). The boundaries must enclose, not cut, the relevant pathways.

Moreover, the very meaning of "survival" becomes different when we stop talking about the survival of something bounded by the skin and start to think of the survival of the system of ideas in circuit. The contents of the skin are randomized at death and the pathways within the skin are randomized. But the ideas, under further transformation, may go on out in the world in books or works of art. Socrates as a bioenergetic individual is dead. But much of him still lives as a component in the contemporary ecology of ideas.[9]

[9] For the phrase "ecology of ideas," I am indebted to Sir Geoffrey Vickers' essay "The Ecology of Ideas" in *Value Systems and Social Process,* Basic Books, 1968.

It is also clear that theology becomes changed and perhaps renewed. The Mediterranean religions for 5000 years have swung to and fro between immanence and transcendence. In Babylon the gods were transcendent on the tops of hills; in Egypt, there was god immanent in Pharoah; and Christianity is a complex combination of these two beliefs.

The cybernetic epistemology which I have offered you would suggest a new approach. The individual mind is immanent but not only in the body. It is immanent also in pathways and messages outside the body; and there is a larger Mind of which the individual mind is only a sub-system. This larger Mind is comparable to God and is perhaps what some people mean by "God," but it is still immanent in the total interconnected social system and planetary ecology.

Freudian psychology expanded the concept of mind in-wards to include the whole communication system within the body—the autonomic, the habitual, and the vast range of unconscious process. What I am saying expands mind out-wards. And both of these changes reduce the scope of the conscious self. A certain humility becomes appropriate, tempered by the dignity or joy of being part of something much bigger. A part—if you will—of God.

If you put God outside and set him vis-à-vis his creation and if you have the idea that you are created *in* his image, you will logically and naturally see yourself as outside and against the things around you. And as you arrogate all mind to yourself, you will see the world around you as mindless and therefore not entitled to moral or ethical consideration. The environment will seem to be yours to exploit. Your survival unit will be you and your folks or conspecifics against the environment of other social units, other races and the brutes and vegetables.

If this is your estimate of your relation to nature *and you have an advanced technology,* your likelihood of survival will be that of a snowball in hell. You will die either of the toxic by-products of your own hate, or, simply, of over-population and overgrazing. The raw materials of the world are finite.

For a more formal discussion of the survival of ideas, see Gordon Pasks' remarks in Wenner-Gren Conference on "Effects of Conscious Purpose on Human Adaptation," 1968

If I am right, the whole of our thinking about what we are and what other people are has got to be restructured. This is not funny, and I do not know how long we have to do it in. If we continue to operate on the premises that were fashionable in the precybernetic era, and which were especially underlined and strengthened during the Indus-trial Revolution, which seemed to validate the Darwinian unit of survival, we may have twenty or thirty years before the logical *reductio ad absurdum* of our old positions destroys us. Nobody knows how long we have, under the present system, before some disaster strikes us, more serious than the destruction of any group of nations. The most important task today is, perhaps, to learn to think in the new way. Let me say that I don't know how to think that way. Intellectually, I can stand here and I can give you a reasoned exposition of this matter; but if I am cutting down a tree, I still think "Gregory Bateson" is cutting down the tree. I am cutting down the tree. "Myself" is to me still an excessively concrete object, different from the rest of what I have been calling "mind."

The step to realizing—to making habitual—the other way of thinking—so that one naturally thinks that way when one reaches out for a glass of water or cuts down a tree—that step is not an easy one.

And, quite seriously, I suggest to you that we should trust no policy decisions which emanate from persons who do not yet have that habit.

There are experiences and disciplines which may help me to imagine what it would be like to have this habit of correct thought. Under LSD, I have experienced, as have many others, the disappearance of the division between self and the music to which I was listening. The perceiver and the thing perceived become strangely united into a single entity. This state is surely more correct than the state in which it seems that "I hear the music." The sound, after all, is *Ding an Bich,* but my perception of it is a part of mind.

It is told of Johann Sebastian Bach that when somebody asked him how he played so divinely, he answered, "I play the notes, in order, as they are written. It is God who makes the music." But not many of us can claim Bach's correctness of epistemology—or that of William Blake, who knew that the Poetic Imagination was the only reality. The poets have known these things all through the ages,

but the rest of us have gone astray into all sorts of false reifications of the "self" and separations between the "self" and "experience."

For me another clue another moment when the nature of mind was for a moment clear—was provided by the famous experiments of Adelbert Ames, Jr. These are optical illusions in depth perception. As Ames' guinea pig, you discover that those mental processes by which you create the world in three-dimensional perspective are within your mind but totally unconscious and utterly beyond voluntary control. Of course, we all know that this is so—that mind creates the images which "we" then see. But still it is a pro-found epistemological shock to have direct experience of this which we always knew.

Please do not misunderstand me. When I say that the poets have always known these things or that most of mental process is unconscious, I am not advocating a greater use of emotion or a lesser use of intellect. Of course, if what I am saying tonight is approximately true, then our ideas about the relation between thought and emotion need to be revised. If the boundaries of the "ego" are wrongly drawn or even totally fictitious, then it may be nonsense to regard emotions or dreams or our unconscious computations of perspective as "ego-alien."

We live in a strange epoch when many psychologists try to "humanize" their science by preaching an anti-intellectual gospel. They might, as sensibly, try to physicalize physics by discarding the tools of mathematics.

It is the attempt to *separate* intellect from emotion that is monstrous, and I suggest that it is equally monstrous—and dangerous—to attempt to separate the external mind from the internal. Or to separate mind from body.

Blake noted that "A tear is an intellectual thing," and Pascal asserted that "The heart has its *reasons of* which the reason knows nothing." We need not be put off by the fact that the reasonings of the heart (or of the hypothalamus) are accompanied by sensations of joy or grief. These computations are concerned with matters which are vital to mammals, namely, matters of *relationship,* by which I mean love, hate, respect, dependency, spectatorship, performance, dominance, and so on. These are central to the life of any mammal and I see no objection to calling these computations "thought,"

though certainly the units of relational computation are different from the units which we use to compute about isolable things.

But there are bridges between the one sort of thought and the other, and it seems to me that the artists and poets are specifically concerned with these bridges. It is not that art is the expression of the unconscious, but rather that it is concerned with the relation *between* the levels of mental process. From a work of art it may be possible to analyze out some unconscious thoughts of the artist, but I believe that, for example, Freud's analysis of Leonardo's *Virgin on the Knees of St. Anne* precisely misses the point of the whole exercise. Artistic skill is the combining of many levels of mind — unconscious, conscious, and external—to make a statement of their combination. It is not a matter of expressing a single level.

Similarly, Isadora Duncan, when she said, "If I could say it, I would not have to dance it," was talking nonsense, be-cause her dance was about combinations of saying and moving.

Indeed, if what I have been saying is at all correct, the whole base of aesthetics will need to be re-examined. It seems that we link feelings not only to the computations of the heart but also to computations in the external pathways

of the mind. It is when we recognize the operations of creatura in the external world that we are aware of "beauty" or "ugliness." The "primrose by the river's brim" is beautiful because we are aware that the combination of differences which constitutes its appearance could only be achieved by information processing, i.e., by *thought*. We recognize an-other mind within our own external mind.

And last, there is death. It is understandable that, in a civilization which separates mind from body, we should either try to forget death or to make mythologies about the survival of transcendent mind. But if mind is immanent not only in those pathways of information which are located in-side the body but also in external pathways, then death takes on a different aspect. The individual nexus of pathways which I call "me" is no longer so precious because that nexus is only part of a larger mind.

The ideas which seemed to be me can also become immanent in you. May they survive if true.