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. According 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 prob-

*This article is reprinted from the *American Behavioral Scientist,* Vol. 10, No. 8, April 1967, pp. 29–32, by permission of the publisher, Sage Publications, Inc.

ability. 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.

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 transformations 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 explanation, "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 suppose 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 letter 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'être* 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 bedeviled 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, however, 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 proposes 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.

Redundancy and Coding*

Discussion of the evolutionary and other relationships between the communication systems of men and those of other animals has made it very clear that the coding devices characteristic of verbal communication differ profoundly from those of kinesics and paralanguage. But the point has been made that there is a great deal of resemblance between the codes of kinesics and paralanguage and the codes of nonhuman mammals.

We may, I think, state categorically that man's verbal system is not derived in any simple way from these preponderantly iconic codes. There is a general popular belief that in the evolution of man, language replaced the cruder systems of the other animals. I believe this to be totally wrong and would argue as follows:

In any complex functional system capable of adaptive evolutionary change, when the performance of a given function is taken over by some new and more efficient method, the old method falls into disuse and decay. The technique of making weapons by the knapping of flint deteriorated when metals came into use.

This decay of organs and skills under evolutionary replace-

*This essay appeared as Chapter 22 in *Animal Communication: Techniques of Study and Results of Research,* edited by Thomas A. Sebeok. Copyright 1968 by Indiana University Press. Reprinted by permission of the publisher.

ment is a necessary and inevitable systemic phenomenon. If, therefore, verbal language were in any sense an evolutionary replacement of communication by means of kinesics and paralanguage, we would expect the old, preponderantly iconic systems to have undergone conspicuous decay. Clearly they have not. Rather, the kinesics of men have become richer and more complex, and paralanguage has blossomed side by side with the evolution of verbal language. Both kinesics and paralanguage have been elaborated into complex forms of art, music, ballet, poetry, and the like, and, even in everyday life, the intricacies of human kinesic communication, facial expression, and vocal intonation far exceed anything that any other animal is known to produce. The logician's dream that men should communicate only by unambiguous digital signals has not come true and is not likely to.

I suggest that this separate burgeoning evolution of kinesics and paralanguage alongside the evolution of verbal language indicates that our iconic communication serves functions totally different from those of language and, indeed, performs functions which verbal language is unsuited to perform.

When boy says to girl, "I love you," he is using words to convey that which is more convincingly conveyed by his tone of voice and his movements; and the girl, if she has any sense, will pay more attention to those accompanying signs than to the words. There are people—professional actors, confidence tricksters, and others—who are able to use kinesics and paralinguistic communication with a degree of voluntary control comparable to that voluntary control which we all think we have over the use of words. For these people who can lie with kinesics, the special usefulness of nonverbal communication is reduced. It is a little more difficult for them to be sincere and still more difficult for them to be believed to be sincere. They are caught in a process of diminishing returns such that, when distrusted, they try to improve their skill in simulating paralinguistic and kinesic sincerity. But this is the very skill which led others to distrust them.

It seems that the discourse of nonverbal communication is precisely concerned with matters of relationship—love, hate, respect, fear, dependency, etc.—between self and

vis-à-vis or between self and environment and that the nature of human society is such that falsification of this discourse rapidly becomes pathogenic. From an adaptive point of view, it is therefore important that this discourse be carried on by techniques which are relatively unconscious and only imperfectly subject to voluntary control. In the language of neurophysiology, the controls of this discourse must be placed in the brain caudad of the controls of true language.

If this general view of the matter be correct, it must follow that to translate kinesics or paralinguistic messages into words is likely to introduce gross falsification due not merely to the human propensity for trying to falsify statements about "feelings" and relationship and to the distortions which arise whenever the products of one system of coding are dissected onto the premises of another, but especially to the fact that all such translation must give to the more or less unconscious and involuntary iconic message the appearance of conscious intent.

As scientists, we are concerned to build a simulacrum of the phenomenal universe in words. That is, our product is to be a verbal transform of the phenomena. It is necessary, therefore, to examine rather carefully the rules of this transformation and the differences in coding between natural phenomena, message phenomena, and words. I know that it is unusual to presume a "coding" of nonliving phenomena and, to justify this phrase, I must expand somewhat on the concept of "redundancy" as this word is used by the communications engineers.

The engineers and mathematicians have concentrated their attention rigorously upon the internal structure of message material. Typically, this material consists of a sequence or collection of events or objects (commonly members of finite sets—phonemes and the like). This sequence is differentiated from irrelevant events or objects occurring in the same region of time-space by the signal/noise ratio and by other characteristics. The message material is said to contain "redundancy" if, when the sequence is received with some items missing, the receiver is able to guess at the missing items with better than random success. It has been pointed out that, in fact, the term "redundancy" so used becomes a synonym for "patterning."[1] It is important to note that this patterning of message material always helps the receiver to differentiate between signal and noise. In fact, the regularity called signal/noise ratio is really only a special case of redundancy. Camouflage (the opposite of communication) is achieved (1) by reducing the signal/noise ratio, (2)by breaking up the patterns and regularities in the signal, or (3) by introducing similar patterns into the noise.

By confining their attention to the internal structure of the message material, the engineers believe that they can avoid the complexities and difficulties introduced into communication theory by the concept of "meaning." I would argue, however, that the concept "redundancy" is at least a partial synonym of "meaning." As I see it, if the receiver can guess at missing parts of the message, then those parts which are received must, in fact, carry a *meaning* which refers to the missing parts and is information about those parts.

If now we turn away from the narrow universe of message structure and consider the outer world of natural phenomena, we observe at once that this outer world is similarly characterized by redundancy, *i.e.*, that when an observer perceives only certain parts of a sequence or configuration of phenomena, he is in many cases able to guess, with better than random success, at the parts which he cannot immediately perceive. It is, indeed, a principal goal of the scientist to elucidate these redundancies or patternings of the phenomenal world.

If we now consider that larger universe of which these two subuniverses are parts, *i.e.*, the system: message *plus* external phenomena, we find that this larger system contains redundancy of a very special sort. The observer's ability to predict external phenomena is very much increased by his receipt of message material. If I tell you that "it is raining" and you look out the window, you will get less information from the perception of raindrops than you would have got had you never received my message. From my message you could have guessed that you would see rain.

In sum, "redundancy" and "meaning" become synonymous whenever both words are applied to the same universe of

[1] F. Attneave, *Applications of Information Theory to Psychology*, New York, Henry Holt and Co., 1959.

discourse. "Redundancy" within the restricted universe of the message sequence is not, of course, synonymous with "meaning" in the wider universe that includes both message and external referent.

It will be noted that this way of thinking about communication groups all methods of coding under the single rubric of part-for-whole. The verbal message "It is raining" is to be seen as a *part* of a larger universe within which that message creates redundancy or predictability. The "digital," the "analogic," the "iconic," the "metaphoric," and all other methods of coding are subsumed under this single heading. (What the grammarians call "synecdoche" is the metaphoric use of the name of a part in place of the name of the whole, as in the phrase "five *head* of cattle.")

This approach to the matter has certain advantages: the analyst is forced at all times to define the universe of discourse within which "redundancy" or "meaning" is supposed to occur. He is forced to examine the "logical typing" of all message material. We shall see that this broad view of the matter makes it easy to identify major steps in the evolution of communication. Let us consider the scientist who is observing two animals in a physical environment. The following components then must be considered:

(1) The physical environment contains internal patterning or redundancy, *i.e.*, the perception of certain events or objects makes other events or objects predictable for the animals and/or for the observer.

(2) Sounds or other signals from one animal may contribute redundancy to the system, *environment plus signal; i.e.*, the signals may be "about" the environment.

(3) The sequence of signals will certainly contain redundancy—one signal from an animal making another signal from the same animal more predictable.

(4) The signals may contribute redundancy to the universe; *A's signals plus B's signals, i.e.*, the signals may be *about* the interaction of which they are component parts.

(5) If all rules or codes of animal communication and understanding were genotypically fixed, the list would end at this point. But some animals are capable of *learning, e.g.*, the repetition of sequences may lead to their becoming effective as patterns. In logic, "every proposition proposes its own truth," but in natural history we deal always with a

converse of this generalization. The perceivable events which accompany a given percept propose that that percept shall "mean" these events. By some such steps an organism may learn to use the information contained in patterned sequences of external events. I can therefore predict with better than random success that in the universe, organism *plus* environment, events will occur to complete patterns or configurations of learned adaptation between organism and environment.

(6) The behavioral "learning" which is usually studied in psychological laboratories is of a different order. The redundancy of that universe, which consists of the animal's actions *plus* external events, is increased, from the animal's point of view, when the animal regularly responds to certain events with certain actions. Similarly, this universe gains redundancy when the animal succeeds in producing those actions which function as regular *precursors* (or causes) of specific external events.

(7) For every organism there are limitations and regularities which define what will be learned and under what circumstances this learning will occur. These regularities and patterns become basic premises for the individual adaptation and social organization of any species.

(8) Last but not least, there is the matter of phylogenetic learning and phylogeny in general. There is redundancy in the system, organism-*plus*-environment, such that from the morphology and behavior of the organism a human observer can guess with better than random success at the nature of the environment. This "information" about the environment has become lodged in the organism through a long phylogenetic process, and its coding is of a very special kind. The observer who would learn about the aquatic environment from the shape of a shark must deduce the hydrodynamics from the adaptation which copes with the water. The information contained in the phenotypic shark is implicit in forms which are complementary to characteristics of other parts of the universe, *phenotype plus environment* whose redundancy is increased by the phenotype.

This very brief and incomplete survey of some of the sorts of redundancy in biological systems and the universes of their relevance indicates that under the general rubric "part-for-whole" a number of different sorts of relationship between

part and whole are included. A listing of some of the characteristics of these formal relations is in order. We consider some of the iconic cases:

(1) The events or objects which we here call the "part" or "signal" may be real components of an existing sequence or whole. A standing trunk of a tree indicates the probable presence of invisible roots. A cloud may indicate the coming storm of which it is a part. The bared fang of a dog may be part of a real attack.

(2) The "part" may have only a conditional relationship to its whole: the cloud may indicate that we shall get wet if we don't go indoors; the bared fang may be the beginning of an attack which will be completed unless certain conditions are met.

(3) The "part" may be completely split from the whole which is its referent. The bared fang at the given instant may *mention* an attack which, if and when it occurs, will include a *new* baring of the fangs. The "part" has now become a true iconic signal.

(4) Once a true iconic signal has evolved—not necessarily through steps 1, 2, or 3, above—a variety of other pathways of evolution become possible:

(*a*) The "part" may become more or less digitalized, so that magnitudes within it no longer refer to magnitudes within the whole which is its referent but, for example, contribute to an improvement of the signal/noise ratio.

(*b*) The "part" may take on special ritual or metaphoric meanings in contexts where the original whole to which it once referred is no longer relevant. The game of mutual mouthing between mother dog and puppy which once followed her weaning of the pup may become a ritual aggregation. The actions of feeding a baby bird may become a ritual of courtship, etc.

Throughout this series, whose branches and varieties are here only briefly indicated, it is notable that animal communication is confined to signals which are derived from actions of the animals themselves, *i.e.*, those which are parts of such actions. The external universe is, as already noted, redundant in the sense that it is replete with part-for-whole messages, and—perhaps for that reason—this basic style of coding is characteristic of primitive animal communication. But in so far as animals can signal at all about the external universe, they do so by means of actions which are parts of their response to that universe. The jackdaws indicate to each other that Lorenz is a "jackdaw-eater" not by simulating some part of the act of eating jackdaws but by simulating part of their aggression vis-à-vis such a creature. Occasionally actual pieces of the external environment—scraps of potential nest-building material, "trophies," and the like—are used for communication, and in these cases again the messages usually contribute redundancy to the universe *message plus the relationship between the organisms* rather than to the universe *message plus external environment.*

In terms of evolutionary theory, it is not simple to explain why over and over again genotypic controls have been evolved to determine such iconic signaling. From the point of view of the human observer such iconic signals are rather easy to interpret, and we might expect iconic coding to be comparatively easy for animals to decode—in so far as the animals must *learn* to do so. But the genome is presumed not capable of learning in this sense, and we might therefore expect genotypically determined signals to be aniconic or arbitrary rather than iconic.

Three possible explanations of the iconic nature of genotypic signals can be offered:

(1) Even genotypically determined signals do not occur as separate and isolated elements in the life of the phenotype but are necessarily components in a complex matrix of behavior some, at least, of which is learned. It is possible that the iconic coding of genotypically determined signals renders these easy to assimilate into this matrix. There may be an experiential "schoolmarm" which acts selectively to favor those genotypic changes which will give rise to iconic rather than arbitrary signaling.

(2) A signal of aggression which places the signaler in a position of readiness to attack probably has more survival value than would a more arbitrary signal.

(3) When the genotypically determined signal affects the behavior of another species—*e.g.*, eye marks or postures which have a warning effect, movements which facilitate camouflage or aposematic mimicry—clearly the signal must be iconic to the perceptive system of that other species. However, an interesting phenomenon arises in many instances where what is achieved is a secondary statistical iconicism.

Labroides dimidiatus, a small Indo-Pacific wrasse, which lives on the ectoparasites of other fishes, is strikingly colored and moves or "dances" in a way which is easily recognized. No doubt these characteristics attract other fish and are part of a signaling system which leads the other fish to permit the approaches of the cleaner. But there is a mimic of this species of *Labroides*, a saber-toothed blenny (*Aspidontus taeniatus*), whose similar coloring and movement permit the mimic to approach—and bite off pieces of the fins of other fishes.[2]

Clearly the coloring and movements of the mimic are iconic and "represent" the cleaner. But what of the coloring and movements of the latter? All that is primarily required is that the cleaner be conspicuous or distinctive. It is not required that it represent something else. But when we consider the statistical aspects of the system, it becomes clear that if the blennies become too numerous, the distinctive features of the wrasses will become iconic warnings and their hosts will avoid them. What is necessary is that the signals of the wrasse shall clearly and indubitably represent wrasse, *i.e.*, the signals, though perhaps aniconic in the first instance, must achieve and maintain by multiple impact a sort of autoiconicism. "When I say it three times, it is true." But this necessity for autoiconicism may also arise within the species. Genotypic control of signaling ensures the necessary repetitiveness (which might be only fortuitous if the signals had to be learned).

(4) There is a case for asserting that the genotypic determination of adaptive characteristics is, in a special sense, more economical than the achievement of similar characteristic by somatic change or phenotypic learning. This matter has been argued elsewhere.[3] Briefly it is asserted that the somatic adaptive flexibility and/or learning capacity of any organism is limited and that the demands placed upon these capacities will be reduced by genotypic change in any appropriate direction. Such changes would therefore have

[2] J. E. Randall and H. S. Randall, "Examples of Mimicry and Protective Resemblance in Tropical Marine Fishes," *Bulletin of Marine Science of the Gulf and Caribbean*, 1960, 10: 444–80.

[3] G. Bateson, "The Role of Somatic Change in Evolution," *Evolution*, 1963, 17: 529–39.

survival value because they set free precious adaptive or learning capacity for other uses. This amounts to an argument for *Baldwin* effects. An extention of this argument would suggest that the iconic character of genotypically controlled signaling characteristics may, in some cases, be explained by supposing that these characteristics were once learned. (This hypothesis does not, of course, imply any sort of Lamarckian inheritance. It is obvious (1) that to fix the value of any variable in a homeostatic circuit by such inheritance would soon gum up the homeostatic system of the body, and (2) that no amount of modification of the dependent variables in a homeostatic circuit will change the bias of the circuit.)

(5) Last, it is unclear at what level genotypic determination of behavior might act. It was suggested above that iconic codes are easier for an organism to learn than more arbitrary codes. It is possible that the genotypic contribution to such an organism might take the form, not of fixing the given behavior, but rather of making this behavior easier to learn—a change in specific learning capacity rather than a change in genotypically determined behavior. Such a contribution from the genotype would have obvious advantages in that it would work along with ontogenetic change instead of working possibly at cross-purposes with it.

To sum up the argument so far:

(1) It is understandable that an early (in an evolutionary sense) method of creating redundancy would be the use of iconic part-for-whole coding. The external nonbiological universe contains redundancy of this kind, and in evolving a code of communication it is expectable that organisms would fall into the same trick. We have noted that the "part" can be split from the whole, so that a showing of the fangs can denote a possible but as yet nonexistent fight. All this provides an explanatory background for communication by means of "intention movements" and the like.

(2) It it partly understandable that such tricks of coding by iconic parts might become genotypically fixed.

(3) It has been suggested that the survival of such primitive (and therefore involuntary) signalling in human communication about personal relationship is explained by a need for honesty in such matters.

But the evolution of aniconic verbal coding remains unexplained.

We know from studies of aphasia, from Hockett's enumeration at this meeting of the characteristics of language and from elementary common sense that the component processes of creating and understanding verbal communication are many and that language fails when any one of those component processes is interrupted. It is possible that each of these processes should be the focus of a separate study. Here, however, I shall consider only one aspect of the matter: the evolution of simple indicative assertion.

An interesting intermediate between the iconic coding of animals and the verbal coding of human speech can be recognized in human dreaming and human myth. In psychoanalytic theory, the productions of dream process are said to be characterized by "primary-process" thinking.[4] Dreams, whether verbal or not, are to be considered as metaphoric statements, *i.e.*, the referents of dream are *relationships* which the dreamer, consciously or unconsciously, perceives in his waking world. As in all metaphor, the relata remain unmentioned and in their places appear other items such that the relationships between these substitute items shall be the same as those between the relata in the waking world.

To identify the relata in the waking world to which the dream refers would convert the metaphor into a simile, and, in general, dreams contain no message material which overtly performs this function. There is no signal in the dream which tells the dreamer that this is metaphor or what the referent of the metaphor may be. Similarly, dream contains no tenses. Time is telescoped, and representations of past events in real or distorted forms may have the present as their referent—or vice versa. The patterns of dream are timeless.

In a theater, the audience is informed by the curtain and the framing of the stage that the action on the stage is "only" a play. From within that frame the producers and actors may attempt to involve the audience in an illusion of reality as seemingly direct as the experience of dream. And,

[4] O. Fenichel, *Psychoanalytic Theory of Neurosis,* New York, Norton, 1945.

as in dream, the play has metaphoric reference to the outside world. But in dream, unless the sleeper be partly conscious of the fact of sleep, there is no curtain and no framing of the action. The partial negative—"This is *only* metaphor"—is absent.

I suggest that this absence of metacommunicative frames and the persistence in dream of pattern recognition are archaic characteristics in an evolutionary sense. If this be correct, then an understanding of dream should throw light both on how iconic communication operates among animals and on the mysterious evolutionary step from the iconic to the verbal.

Under the limitation imposed by the lack of a metacommunicative frame, it is clearly impossible for dream to make an indicative statement, either positive or negative. As there can be no frame which labels the content as "metaphoric," so there can be no frame to label the content as "literal." Dream can imagine rain or drought, but it can never assert "It is raining" or "It is not raining." Therefore, as we have seen, the usefulness in imagining "rain" or "drought" is limited to their metaphoric aspects.

Dream can *propose* the applicability of pattern. It can never assert or deny this applicability. Still less can it make an indicative statement about any identified referent, since no referent is identified.

The pattern is the thing.

These characteristics of dream may be archaic, but it is important to remember that they are not obsolete: that, as kinesic and paralinguistic communication has been elaborated into dance, music, and poetry, so also the logic of dream has been elaborated into theater and art. Still more astonishing is that world of rigorous fantasy which we call mathematics, a world forever isolated by its axioms and definitions from the possibility of making an indicative statement about the "real" world. Only *if* a straight line is the shortest distance between two points is the theorem of Pythagoras asserted.

The banker manipulates numerals according to rules supplied by the mathematician. These numerals are the names of numbers, and the numbers are somehow embodied in (real or fictitious) dollars. To remember what he is doing, the banker marks his numerals with labels, such as the dollar sign, but these are nonmathematical and no computer needs them. In the strictly mathematical procedure, as in the process of dream, the pattern of relationships controls all operations, but the relata are unidentified.

We return now to the contrast between the iconic method of creating redundancy in the universe, organism *plus* other organism, by the emission of parts of interactive patterns and the linguistic device of naming the relata. We noted above that the human communication which creates redundancy in the relationships between persons is still preponderantly iconic and is achieved by means of kinesics, paralinguistics, intention movements, actions, and the like. It is in dealing with the universe, message *plus* environment, that the evolution of verbal language has made the greatest strides.

In animal discourse, redundancy is introduced into this universe by signals which are iconic parts of the signaler's probable response. The environmental items may serve an ostensive function but cannot, in general, be mentioned. Similarly, in iconic communication about relationship, the relata—the organisms themselves—do not have to be identified because the subject of any predicate in this iconic discourse is the emitter of the signal, who is always ostensively present.

It appears then that at least two steps were necessary to get from the iconic use of parts of patterns of own behavior to the naming of entities in the external environment: there was both a change in coding and a change in the centering of the subject-predicate frame.

To attempt to reconstruct these steps can only be speculative, but some remarks may be offered:

(1) Imitation of environmental phenomena makes it possible to shift the subject-predicate frame from the self to some environmental entity while still retaining the iconic code.

(2) A similar shifting of the subject-predicate frame from self to other is latent in those interactions between animals in which A proposes a pattern of interaction and B negates this with an iconic or ostensive "don't." The subject of B's message here verbalized as "don't" is A.

(3) It is possible that the paradigms of interaction which are basic to iconic signaling about relationship could serve

as evolutionary models for the paradigms of verbal grammar. We should not, I suggest, think of the earliest rudiments of verbal communication as resembling what a man does with only a few words of a foreign language and no knowledge of its grammar and syntax. Surely, at all stages of the evolution of language, the communication of our ancestors was structured and formed—complete in itself, not made of broken pieces. The antecedents of grammar must surely be as old or older than the antecedents of words.

(4) For actions of the self, iconic abbreviations are readily available, and these control the vis-à-vis by implicit reference to interactional paradigms. But all such communication is necessarily positive. To show the fangs is to mention combat, and to mention combat is to propose it. There can be no simple iconic representation of a negative: no simple way for an animal to say "I will not bite you." It is easy, however, to imagine ways of communicating negative commands if (and *only* if) the other organism will first propose the pattern of action which is to be forbidden. By threat, by inappropriate response and so on, "don't" can be communicated. A pattern of interaction, offered by one organism, is negated by the other, who disrupts the proposed paradigm.

But "don't" is very different from "not." Commonly, the important message "I will not bite you" is generated as an *agreement* between two organisms following real or ritual combat. That is, the opposite of the final message is worked through to reach a *reductio ad absurdum* which can then be the basis of mutual peace, hierarchic precedence, or sexual relations. Many of the curious interactions of animals, called "play," which resemble (but are not) combat are probably the testing and reaffirmation of such negative agreement.

But these are cumbersome and awkward methods of achieving the negative.

(5) It was suggested above that the paradigms of verbal grammar might somehow be derived from the paradigms of interaction. We, therefore, look for the evolutionary roots of the simple negative among the paradigms of interaction. The matter, however, is not simple. What is known to occur at the animal level is the simultaneous presentation of contradictory signals—postures which mention both aggression and flight, and the like. These ambiguities are, however, quite different from the phenomenon familiar among humans where the friendliness of a man's words may be contradicted by the tension or aggressiveness of his voice or posture. The man is engaging in a sort of deceit, an altogether more complex achievement, while the ambivalent animal is offering positive alternatives. From neither of these patterns is it easy to derive a simple "not."

(6) From these considerations it appears likely that the evolution of the simple negative arose by introjection or imitation of the vis-à-vis, so that "not" was somehow derived from "don't."

(7) This still leaves unexplained the shift from communication about interaction patterns to communication about things and other components of the external world. This is the shift which determines that language would never make obsolete the iconic communication about the contingency patterns of personal relationship.

Further than that we cannot at present go. It is even possible that the evolution of verbal naming preceded the evolution of the simple negative. It is, however, important to note that evolution of a simple negative would be a decisive step toward language as we know it. This step would immediately endow the signals—be they verbal or iconic—with a degree of separateness from their referents, which would justify us in referring to the signals as "names." The same step would make possible the use of negative aspects of classification: those items which are not members of an identified class would become identifiable as nonmembers. And, lastly, simple affirmative indicative statements would become possible.