31

THE CYBERNETICS OF "SELF": A THEORY OF ALCOHOLISM

GREGORY BATESON

The present essay is based upon ideas which are, perhaps all of them, familiar either to psychiatrists who have had dealings with alcoholics, or to philosophers who have thought about the implications of cybernetics and systems theory. The only novelty which can be claimed for the thesis here offered derives from treating these ideas seriously as premises of argument and from the bringing together of commonplace ideas from two too-separate fields of thought.

In its first conception, this essay was planned to be a systems-theoretic study of alcoholic addiction, in which I would use data from the publications of Alcoholics Anonymous, which has the only outstanding record of success in dealing with alcoholics. It soon became evident, however, that the religious views and the organizational structure of A.A. presented points of great interest to systems theory, and that the correct scope of my study should include not only the premises of alcoholism but also the premises of the A.A. system of treating it and the premises of A.A. organization.

My debt to A.A. will be evident throughout—also, I hope, my respect for that organization and especially for the extraordinary wisdom of its co-founders, Bill W. and Dr. Bob.

In addition, I have to acknowledge a debt to a small sample of alcoholic patients with whom I worked intensively for about two years in 1949–52, in the Veterans Administration Hospital, Palo Alto, California. These men, it should be mentioned, carried other diagnoses—mostly of "schizophrenia"—in addition to the pains of alcoholism. Several were members of A.A. I fear that I helped them not at all.

Reprinted from *Psychiatry*, 34, 1–18. (1971).

THE PROBLEM

It is rather generally believed that "causes" or "reasons" for alcoholism are to be looked for in the sober life of the alcoholic. Alcoholics, in their sober manifestations, are commonly dubbed "immature," "maternally fixated," "oral," "homosexual," "passive-aggressive," "fearful of success," "over-sensitive," "proud," "affable," or simply "weak." But the logical implications of this belief are usually not examined:

(1) If the sober life of the alcoholic somehow drives him to drink or proposes the first step toward intoxication, it is not to be expected that any procedure which reinforces his particular style of sobriety will reduce or control his alcoholism.

(2) If his style of sobriety drives him to drink, then that style must contain error or pathology; and intoxication must provide some—at least subjective—correction of this error. In other words, compared with his sobriety, which is in some way "wrong," his intoxication must be in some way "right." The old tag "*In vino veritas*" may contain a truth more profound than is usually attributed to it.

(3) An alternative hypothesis would suggest that when sober, the alcoholic is somehow more sane than the people around him, and that this situation is intolerable. I have heard alcoholics argue in favor of this possibility, but I shall ignore it in this essay. I think that Bernard Smith, the non-alcoholic legal representative of A.A., came close to the mark when he said, "the [A.A.] member was never enslaved by alcohol. Alcohol simply served as an escape from *personal* enslavement in the false ideals of a meterialistic society" (A.A., 1957, p. 279, italics added). It is not a matter of revolt against insane ideals around him but of escaping from his own insane premises, which are continually reinforced by the surrounding society. It is possible, however, that the alcoholic is in some way more vulnerable or sensitive than the normal to the fact that his insane (but conventional) premises lead to unsatisfying results.

(4) The present theory of alcoholism, therefore, will provide a *converse matching* between the sobriety and the intoxication, such that the latter may be seen as an appropriate subjective correction for the former.

(5) There are, of course, many instances in which people resort to alcohol and even to extreme intoxication as an anesthetic giving release from ordinary grief, resentment, or physical pain. It might be argued that the anesthetic action of alcohol provides a sufficient converse matching for our theoretical purposes. I shall, however, specifically exclude these cases from consideration as being not relevant to the problem of addictive or repetitive alcoholism; and this in spite of the undoubted fact that "grief," "resentment," and "frustration" are commonly used by addicted alcoholics as *excuses* for drinking.

I shall demand, therefore, a converse matching between sobriety and intoxication more specific than that provided by mere anesthesia.

SOBRIETY

The friends and relatives of the alcoholic commonly urge him to be "strong," and to "resist temptation." What they mean by this is not very clear, but it is significant that the alcoholic himself—while sober—commonly agrees with their view of his "problem." He believes that he could be, or, at least, ought to be "the captain of his soul."[1] But it is a cliché of alcoholism that

[1]This phrase is used by A.A. in derision of the alcoholic who tries to use will power against the bottle. The quotation, along with the line, "my head is bloody but unbowed," comes from the poem "Invictus" by William Ernest Henley, who was a cripple but not an alcoholic. The use of the will to conquer pain and physical disability is probably not comparable to the alcoholic's use of will.

after "that first drink," the motivation to stop drinking is zero. Typically the whole matter is phrased overtly as a battle between "self" and "John Barleycorn." Covertly the alcoholic may be planning or even secretly laying in supplies for the next binge, but it is almost impossible (in the hospital setting) to get the sober alcoholic to plan his next binge in an overt manner. He cannot, seemingly, be the "captain" of his soul and overtly will or command his own drunkenness. The "captain" can only command sobriety—and then not be obeyed.

Bill W., the co-founder of Alcoholics Anonymous, himself an alcoholic, cut through all this mythology of conflict in the very first of the famous "Twelve Steps" of A.A. The first step demands that the alcoholic agree that he is powerless over alcohol. This step is usually regarded as a "*surrender*" and many alcoholics are either unable to achieve it or achieve it only briefly during the period of remorse following a binge. A.A. does not regard these cases as promising: they have not yet "hit bottom"; their despair is inadequate and after a more or less brief spell of sobriety they will again attempt to use "self-control" to fight the "temptation." They will not or cannot accept the premise that, drunk or sober, the total personality of an alcoholic is an alcoholic personality which cannot conceivably fight alcoholism. As an A.A. leaflet puts it, "trying to use will power is like trying to lift yourself by your bootstraps."

The first two steps of A.A. are as follows:

> 1. We admitted we were powerless over alcohol—that our lives had become unmanageable.
> 2. Came to believe that a Power greater than ourselves could restore us to sanity. [A.A., 1939]

Implicit in the combination of these two steps is an extraordinary—and I believe correct—idea: the experience of defeat not only serves to convince the alcoholic that change is necessary; it *is* the first step in that change. To be defeated by the bottle and to know it is the first "spiritual experience." The myth of self-power is thereby broken by the demonstration of a greater power.

In sum, I shall argue that the "sobriety" of the alcoholic is characterized by an unusually disastrous variant of the Cartesian dualism, the division between Mind and Matter, or, in this case, between conscious will, or "self," and the remainder of the personality. Bill W.'s stroke of genius was to break up with the first "step" the structuring of this dualism.

Philosophically viewed, this first step is *not* a surrender; it is simply a change in epistemology, a change in how to know about the personality-in-the-world. And, notably, the change is from an incorrect to a more correct epistemology.

EPISTEMOLOGY AND ONTOLOGY

Philosophers have recognized and separated two sorts of problem. There are first the problems of how things are, what is a person, and what sort of a world this is. These are the problems of ontology. Second, there are the problems of how we know anything, or more specifically, how we know what sort of a world it is and what sort of creatures we are that can know something (or perhaps nothing) of this matter. These are the problems of epistemology. To these questions, both ontological and epistemological, philosophers try to find true answers.

But the naturalist, observing human behavior, will ask rather different questions. If he be a cultural relativist, he may agree with those philosophers who hold that a "true" ontology is conceivable, but he will not ask whether the ontology of the people he observes is "true." He will expect their epistemology to be culturally determined or even idiosyncratic, and he will expect the culture as a whole to make sense in terms of their particular epistemology and ontology.

If, on the other hand, it is clear that the local epistemology is *wrong*, then the naturalist should be alert to the possibility that the culture as a whole will never really make "sense," or

will make sense only under restricted circumstances, which contact with other cultures and new technologies might disrupt.

In the natural history of the living human being, ontology and epistemology cannot be separated. His (commonly unconscious) beliefs about what sort of world it is will determine how he sees it and acts within it, and his ways of perceiving and acting will determine his beliefs about its nature. The living man is thus bound within a net of epistemological and ontological premises which—regardless of ultimate truth or falsity—become partially self-validating for him (Ruesch and Bateson, 1951, ch. 8).

It is awkward to refer constantly to both epistemology and ontology and incorrect to suggest that they are separable in human natural history. There seems to be no convenient word to cover the combination of these two concepts. The nearest approximations are "cognitive structure" or "character structure," but these terms fail to suggest that what is important is a body of habitual assumptions or premises implicit in the relationship between man and environment, and that these premises may be true or false. I shall therefore use the single term "epistemology" in this essay to cover both aspects of the net of premises which govern adaptation (or maladaptation) to the human and physical environment. In George Kelly's vocabulary, these are the rules by which an individual "construes" his experience.

I am concerned especially with that group of premises upon which Occidental concepts of the "self" are built, and, conversely, with premises which are corrective to some of the more gross Occidental errors associated with that concept.

THE EPISTEMOLOGY OF CYBERNETICS

What is new and surprising is that we now have partial answers to some of these questions. In the last twenty-five years extraordinary advances have been made in our knowledge of what sort of thing the environment is, what sort of thing an organism is, and, especially, what sort of thing a *mind* is. These advances have come out of cybernetics, systems theory, information theory, and related sciences.

We now know, with considerable certainty, that the ancient problem of whether the mind is immanent or transcendent can be answered in favor of immanence, and that this answer is more economical of explanatory entities than any transcendent answer: it has at least the negative support of "Occam's Razor."

On the positive side, we can assert that *any* on-going ensemble of events and objects which has the appropriate complexity of causal circuits and the appropriate energy relations will surely show mental characteristics. It will *compare*, that is, be responsive to *difference* (in addition to being affected by the ordinary physical "causes" such as impact or force). It will "process information" and will inevitably be self-corrective either toward homeostatic optima or toward the maximization of certain variables.

A "bit" of information is definable as a difference which makes a difference. Such a difference, as it travels and undergoes successive transformation in a circuit, is an elementary idea.

But, most relevant in the present context, we know that no part of such an internally interactive system can have unilateral control over the remainder or over any other part. The mental characteristics are inherent or immanent in the ensemble as a *whole*.

Even in very simple self-corrective systems, this holistic character is evident. In the steam engine with a "governor," the very word "governor" is a misnomer if it be taken to mean that this part of the system has unilateral control. The governor is, essentially, a sense organ or transducer which receives a transform of the *difference* between the actual running speed of the

engine and some ideal or preferred speed. This sense organ transforms these differences into differences in some efferent message, for example, to fuel supply or to a brake. The behavior of the governor is determined, in other words, by the behavior of the other parts of the system, and indirectly by its own behavior at a previous time.

The holistic and mental character of the system is most clearly demonstrated by this last fact, that the behavior of the governor (and, indeed, of every part of the causal circuit) is partially determined by its own previous behavior. Message material (i.e., successive transforms of difference) must pass around the total circuit, and the *time* required for the message material to return to the place from which it started is a basic characteristic of the total system. The behavior of the governor (or any other part of the circuit) is thus in some degree determined not only by its immediate past, but by what it did at a time which precedes the present by the interval necessary for the message to complete the circuit. There is thus a sort of determinative *memory* in even the simplest cybernetic circuit.

The stability of the system (i.e., whether it will act self-correctively or oscillate or go into run-away) depends upon the relation between the operational product of all the transformations of difference around the circuit and upon this characteristic time. The "governor" has no control over these factors. Even a human governor in a social system is bound by the same limitations. He is controlled by information from the system and must adapt his own actions to its time characteristics and to the effects of his own past action.

Thus: in no system which shows mental characteristics can any part have unilateral control over the whole. In other words, *the mental characteristics of the system are immanent, not in some part, but in the system as a whole.*

The significance of this conclusion appears when we ask, "Can a computer think?" or, "Is the mind in the brain?" And the answer to both questions will be negative unless the question is focused upon one of the few mental characteristics which are contained within the computer or the brain. A computer is self-corrective in regard to some of its internal variables. It may, for example, include thermometers or other sense organs which are affected by differences in its working temperature and the response of the sense organ to these differences may affect the action of a fan which in turn corrects the temperature. We may therefore say that the system shows mental characteristics in regard to its internal temperature. But it would be incorrect to say that the main business of the computer—the transformation of input differences into output differences—is "a mental process." The computer is only an arc of a larger circuit which always includes a man and an environment from which information is received and upon which efferent messages from the computer have effect. This total system, or ensemble, may legitimately be said to show mental characteristics. It operates by trial and error and has creative character.

Similarly, we may say that "mind" is immanent in those circuits of the brain which are complete within the brain. Or that mind is immanent in circuits which are complete within the system, brain *plus* body. Or, finally, that mind is immanent in the larger system—man *plus* environment.

In principle, if we desire to explain or understand the mental aspect of any biological event, we must take into account the system—that is, the network of *closed* circuits, within which that biological event is determined. But when we seek to explain the behavior of a man or any other organism, this "system" will usually *not* have the same limits as the "self"—as this term is commonly (and variously) understood.

Consider a man felling a tree with an axe. Each stroke of the axe is modified or corrected, according to the shape of the cut face of the tree left by the previous stroke. This self-corrective (i.e., mental) process is brought about by a total system, tree-eyes-brain-muscles-axe-stroke-tree; and it is this total system that has the characteristics of immanent mind.

More correctly, we should spell the matter out as: (differences in tree)-(differences in retina)-(differences in brain)-(differences in muscles)-(differences in movement of axe)-(differences in tree), etc. What is transmitted around the circuit is transforms of differences. And, as noted above, a difference which makes a difference in an *idea* or unit of information.

But this is *not* how the average Occidental sees the event-sequence of tree-felling. He says, "*I* cut down the tree" and he even believes that there is a delimited agent, the "self," which performed a delimited "purposive" action upon a delimited object.

It is all very well to say that "billiard ball A hit billiard ball B and sent it into the pocket"; and it would perhaps be all right (if we could do it) to give a complete hard science account of the events all around the circuit containing the man and the tree. But popular parlance includes *mind* in its utterance by invoking the personal pronoun, and then achieves a mixture of mentalism and physicalism by restricting mind within the man and reifying the tree. Finally the mind itself becomes reified by the notion that, since the "self" acted upon the axe which acted upon the tree, the "self" must also be a "thing." The parallelism of syntax between "*I* hit the billiard ball" and "the ball hit another ball" is totally misleading.

If you ask anybody about the localization and boundaries of the self, these confusions are immediately displayed. Or consider a blindman with a stick. Where does the blindman's self begin? At the tip of the stick? At the handle of the stick? Or at some point halfway up the stick? These questions are nonsense, because the stick is a pathway along which differences are transmitted under transformation, so that to draw a delimiting line *across* this pathway is to cut off a part of the systemic circuit which determines the blindman's locomotion.

Similarly, his sense organs are transducers or pathways for information, as also are his axons, etc. From a systems-theoretic point of view, it is a misleading metaphor to say that what travels in an axon is an "impulse." It would be more correct to say that what travels is a difference, or a transform of a difference. The metaphor of "impulse" suggests a hard-science line of thought which will ramify only too easily into nonsense about "psychic energy," and those who talk this kind of nonsense will disregard the information content of *quiescence*. The quiescence of an axon *differs* as much from activity as its activity does from quiescence. Therefore quiescence and activity have equal informational relevance. The message of activity can only be accepted as valid if the message of quiescence can also be trusted.

It is even incorrect to speak of the "message of activity" and the "message of quiescence." Always the fact that information is a transform of difference should be remembered, and we might better call the one message "activity—not quiescence" and the other "quiescence—not activity."

Similar considerations apply to the repentant alcoholic. He cannot simply elect "sobriety." At best he could only elect "sobriety—not drunkenness," and his universe remains polarized, carrying always both alternatives.

The total self-corrective unit which processes information, or, as I say, "thinks" and "acts" and "decides," is a *system* whose boundaries do not at all coincide with the boundaries either of the body or of what is popularly called the "self" or "consciousness"; and it is important to notice that there are *multiple* differences between the thinking system and the "self" as popularly conceived:

(1) The system is not a transcendent entity as the "self" is commonly supposed to be.

(2) The ideas are immanent in a network of causal pathways along which transforms of difference are conducted. The "ideas" of the system are in all cases at least binary in structure. They are not "impulses" but "information."

(3) This network of pathways is not bounded with consciousness but extends to include the pathways of all unconscious mentation—both autonomic and repressed, neural and hormonal.

(4) The network is not bounded by the skin but includes all external pathways along which information can travel. It also includes those effective differences which are immanent in the "objects" of such information. It includes the pathways of sound and light along which travel transforms of differences originally immanent in things and other people—and especially *in our own actions.*

It is important to note that the basic—and I believe erroneous—tenets of popular epistemology are mutually reinforcing. If, for example, the popular premise of transcendence is discarded, the immediate substitute is a premise of immanence in the body. But this alternative will be unacceptable because large parts of the thinking network are located outside the body. The so-called "Body-Mind" problem is wrongly posed in terms which force the argument toward paradox: if mind be supposed immanent in the body, then it must be transcendent. If transcendent, it must be immanent. And so on (cf. Collingwood).

Similarly, if we exclude the unconscious processes from the "self" and call them "ego-alien," then these processes take on the subjective coloring of "urges" and "forces"; and this pseudo-dynamic quality is then extended to the conscious "self" which attempts to "resist" the "forces" of the unconscious. The "self" thereby becomes itself an organization of seeming "forces." The popular notion which would equate "self" with consciousness thus leads into the notion that ideas are "forces"; and this fallacy is in turn supported by saying that the axon carries "impulses." To find a way out of this mess is by no means easy.

We shall proceed by first examining the structure of the alcoholic's polarization. In the epistemologically unsound resolution, "I will fight the bottle," what is supposedly lined up against what?

ALCOHOLIC "PRIDE"

Alcoholics are philosophers in that universal sense in which all human beings (and all mammals) are guided by highly abstract principles of which they are either quite unconscious, or unaware that the principle governing their perception and action is philosophic. A common misnomer for such principles is "feelings" (Bateson, 1963).

This misnomer arises naturally from the Anglo-Saxon epistemological tendency to reify or attribute to the body all mental phenomena which are peripheral to consciousness. And the misnomer is, no doubt, supported by the fact that the exercise and/or frustration of these principles is often accompanied by visceral and other bodily sensations. I believe, however, that Pascal was correct when he said, "The heart has its *reasons* of which the reason knows nothing."

But the reader must not expect the alcoholic to present a consistent picture. When the underlying epistemology is full of error, derivations from it are inevitably either self-contradictory or extremely restricted in scope. A consistent corpus of theorems cannot be derived from an inconsistent body of axioms. In such cases, the attempt to be consistent leads either to the great proliferation of complexity characteristic of psychoanalytic theory and Christian theology or to the extremely narrow view characteristic of contemporary behaviorism.

I shall therefore proceed to examine the "pride" which is characteristic of alcoholics to show that this principle of their behavior is derived from the strange dualistic epistemology characteristic of Occidental civilization.

A convenient way of describing such principles as "pride," "dependency," "fatalism," and

so forth, is to examine the principle as if it were a result of deutero-learning[2] (Bateson, 1942) and to ask what contexts of learning might understandably inculcate this principle.

(1) It is clear that the principle of alcoholic life which A.A. calls "pride" is not contextually structured around past achievement. They do not use the word to mean pride in something accomplished. The emphasis is not upon "I succeeded," but rather upon "I can . . ." It is an obsessive acceptance of a challenge, a repudiation of the proposition "I cannot."

(2) After the alcoholic has begun to suffer from—or be blamed for—alcoholism, this principle of "pride" is mobilized behind the proposition "I can stay sober." But, noticeably, success in this achievement destroys the "challenge." The alcoholic becomes "cocksure," as A.A. says. He relaxes his determination, risks a drink, and finds himself on a binge. We may say that the contextual structure of sobriety changes with its achievement. Sobriety, at this point, is no longer the appropriate contextual setting for "pride." It is the risk of the drink that now is challenging and calls out the fatal "I can. . . ."

(3) A.A. does its best to insist that this change in contextual structure shall never occur. They restructure the whole context by asserting over and over again that "*once an alcoholic, always an alcoholic.*" They try to have the alcoholic place alcoholism within the self, much as a Jungian analyst tries to have the patient discover his "psychological type" and to learn to live with the strengths and weaknesses of that type. In contrast, the contextual structure of alcoholic "pride" places the alcoholism *outside* the self: "*I* can resist drinking."

(4) The challenge component of alcoholic "pride" is linked with *risk-taking.* The principle might be put in words: "I can do something where success is improbable and failure would be disastrous." Clearly this principle will never serve to maintain continued sobriety. As success begins to appear probable, the alcoholic must challenge the risk of a drink. The element of "bad luck" or "probability" of failure places failure beyond the limits of the self. "If failure occurs, it is not *mine.*" Alcoholic "pride" progressively narrows the concept of "self," placing what happens outside its scope.

(5) The principle of pride-in-risk is ultimately almost suicidal. It is all very well to test once whether the universe is on your side, but to do so again and again, with increasing stringency of proof, is to set out on a project which can only prove that the universe hates you. But, still and all, the A.A. narratives show repeatedly that, at the very bottom of despair, *pride* sometimes prevents suicide. The final quietus must not be delivered by the "self" (cf. Bill's Story, A.A., 1939).

PRIDE AND SYMMETRY

The so-called pride of the alcoholic always presumes a real or fictitious "other" and its complete contextual definition therefore demands that we characterize the real or imagined relationship to this "other." A first step in this task is to classify the relationship as either

[2]This use of formal contextual structure as a descriptive device does not necessarily assume that the principle discussed is wholly or in part actually *learned* in contexts having the appropriate formal structure. The principle could have been genetically determined and it might still follow that the principle is best described by the formal delineation of the contexts in which it is exemplified. It is precisely this fitting of behavior to context that makes it difficult or impossible to determine whether a principle of behavior was genetically determined or learned in that context.

"symmetrical" or "complementary" (Bateson, 1936). To do this is not entirely simple when the "other" is a creation of the unconscious, but we shall see that the indications for such a classification are clear.

An explanatory digression is, however, necessary. The primary criterion is simple:

If, in a binary relationship, the behaviors of A and B are regarded (by A and B) as *similar* and are linked so that more of the given behavior by A stimulates more of it in B, and *vice versa,* then the relationship is "symmetrical" in regard to these behaviors.

If, conversely, the behaviors of A and B are *dissimilar* but mutually fit together (as, for example, spectatorship fits exhibitionism), and the behaviors are linked so that more of A's behavior stimulates more of B's fitting behavior, then the relationship is "complementary" in regard to these behaviors.

Common examples of simple symmetrical relationship are: armaments races, keeping up with the Joneses, athletic emulation, boxing matches, and the like. Common examples of complementary relationship are: dominance-submission, sadism-masochism, nurturance-dependency, spectatorship-exhibitionism, and the like.

More complex considerations arise when higher logical typing is present. For example: A and B may compete in gift-giving, thus superposing a larger symmetrical frame upon primarily complementary behaviors. Or, conversely, a therapist might engage in competition with a patient in some sort of play therapy, placing a complementary nurturant frame around the primarily symmetrical transactions of the game.

Various sorts of "double binds" are generated when A and B perceive the premises of their relationship in different terms—A may regard B's behavior as competitive when B thought he was helping A. And so on.

With these complexities we are not here concerned, because the imaginary "other" or counterpart in the "pride" of the alcoholic does not, I believe, play the complex games which are characteristic of the "voices" of schizophrenics.

Both complementary and symmetrical relationships are liable to progressive changes of the sort which I have called *schismogenesis* (Bateson, 1936). Symmetrical struggles and armaments races may, in the current phrase, "escalate"; and the normal pattern of succoring-dependency between parent and child may become monstrous. These potentially pathological developments are due to undamped or uncorrected positive feedback in the system, and may—as stated—occur in either complementary or symmetrical systems. However, in *mixed* systems schismogenesis is necessarily reduced. The armaments race between two nations will be slowed down by acceptance of complementary themes such as dominance, dependency, admiration, and so forth, between them. It will be speeded up by the repudiation of these themes.

This antithetical relationship between complementary and symmetrical themes is, no doubt, due to the fact that each is the logical opposite of the other. In a merely symmetrical armaments race, nation A is motivated to greater efforts by its estimate of *the greater strength* of B. When it estimates that B is weaker, nation A will relax its efforts. But the exact opposite will happen if A's structuring of the relationship is complementary. Observing that B is *weaker* than they, A will go ahead with hopes of conquest (cf. Bateson, 1946, and Richardson, 1935).

This antithesis between complementary and symmetrical patterns may be more than simply logical. Notably, in psychoanalytic theory (cf. Erikson, 1937), the patterns which are called "libidinal" and which are modalities of the erogenous zones are all *complementary.* Intrusion, inclusion, exclusion, reception, retention, and the like—all of these are classed as "libidinal." Whereas rivalry, competition, and the like fall under the rubric of "ego" and "defense."

It is also possible that the two antithetical codes—symmetrical and complementary—may be physiologically represented by contrasting states of the central nervous system. The progressive changes of schismogenesis may reach climactic discontinuities and sudden reversals. Symmetrical

rage may suddenly turn to grief; the retreating animal with tail between his legs may suddenly "turn at bay" in a desperate battle of symmetry to the death. The bully may suddenly become the coward when he is challenged, and the wolf who is beaten in a symmetrical conflict may suddenly give "surrender" signals which prevent further attack.

The last example is of special interest. If the struggle between the wolves is symmetrical—that is, if wolf A is stimulated to more aggressive behavior by the aggressive behavior of B—then if B suddenly exhibits what we may call "negative aggression," A will not be able to continue to fight unless he can quickly switch over to that complementary state of mind in which B's weakness would be a stimulus for his aggression. Within the hypothesis of symmetrical and complementary modes, it becomes unnecessary to postulate a specifically "inhibitory" effect for the surrender signal.

Human beings who possess language can apply the label "aggression" to all attempts to damage the other, regardless of whether the attempt is prompted by the other's strength or weakness; but at the prelinguistic mammalian level these two sorts of "aggression" must appear totally different. We are told that from the lion's point of view, an "attack" on a zebra is totally different from an "attack" on another lion (Lorenz).

Enough has now been said so that the question can be posed: Is alcoholic pride contextually structured in symmetrical or complementary form?

First, there is a very strong tendency toward symmetry in the normal drinking habits of Occidental culture. Quite apart from addictive alcoholism, two men drinking together are impelled by convention to match each other, drink for drink. At this stage, the "other" is still real and the symmetry, or rivalry, between the pair is friendly.

As the alcoholic becomes addicted and tries to resist drinking, he begins to find it difficult to resist the social context in which he should match his friends in their drinking. The A.A. says, "Heaven knows, we have tried hard enough and long enough to drink like other people!"

As things get worse, the alcoholic is likely to become a solitary drinker and to exhibit the whole spectrum of response to challenge. His wife and friends begin to suggest that his drinking is a *weakness,* and he may respond, with symmetry, both by resenting them and by asserting his strength to resist the bottle. But, as is characteristic of symmetrical responses, a brief period of successful struggle weakens his motivation and he falls off the wagon. Symmetrical effort requires continual opposition from the opponent.

Gradually the focus of the battle changes, and the alcoholic finds himself committed to a new and more deadly type of symmetrical conflict. He must now prove that the bottle cannot kill him. His "head is bloody but unbowed." He is still the "captain of his soul"—for what it's worth.

Meanwhile, his relationships with wife and boss and friends have been deteriorating. He never did like the complementary status of his boss as an authority; and now as he deteriorates his wife is more and more forced to take a complementary role. She may try to exert authority, or she becomes protective, or she shows forbearance, but all those provoke either rage or shame. His symmetrical "pride" can tolerate no complementary role.

In sum, the relationship between the alcoholic and his real or fictitious "other" is clearly symmetrical and clearly schismogenic. It escalates. We shall see that the religious conversion of the alcoholic when saved by A.A. can be described as a dramatic shift from this symmetrical habit, or epistemology, to an almost purely complementary view of his relationship to others and to the universe or God.

PRIDE OR INVERTED PROOF?

Alcoholics may appear to be stiffnecked, but they are not stupid. The part of the mind in which their policy is decided certainly lies too deep for the word "stupidity" to be applicable.

These levels of the mind are prelinguistic and the computation which goes on there is coded in *primary process.*

Both in dream and in mammalian interaction, the only way to achieve a proposition which contains its own negation ("I will not bite you," or "I am not afraid of him") is by an elaborate imagining or acting out of the proposition to be negated, leading to a *reductio ad absurdum.* "I will not bite you" is achieved between two mammals by an experimental combat which is a "not combat," sometimes called "play." It is for this reason that "agnostic" behavior commonly evolves into friendly greeting (Bateson, 1969).

In this sense, the so-called pride of the alcoholic is in some degree ironic. It is a determined effort to test something like "self-control" with an ulterior but unstateable purpose of proving that "self-control" is ineffectual and absurd. "It simply won't work." This ultimate proposition, since it contains a simple negation, is not to be expressed in primary process. Its final expression is in an action—the taking of a drink. The heroic battle with the bottle, that fictitious "other," ends up in a "kiss and make friends."

In favor of this hypothesis, there is the undoubted fact that the testing of self-control leads back into drinking. And, as I have argued above, the whole epistemology of self-control which his friends urge upon the alcoholic is monstrous. If this be so, then the alcoholic is right in rejecting it. He has achieved a *reductio ad absurdum* of the conventional epistemology.

But this description of achieving a *reductio ad absurdum* verges upon teleology. If the proposition "It won't work" cannot be entertained within the coding of primary process, how then can the computations of primary process direct the organism to try out those courses of action which will demonstrate that "It won't work"?

Problems of this general type are frequent in psychiatry and can perhaps only be resolved by a model in which, under certain circumstances, the organism's discomfort activates a positive feedback loop to *increase* the behavior which preceded the discomfort. Such positive feedback would provide a verification that it was really that particular behavior which brought about the discomfort, and might increase the discomfort to some threshold level at which change would become possible.

In psychotherapy such a positive feedback loop is commonly provided by the therapist who pushes the patient in the direction of his symptoms—a technique which has been called the "therapeutic double bind." An example of this technique is quoted later in this essay, where the A.A. member challenges the alcoholic to go and do some "controlled drinking" in order that he may discover for himself that he has no control.

It is also usual that the symptoms and hallucinations of the schizophrenic—like dreams—constitute a corrective experience, so that the whole schizophrenic episode takes on the character of a self-initiation. Barbara O'Brien's account of her own psychosis is perhaps the most striking example of this phenomenon, which has been discussed elsewhere (Bateson, 1961, Introduction).

It will be noted that the possible existence of such a positive feedback loop, which will cause a runaway in the direction of increasing discomfort up to some threshold (which might be on the other side of death), is not included in conventional theories of learning. But a tendency to verify the unpleasant by seeking repeated experience of it is a common human trait. It is perhaps what Freud called the "death instinct."

THE DRUNKEN STATE

What has been said above about the treadmill of symmetrical pride is only one half of the picture. It is the picture of the state of mind of the alcoholic *battling* with the bottle. Clearly this state is very unpleasant and clearly it is also unrealistic. His "others" are either totally

imaginary or are gross distortions of persons on whom he is dependent and whom he may love. He has an alternative to this uncomfortable state—he can get drunk. Or, "*at least*," have a drink.

With this complementary surrender, which the alcoholic will often see as an act of spite—a Parthian dart in a symmetrical struggle—his entire epistemology changes. His anxieties and resentments and panic vanish as if by magic. His self-control is lessened, but his need to compare himself with others is reduced even further. He feels the physiological warmth of alcohol in his veins and, in many cases, a corresponding psychological warmth toward others. He may be either maudlin or angry, but he has at least become again a part of the human scene.

Direct data bearing upon the thesis that the step from sobriety into intoxication is also a step from symmetrical challenge into complementarity are scarce, and always confused both by the distortions of recall and by the complex toxicity of the alcohol. But there is strong evidence from song and story to indicate that the step is of this kind. In ritual, partaking of wine has always stood for the social aggregation of persons united in religious "communion" or secular *Gemütlichkeit*. In a very literal sense, alcohol supposedly makes the individual see himself as and act as *a part of* the group. That is, it enables complementarity in the relationships which surround him.

HITTING BOTTOM

A.A. attaches great importance to this phenomenon and regards the alcoholic who has not hit bottom as a poor prospect for their help. Conversely, they are inclined to explain their failures by saying that the individual who goes back to his alcoholism has not yet "hit bottom."

Certainly many sorts of disaster may cause an alcoholic to hit bottom. Various sorts of accidents, an attack of delirium tremens, a patch of drunken time of which he has no memory, rejection by wife, loss of job, hopeless diagnosis, and so on—any of these may have the required effect. A.A. says that "bottom" is different for different men and some may be dead before they reach it.[3]

It is possible, however, that "bottom" is reached many times by any given individual; that "bottom" is a spell of panic which provides a favorable moment for change, but not a moment at which change is inevitable. Friends and relatives and even therapists may pull the alcoholic out of his panic, either with drugs or reassurance, so that he "recovers" and goes back to his "pride" and alcoholism—only to hit a more disastrous "bottom" at some later time, when he will again be ripe for a change. The attempt to change the alcoholic in a period *between* such moments of panic is unlikely to succeed.

The nature of the panic is made clear by the following description of a "test."

> We do not like to pronounce any individual as alcoholic, but you can quickly diagnose yourself. Step over to the nearest barroom and try some controlled drinking. Try to drink and stop abruptly. Try it more than once. It will not take long for you to decide, if you are honest with yourself about it. It may be worth a bad case of jitters if you get a full knowledge of your condition. [A.A., 1939, p. 43]

We might compare the test quoted above to commanding a driver to brake suddenly when travelling on a slippery road: he will discover fast that his control is limited. (The metaphor "skid row" for the alcoholic section of town is not inappropriate.)

[3]Personal communication from a member.

The panic of the alcoholic who has hit bottom is the panic of the man who thought he had control over a vehicle but suddenly finds that the vehicle can run away with him. Suddenly, pressure on what he knows is the brake seems to make the vehicle go faster. It is the panic of discovering that *it* (the system, self *plus* vehicle) is bigger than he is.

In terms of the theory here presented, we may say that hitting bottom exemplifies systems theory at three levels:

(1) The alcoholic works on the discomforts of sobriety to a threshold point at which he has bankrupted the epistemology of "self-control." He then gets drunk—because the "system" is bigger than he is—and he may as well surrender to it.

(2) He works repeatedly at getting drunk until he proves that there is a still larger system. He then encounters the panic of "hitting bottom."

(3) If friends and therapists reassure him, he may achieve a further unstable adjustment—becoming addicted to their help—until he demonstrates that this system won't work, and "hits bottom" again but at a lower level. In this, as in all cybernetic systems, the sign (plus or minus) of the effect of any intrusion upon the system depends upon timing.

(4) Lastly, the phenomenon of hitting bottom is complexly related to the experience of double bind (Bateson et al., 1956). Bill W. narrates that he hit bottom when diagnosed as a hopeless alcoholic by Dr. William D. Silkworth in 1939, and this event is regarded as the beginning of A.A. history (A.A., 1957, p. vii). Dr. Silkworth also "supplied us with the tools with which to puncture the toughest alcoholic ego, those shattering phrases by which he described our illness: *the obsession of the mind* that compels us to drink and *the allergy of the body* that condemns us to go mad or die" (Bill W. in A.A., 1957, p. 13; italics in the original). This is a double bind correctly founded upon the alcoholic's dichotomous epistemology of mind *versus* body. He is forced by these words back and back to the point at which only an involuntary change in deep unconscious epistemology—a spiritual experience—will make the lethal description irrelevant.

THE THEOLOGY OF ALCOHOLICS ANONYMOUS

Some outstanding points of the theology of A.A. are:

(1) *There is a Power greater than the self.* Cybernetics would go somewhat further and recognize that the "self" as ordinarily understood is only a small part of a much larger trial-and-error system which does the thinking, acting, and deciding. This system includes all the informational pathways which are relevant at any given moment to any given decision. The "self" is a false reification of an improperly delimited part of this much larger field of interlocking processes. Cybernetics also recognizes that two or more persons—any group of persons—may together form such a thinking-and-acting system.

(2) This Power is felt to be personal and to be intimately linked with each person. It is "God as *you* understand him to be."

Cybernetically speaking, "my" relation to any larger system around me and including other things and persons will be different from "your" relation to some similar system around you. The relation "part of" must necessarily and logically always be complementary but the meaning of the phrase "part of" will be different for every person.[4] This difference will be especially important in systems containing more than one person. The system or "power" must necessarily appear different from where each person sits. Moreover, it is expectable that such systems, when they

[4]This diversity in styles of integration could account for the fact that some persons become alcoholic while others do not.

encounter each other, will recognize each other as systems in this sense. The "beauty" of the woods through which I walk is my recognition both of the individual trees and of the total ecology of the woods as *systems*. A similar esthetic recognition is still more striking when I talk with another person.

(3) A favorable relationship with this Power is discovered through "hitting bottom" and "surrender."

(4) By resisting this Power, men and especially alcoholics bring disaster upon themselves. The materialistic philosophy which sees "man" as pitted against his environment is rapidly breaking down as technological man becomes more and more able to oppose the largest systems. Every battle that he wins brings a threat of disaster. The unit of survival—either in ethics or in evolution—is not the organism or the species but the largest system or "power" within which the creature lives. If the creature destroys its environment, it destroys itself.

(5) But—and this is important—the Power does not reward and punish. It does not have "power" in that sense. In the biblical phrase, "All things work together for good to them that love God." And, conversely, to them that do not. The idea of power in the sense of unilateral control is foreign to A.A. Their organization is strictly "democratic" (their word), and even their deity is still bound by what we might call a systemic determinism. The same limitation applies both to the relationship between the A.A. sponsor and the drunk whom he hopes to help and to the relationship between A.A. central office and every local group.

(6) The first two "steps" of Alcoholics Anonymous taken together identify the addiction as a manifestation of this Power.

(7) The healthy relation between each person and this Power is complementary. It is in precise contrast to the "pride" of the alcoholic, which is predicated upon a symmetrical relationship to an imagined "other." The schismogenesis is always more powerful than the participants in it.

(8) The quality and content of each person's relation to the Power is indicated or reflected in the social structure of A.A. The secular aspect of this system—its governance—is delineated in "Twelve Traditions" (A.A., 1957), which supplement the "Twelve Steps," the latter developing man's relationship to the Power. The two documents overlap in the Twelfth Step, which enjoins aid to other alcoholics as a necessary spiritual exercise without which the member is likely to relapse. The total system is a Durkheimian religion in the sense that the relationship between man and his community parallels the relationship between man and God. "A.A. is a power greater than any of us" (A.A., 1957, p. 288).

In sum, the relationship of each individual to the "Power" is best defined in the words "*is part of.*"

(9) Anonymity. It must be understood that anonymity means much more in A.A. thinking and theology than the mere protection of members from exposure and shame. With increasing fame and success of the organization as a whole, it has become a temptation for members to use the fact of their membership as a positive asset in public relations, politics, education, and many other fields. Bill W., the co-founder of the organization, was himself caught by this temptation in early days and has discussed the matter in a published article (A.A., 1957, pp. 286–294). He sees first that any grabbing of the spotlight must be a personal and spiritual danger to the member, who cannot afford such self-seeking; and beyond this that it would be fatal for the organization as a whole to become involved in politics, religious controversy, and social reform. He states clearly that the errors of the alcoholic are the same as the "forces which are today ripping the world apart at its seams," but that it is not the business of A.A. to save the world. Their single purpose is "to carry the A.A. message to the sick alcoholic who wants it" (loc. cit.). He concludes that anonymity is "the greatest symbol of self-sacrifice that we know."

Elsewhere the twelfth of the "Twelve Traditions" states that "anonymity is the spiritual foundation of our traditions, ever reminding us to place principles before personalities."

To this we may add that anonymity is also a profound statement of the systemic relation, part-to-whole. Some systems theorists would go even further, because a major temptation for systems theory lies in the reification of theoretical concepts. Anatol Holt says he wants a bumper sticker which would (paradoxically) say, "Stamp out nouns" (Wenner-Gren Foundation).

(10) Prayer. The A.A. use of prayer similarly affirms the complementarity of part-whole relationship by the very simple technique of asking for that relationship. They ask for those personal characteristics, such as humility, which are in fact exercised in the very act of prayer. If the act of prayer be sincere (which is not so easy), God cannot but grant the request. And this is peculiarly true of "God, *as you understand him.*" This self-affirming tautology, which contains its own beauty, is precisely the balm required after the anguish of the double binds which went with hitting bottom.

Somewhat more complex is the famous "Serenity Prayer": "God grant us the serenity to accept the things we cannot change, courage to change the things we can, and wisdom to know the difference."[5]

If double binds cause anguish and despair and destroy personal epistemological premises at some deep level, then it follows, conversely, that for the healing of these wounds and the growth of a new epistemology, some converse of the double bind will be appropriate. The double bind leads to the conclusion of despair, "There are no alternatives." The Serenity Prayer explicitly frees the worshipper from these maddening bonds.

In this connection it is worth mentioning that the great schizophrenic, John Perceval, observed a change in his "voices." In the beginning of his psychosis they bullied him with "contradictory commands" (or as I would say, double binds), but later he began to recover when they offered him choice of clearly defined alternatives (cf. Bateson, 1961).

(11) In one characteristic, A.A., differs profoundly from such natural mental systems as the family or the redwood forest. It has a *single* purpose—"to carry the A.A. message to the sick alcoholic who wants it"—and the organization is dedicated to the maximization of that purpose. In this respect, A.A. is no more sophisticated than General Motors or an Occidental nation. But biological systems, other than those premised upon Occidental ideas (and especially *money*), are multipurposed. There is no single variable in the redwood forest of which we can say that the whole system is oriented to maximizing that variable and all other variables are subsidiary to it; and, indeed, the redwood forest works toward optima, not maxima. Its needs are satiable, and too much of anything is toxic.

There is, however, this: that the single purpose of A.A. is directed outward and is aimed at a noncompetitive relationship to the larger world. The variable to be maximized is a complementarity and is of the nature of "service" rather than dominance.

THE EPISTEMOLOGICAL STATUS OF COMPLEMENTARY AND SYMMETRICAL PREMISES

It was noted above that in human interaction, symmetry and complementarity may be complexly combined. It is therefore reasonable to ask how it is possible to regard these themes as so fundamental that they shall be called "epistemological," even in a natural history study of cultural and interpersonal premises.

[5]This was not originally an A.A. document and its authorship is unknown. Small variations in the text occur. I have quoted the form which I personally prefer (A.A., 1957, p. 196).

The answer seems to hang upon what is meant by "fundamental" in such a study of man's natural history; and the word seems to carry two sorts of meaning.

First, I call *more fundamental* those premises which are the more deeply embedded in the mind, which are the more "hard programmed" and the less susceptible to change. In this sense, the symmetrical pride or hubris of the alcoholic is fundamental.

Second, I shall call more fundamental those premises of mind which refer to the larger rather than the smaller systems or gestalten of the universe. The proposition "grass is green" is less fundamental than the proposition "color differences make a difference."

But, if we ask about what happens when premises are changed, it becomes clear that these two definitions of the "fundamental" overlap to a very great extent. If a man achieves or suffers change in premises which are deeply embedded in his mind, he will surely find that the results of that change will ramify throughout his whole universe. Such changes we may well call "epistemological."

The question then remains regarding what is epistemologically "right" and what is epistemologically "wrong." Is the change from alcoholic symmetrical "pride" to the A.A. species of complementarity a correction of his epistemology? And is complementarity *always* somehow better than symmetry?

For the A.A. member, it may well be true that complementarity is always to be preferred to symmetry and that even the trivial rivalry of a game of tennis or chess may be dangerous. The superficial episode may touch off the deeply embedded symmetrical premise. But this does not mean that tennis and chess propose epistemological error for everybody.

The ethical and philosophic problem really concerns only the widest universe and the deepest psychological levels. If we deeply and even unconsciously believe that our relation to the largest system which concerns us—the "Power greater than self"—is symmetrical and emulative, then we are in error.

LIMITATIONS OF THE HYPOTHESIS

Finally, the above analysis is subject to the following limitations and implications:

(1) It is not asserted that all alcoholics operate according to the logic which is here outlined. It is very possible that other types of alcoholics exist and almost certain that alcoholic addiction in other cultures will follow other lines.

(2) It is not asserted that the way of Alcoholics Anonymous is the *only* way to live correctly or that their theology is the only correct derivation from the epistemology of cybernetics and systems theory.

(3) It is not asserted that all transactions between human beings ought to be complementary, though it is clear that the relation between the individual and the larger system of which he is a part must necessarily be so. Relations between persons will (I hope) always be complex.

(4) It is, however, asserted that the non-alcoholic world has many lessons which it might learn from the epistemology of systems theory and from the ways of A.A. If we continue to operate in terms of a Cartesian dualism of mind *versus* matter, we shall probably also continue to see the world in terms of God *versus* man; elite *versus* people; chosen race *versus* others; nation *versus* nation; and man *versus* environment. It is doubtful whether a species having *both* an advanced technology *and* this strange way of looking at its world can endure.

REFERENCES

[Alcoholics Anonymous.] *Alcoholics Anonymous*; New York, Works Publishing, 1939.

[Alcoholics Anonymous.] *Alcoholics Anonymous Comes of Age*; Harper, 1957.

Bateson, Gregory, *Naven*; Cambridge Univ. Press, 1936.

Bateson, Gregory. "Social Planning and the Concept of 'Deutero-learning,' " in *Conference on Science, Philosophy, and Religion, Second Symposium*; Harper, 1942.

Bateson, Gregory. "The Pattern of an Armaments Race—Part I: An Anthropological Approach," *Bull. Atomic Scientists* (1946) 2(5):10–11.

Bateson, Gregory, et al. "Toward a Theory of Schizophrenia," *Behavioral Science* (1956) 1:251–264.

Bateson, Gregory. (Ed.) *Perceval's Narrative*; Stanford Univ. Press, 1961.

Bateson, Gregory. "A Social Scientist Views the Emotions," in P. Knapp (Ed.), *Expression of the Emotions in Man*; Internat. Univ. Press, 1963.

Bateson, Gregory. "Metalogue: What Is an Instinct?," in T. Sebeok (Ed.), *Approaches to Animal Communication*; The Hague, Mouton, 1969.

Collingwood, Robin George. *The Idea of Nature*; Oxford Univ. Press, 1945.

Erikson, Erik H. "Configurations in Play—Clinical Notes," *Psychoanal. Quart.* (1937) 6:139–214.

Lorenz, Konrad. *On Aggression*; Harcourt, Brace & World, 1966.

O'Brien, Barbara. *Operators and Things: The Inner Life of a Schizophrenic*; Cambridge, Mass., Arlington Books, 1958.

Richardson, Lewis Fry. "Generalized Foreign Politics," *British J. Psychology*, Monogr. Supplements, 1939.

Ruesch, Jurgen, and Bateson, Gregory. *Communication: The Social Matrix of Psychiatry*; Norton, 1951.

Wenner-Gren Foundation. Conference on the effects of conscious purpose on human adaptation, 1968; Knopf, in press.