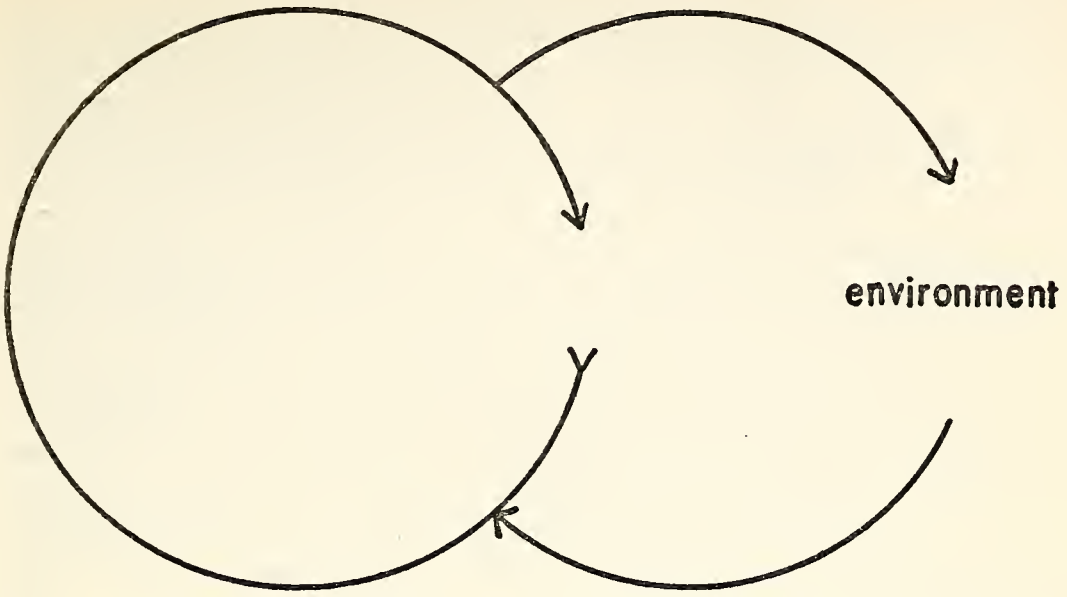
gressional spontaneity which many Occidentals envy. Occidental cultures seem often to promote a compulsive categorization of the details of behavior, while leaving the individual a greater freedom to act in terms of progressional integration in regard to the wider decisions. These generalizations are, however, liable to be reversed or modified from individual to individual.

DIVERSITIES OF CODIFICATION

As mentioned above, people vary in the degree to which they perceive and act upon their own and other people's words and actions as "reports" or commands. Likewise, people vary in the degree to which they operate selectively or progressionally. The present section follows up these two statements of variation among individuals with an attempt to construct a general scheme, sufficiently abstract to classify the various orders of contrast in codification and evaluation which may conceivably occur among individuals. To enumerate all varieties of human codification and evaluation would be a superhuman task; all that we here attempt is to set up a framework of categories within which the many varieties may be interrelated—a task which should not be beyond the wit of man.

In order to set up such a scheme, it is convenient to start from a non-human model which will be totally incapable of the complex codifications and evaluations which are characteristic of man. We can then build up our notions about man by deliberately and systematically anthropomorphizing the model.

Figure 1 is offered as such a model, and represents the minimum system about which we can meaningfully talk. The arrows represent causal chains, and the entire diagram represents an entity consisting of an internal self-corrective causal circuit which acts upon and is acted upon by an environment. The reader who desires a more concrete picture may imagine, if he pleases, either a protozoan engaged in positive heliotropism or a servomechanism seeking a target. In any case, however the model is embodied, it is certainly incapable of the complex codifications and evaluations which are relevant for this study. At most it may distinguish elements of the environment ("light,"



"no light," etc.), but certainly it will not be capable of conceptualizing such notions as "I perceive light," "I seek light," "The perception of light is a pleasure," or "The light compels me to go toward it." In the actions and self-corrections of such simple tropic systems, no evaluative principles of these higher types can be perceived, and it is such principles as these that we seek to classify. The model, then, has this usefulness: it presents us with a *carte blanche* for systematic anthropomorphizing.

We shall later attempt to consider the more complex case of interaction between two such models, exemplifying the possibilities for codification and evaluation of the processes of interaction between persons. Here for heuristic simplicity, we deal first with the organisms vis-à-vis an impersonal environment and mention those types of complexity conceivable in this simple but unreal case.

The following types of codification-evaluation present themselves, and are listed according to a logical order. (We do not suggest that evolution followed this course.)

1. *Discrimination of perceived entities in that arc of the total circuit which we call environment.* We here refer to the recognition of environmental Gestalten of various types and to the classification and delimitation of these ("oak trees," "light,"

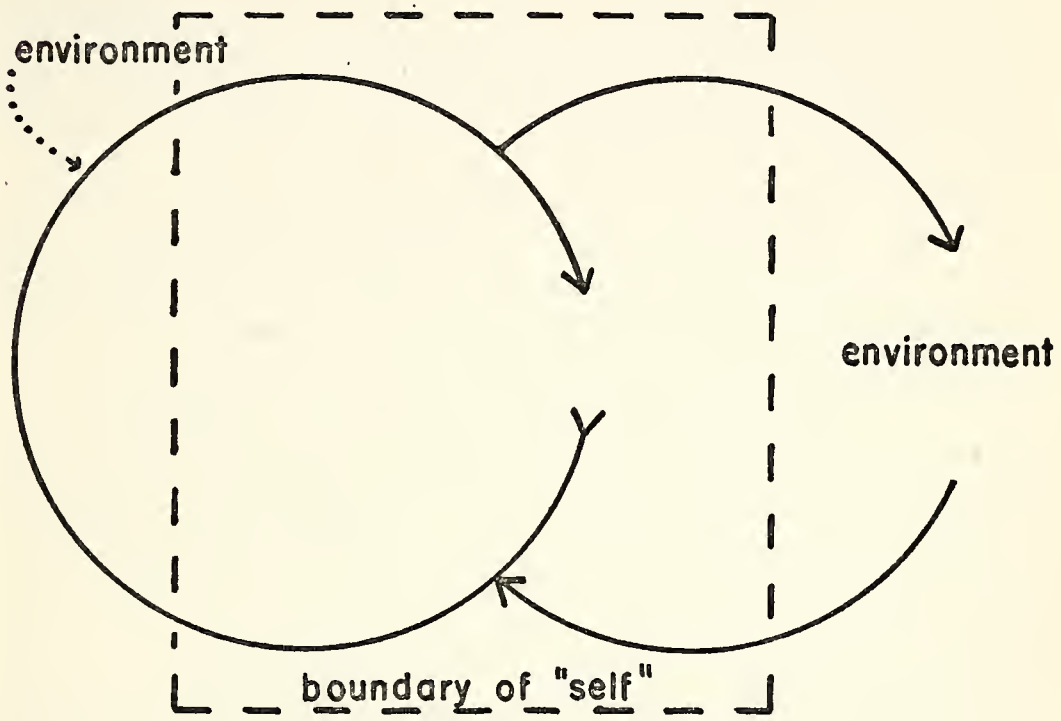
etc.). This categorization is trivial so far as our present study is concerned, because it is relatively easy in ongoing communication between persons to iron out misunderstandings at this level, and especially easy to do so when verbal communication can be supplemented by pointing at concrete objects and events. For purposes of mapping the possible diversities of codification, we note that the organism's perceived Gestalten are in all cases arbitrary but interdependent. The organism is, like the scientist, free to delimit whatever systems and entities it pleases in the external world; but certain entities having been discriminated, later discriminations will follow according to the system of discrimination to which the organism is committed by the earlier discriminations. Having discriminated "oak trees" from "light," the discrimination of "elms" from "oaks" and of "blue" from "red" is likely to follow.

2. *Subdivision of that sub-circuit which we call the organism.* Here we refer to the organism's recognition and discrimination of parts of the body, sensations, actions, and the like. With this step it becomes possible to associate sensations with parts of the body, and perhaps to conceptualize purposes in terms of localized sensation—already a conceptual narrowing of the totality of interaction between the organism and the environment. Also, perhaps with this step it becomes possible to falsify parts of the body image and to project such falsification onto other parts of the total circuit, especially onto the environment.

3. *Subdivision of the total circuit into two parts, the self and the environment.* In ontogeny this step is no doubt facilitated by the presence of other similar organisms and by the recognition of these as similar to the self, and it is possible that no concept of the self could be reached in the absence of other similar organisms. But, be that as it may, the differentiation of the self from the environment is conceptually possible without the presence of other organisms and therefore is discussed at this point. Such a differentiation, like all those with which we deal, is in a certain sense arbitrary, and its arbitrary nature is clear when we consider the simplified model from which we started. The imaginary organism is free to draw a closed line anywhere on this diagram and to regard everything inside this

line as "self" while everything on the outside is "environment"; and indeed the usefulness of the model is that it stresses this freedom.

Figure 2 represents the case in which the organism includes



within the self various objects and events outside his skin but intimately connected with him, while he labels as parts of the environment certain of his own body parts or functions of which he is perhaps dimly aware or over which he feels that he has no control. There is, in fact, no right way of delimiting the self; failure to communicate, frustration, and ultimately hostility and pathology may follow if organisms who have conflicting premises on this subject seek to communicate. Further, their communication will be rendered the more difficult inasmuch as neither is likely to be fully aware of what he himself includes in his concept of self. To be able to conceptualize "I include such-and-such in my 'self' " is already a much more complex achievement than the simpler "I am and there are things which are not me."

Recognition of the phenomena of conceptualization will be discussed below.

4. *Conceptualization of control between self and environment.* This again is a step beyond the mere differentiation of self from environment and enables the organism to perceive the environment as coercive or to see himself as coercing the environment—either of which notions is usually a false simplification of the reality of interaction. The diversities of codification will include all those attributions of passivity and activity to the self and the environment.

5. *Conceptualization of separate causal arcs within the self.* Here we refer to such premises as "I am the captain of my soul" and the mind-body dichotomy. All of these are perhaps derivative from a codification, which would identify parts of the organism itself as "environmental," combined with premises about control of the environment or control by the environment. In fact, the supposed internal splitting of the individual is likely to be a symbolic echo of the presumed relations between the self and the environment; or, vice versa, splits within the self may be expressed in premises regarding the relationship between self and environment.

6. *Multiple levels of abstraction.* In the preceding paragraph a step of a special kind was introduced. It was there supposed that the organism may adopt as a device for codifying the relations between causal arcs within the self, certain premises formerly used in the codification of a relationship between the self and the environment. We note in passing that such a proceeding, based on analogy, may well lead to error, but does not of necessity do so. It is, in fact, the possibility and nature of such steps in codification that concern us, rather than their validity. In sum, the organism perceives Gestalt A and Gestalt B and codifies on the assumption that there is a relationship (of likeness or unlikeness) between A and B. This proceeding involves, either explicitly or implicitly, a higher level of abstraction than that involved in the primary codification of the two Gestalten. The "likeness" or "unlikeness" is more abstract than either Gestalt A or Gestalt B. By steps of this kind the codification system of the organism becomes more and more elaborate

and may contain many levels of abstraction, the interrelations of which are capable of great diversity.

7. *Gestalten involving time spans of varying lengths.* The organism may see a single movement as an "act" or may see whole sequences of events, including both his own actions and results of that action as having a unity or purpose or failure. He may even conceptualize the limitations of his own life and with such notions, symbolically extended, he may see an initiation ceremony or even his own therapeutic experience as a sort of death and rebirth.

8. *The reification of concepts.* Finally the organism may turn upon its own codification system in various ways. As soon as sufficient complexity is reached to permit of two or more levels of abstraction, the organism becomes able to treat abstractions of a higher level as though they were equivalent to abstractions of a lower level. In brief, the organism may reify any concept within the wide scope already indicated above, and may endow this concept with, for example, causal or controlling efficacy. "Morality" (an abstraction derived from the actions and the words of the self and others) may be seen as "binding" upon actions of the self; the organism may respect or may revolt against "cultural conventions"; he may even sneer at or repudiate "death" (an abstraction about which he can, in the nature of the case, have no subjective knowledge).

This very brief survey of the orders of complexity and possible diversity in codification will serve to prepare the reader for the generalization that it is possible for the organism to commit many types of error in its codification and interpretation of the world. The next section will attempt to define some of these types of error.

INTERNAL CONTRADICTIONS OF CODIFICATION-EVALUATION

The previous section discusses the conceivable scope of variation in codification-evaluation but stops short of considering the discrepancies which may occur within such systems. The present section adds a new level of complexity by asserting that, conceivably, contradiction (i.e., ambivalence) may occur in any of the types of codification there outlined; that such internal

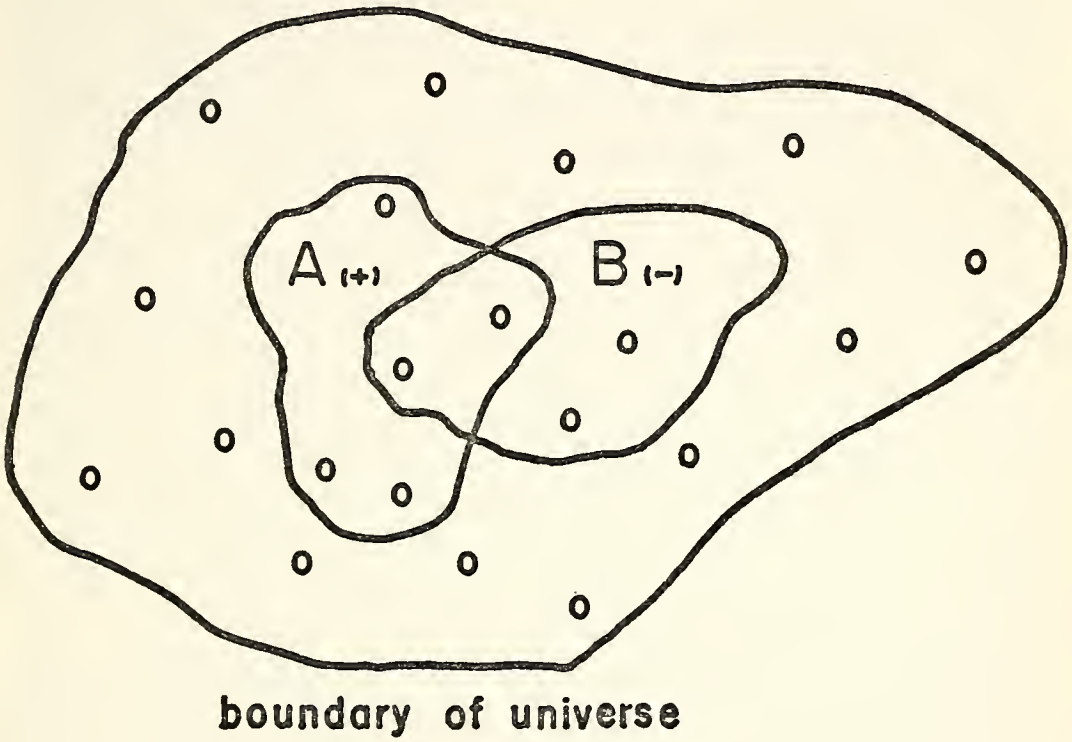
contradictions may occur at any level of abstraction; and that a given contradiction may in fact involve two or more such levels.

In daily life and psychiatric experience, it is common to observe that a person may see and evaluate similar events in one way in one set of circumstances and in quite a different way in another set of circumstances; and the contrast of circumstances which determines such a change may be either internal (for example, a shift of mood) or external (that which is approved and valued in war may be regarded with horror in time of peace). Trouble arises the moment the individual fails to make due allowance for the contexts of his evaluation and equates, for example, certain actions which are appropriate in war with certain similar actions in peacetime. He thus creates for himself a concept or Gestalt (e.g., "violence") which is charged with both positive and negative value.

Perhaps if human beings were capable of maintaining clarity about the contexts of perception and evaluation, they might avoid the complex internal and interpersonal conflicts which result from such contradictions. But this they cannot achieve. If it were possible never to confuse a given type of event (E^1) in one set of internal or external circumstances (C^1) with similar events (E^2) in other sets of circumstances (C^2, etc.), all would be well. But this is impossible, short of sacrificing the whole of Gestalt codification. The price which man pays for the economy which Gestalt codification permits is his proneness to ambivalence. After all, the great economy which this type of codification permits is due precisely to the fact that it permits us to identify E^1 with E^2 (e.g., to recognize a square as a square though it is presented in many different ways). Codification in terms of Gestalten permits us to summarize experience, and it is this summarization of experience that results in ambivalence.

Further, a second sort of internal discrepancy in the codification-evaluation system follows from the fact that every summary is an arbitrary condensation of the unsummarized data. Every Gestalt label is a man-made categorization of events in a universe which might be categorized in infinitely various ways. Even in the instant of perception or action, the individual is applying many such labels to the given set of events or objects. Inevitably

there will be cases in which such overlapping labels will have contrary value or contrary implications for action. With such a variety of possibilities for internal discrepancy, it is perhaps hopeless to attempt any complete survey of the possibilities for



ambivalence. However, since certain types can be defined with some rigor, these will be listed:

(a) Cases of overlapping Gestalt labeling in which the Gestalten are of the same level of abstraction. These cases may be described in a diagram. Figure 3 represents a universe of objects and events as perceived by an individual, including the individual's own actions among these. Within this universe, he perceives a sub-set of items as together making a Gestalt unity, A; he also perceives another sub-set which make up another Gestalt unity, B. Now, if there are items in common between A and B, and if A is positively valued and B negatively valued, the result will be a form of ambivalence. The items in the area of overlap will be positively valued when they are perceived as

parts of A, but negatively valued when they are perceived as parts of B.

In this type of contradiction, it is important to note that there is no necessary tendency for the perception of the Gestalt A to promote perception of the Gestalt B and vice versa; rather, we would expect the perception of the one to hinder the perception of the other. It is, however, easy to imagine instances in which the perception of either Gestalt might drive the individual to the perception of the other. These would be cases intermediate between the first types of contradiction and the second, which we now describe.

(b) Cases which are comparable in form to the famous Russellian paradox (177). The paradox may be presented as follows: A man classifies entities into classes, and every class which he defines establishes a class of other entities which are non-members. He notes that the class of elephants is itself not an elephant but that the class of not-elephants is itself a not-elephant. He generalizes that some classes are members of themselves while others are not. Thereby, he establishes two larger classes of classes. He then must decide: is the class of classes which are not members of themselves a member of itself?

If the answer to this question is "yes," then it follows that this class must be one of those which are not members of themselves, since all the members are of this type—and therefore the answer must really be "no." If on the other hand the answer is "no," then the class must be a member of that other class whose characteristic it is that its members are members of themselves —so the answer must be "yes"; and so on. If the answer is "yes" then it must be "no"—but if it is "no" then it must be "yes."

Another paradox with essentially the same structure is that presented by a man who says, "I am lying." Is he telling the truth?

A mechanical model of such an oscillating or paradoxical system may be of use to the reader. Such a model is the ordinary electric buzzer or house bell. This machine consists of an electromagnet acting upon an armature (a light metal spring) through which the current which activates the magnet must pass. The armature is so placed that the circuit is broken whenever the magnet is active and causes the spring to bend. But the current

is re-established by the relaxation of the spring when the magnet ceases to act. We may translate this system into logical propositions by labeling that position of the spring which closes the circuit as "yes"; and labeling the contrasting position which breaks the circuit as "no." The following pair of propositions can then be stated:

1. If the spring is in "yes," the circuit is closed and the electromagnet operates; therefore the spring must go to "no."
2. But if the spring is in "no," the magnet is not operating, and the spring must therefore go to "yes."

Thus the implications of "yes" involve "no"; and the implications of "no" involve "yes." The model illustrates precisely the Russellian paradox, inasmuch as the "yes" and "no" are each of them being applied at two levels of abstraction. In Proposition 1, "yes" refers to position, while "no" refers to direction of change; in proposition 2, "no" refers to position, while "yes" refers to direction of change. The "no" to which "yes" is an answer is therefore not the same as the "no" which is an answer to "yes."

Similarly the paradox presented by the statement "I am lying" can be traced to a confusion of levels of abstraction. The three words which are all we have to go on are simultaneously both a statement (level 1) and a statement (level 2) about the falsity of this first statement; and the second statement is of a higher order of abstraction than the first. In Russell's formal presentation of the paradox in terms of "classes of classes" the levels of abstraction are made explicit and the paradox is ruled out of order.

This matter of the paradoxes is here discussed at some length because it is impossible to go far in thinking about communication and codification without running into tangles of this type and because similar tangles of levels of abstraction are common in the premises of human culture (Chapter 8) and in psychiatric patients. In fact, this is the type of internal contradiction which Korzybski (90) and the school of general semantics attempt to correct in their therapy. Their treatment consists in training the patient not to confuse his levels of abstraction. In fact, their treatment follows the lines of Russell's resolution of the paradox, which he attempted by asserting the rule that no class shall ever

be regarded as a member of itself—because to do so would be to confuse levels of abstraction.

The reader, if he submits himself to the experience of thinking through the Russellian paradox, will observe that a time element is involved. For a moment it is satisfactory to accept the answer "yes," but he will observe that as he perceives the more intimate details of the Gestalt set up by this answer, he is driven to reject it. Then for a moment the answer "no" is acceptable until its implications are perceived—and so on. From the psychological point of view, this time characteristic is important: the phenomenon is not one of static indecision, but one of "oscillation" in time. Probably everybody has had the experience of similar sequences in real life in which increasing familiarity with a Gestalt leads to its rejection in favor of some other which in turn later becomes unacceptable. These, in fact, are systems of contradiction in which temporary acceptance of one pole promotes preference for the other and vice versa. The attempt to resolve the conflict in favor of one of its polarities *ipso facto* generates a preference for the contrasting pole. This mechanism is therefore very different from that discussed in (a) above, though the two mechanisms may conceivably work in combination.

(c) A third form of apparent internal contradiction in the codification-evaluation system is that of circular preference (106). When possibilities are offered in pairs, it may happen that A is preferred to B; and B is preferred to C; and C is preferred to A. In such a system, there will presumably be impossibility of decision when the three possibilities are simultaneously present. The data upon preference systems of this kind are meager, but the phenomenon is of great theoretical importance. The phenomenon is said to occur in experiments with aesthetic preference—e.g., when rectangles are presented in pairs, and the subject is asked to express preference for one member of each pair.

The mechanisms involved in circular preference may be various: (a) It may be that in the Gestalt "A plus B" the evaluation of each member is a function of the presence of the other; so that when B is presented in the presence of C, and the Gestalt "B plus C" is seen, B will be evaluated differently and perhaps

according to different criteria. (b) The mechanism of decision may consist of multiple linked sub-entities each with its own preference. The total mechanism of decision would then be something like a voting population; and notoriously in such systems, if there are three voters (or three equal parties) it is possible for that voter who sends no candidate to the election to swing the decision to that one of the other two candidates whom he prefers. (c) McCulloch has outlined a possible type of neurological circuit which will produce a similar result, and a related phenomenon obtains among the alternative solutions to certain types of Von Neumannian games (168).

Whether there are other types of internal contradiction or whether all are ultimately reducible to the three types mentioned above, is unknown.

ONE-WAY COMMUNICATION: THE UNOBSERVED OBSERVER

Before considering the more complex case of two or more organisms in mutual communication, it is worth while to ask how an observer, of whom the single observed organism is unaware, might build up inferences about that organism's system of codification-evaluation. This case, of involuntary one-way communication, will provide an important generalization relevant to more complex cases.

If, for example, the observer sees the organism moving in a straight line toward some target such as a source of light, he will not with this limited observation be able even to recognize a tropism. Repeated observations of the same type will still not help him, except to verify a very broad hypothesis that the coincidence between the direction of motion of the organism and the direction in which the light is located is due to something more than chance. He will not know that the direction of motion is selected by any process within the organism. To learn more, either the observer must perform repeated experiments, or he must observe repeatedly that the organism corrects itself every time the course deviates from the direction of the target. Moreover, the observer's experiments will necessarily take the form of placing the organism in error (e.g., he will place the light somewhere not in the direction in which the organism is travel-

ing and will then look to see what the organism will do). And from this it follows that the data which the observer obtains by experiment are of the same general type as those which he would obtain by observing the organism's self-correction in a variety of circumstances.

From these arguments, the major conclusion follows: That the correction of mistakes is a basic means of communication and is actually the only sort of communication which will permit an unobserved observer to form inferences about the codification-evaluation system of the observed.

The special case in which an organism while performing some action talks to itself and is overheard by an observer is interesting but does not invalidate the above argument: (a) If the observer is unfamiliar with the language of the observed, his only way of reaching an understanding will be by regarding the verbal behavior and the action sequences of the observed as a single system of ongoing self-correction. In this way he will finally discover what the words mean, and each utterance will be meaningful as selected by the observed in a process of self-correction. The observer will learn, in fact, the polarities of codification in the language; he will be able to assign to each word positive meaning, in so far as the word delimits a set of negated alternatives. (b) If the observer already partly or fully knows the language of the observed, he will tend to operate by identifying himself with the latter, ascribing to the observed his own understanding of the words used. In doing this he will either learn nothing new about the latter's codification system, or he will discover that he himself is making mistakes in interpreting what is said; as regards these parts of the codification-evaluation system of the observed, he will be, in fact, in the position of not knowing the language. In any case, unless the verbal stream is accompanied by action and is self-corrective in regard to the action, the observer will be able to learn nothing about the language.

This generalization echoes what has been said earlier as to the nature of codification-evaluation. It was there stated that in both codification and evaluation the universe is structured into a network. In the case of codification, the network is one whose nodes are the bi-polar or multi-polar discriminations of perception. In

the case of evaluation, the network has nodes which define the polarities of preference. In studying the mistakes and self-corrections of the organism, the observer is in fact getting the data necessary to map out the polarities of these networks. He will learn—albeit laboriously—what discriminations the organism can make, upon what cues it acts, whether it is able to perceive any characteristics of its own actions, how the action system is related to cues given, and so on.

INTRAPERSONAL COMMUNICATION: THE SELF-OBSERVER

A question of very great practical and theoretical importance in the psychiatric world concerns the limitations of self-observation and self-therapy. A corresponding question strikes anthropologists, who are well aware that it is especially difficult for a student to obtain insight into his own culture. It is orthodox among anthropologists today to believe that deep and articulate insight into one's own culture can only be obtained through the comparative method. Be that as it may, it is certain that awareness of his own cultural premises has come to man late in his history and has been aided by comparison between cultures. It is natural to draw an analogy between anthropology and psychiatry in this matter, and to suggest that the need for a comparative approach in anthropology is comparable to the need, in therapy, for another human being (the therapist) different from the self, against whom as a background the peculiarities of the self can be seen. The analogy cannot, however, be pushed too far. It is curious that while the especially important item in an anthropologist's training is his first-hand experience of a completely foreign culture, the corresponding item in the training of a psychiatrist is his own psychoanalysis.

Opinions differ as to what an individual can do in the way of obtaining insight into his own personality unaided by any therapist, and the problem is complicated by the evident fact that therapeutic progress may, under certain circumstances, occur without insight. It is possible that a therapist is necessary if the patient is to achieve insight but that other types of progress can occur without his presence.

In the present study the problems of self-observation are

relevant as part of the base from which we go on to inquire into interpersonal communication. Briefly, the question is: What are the limitations of self-observation as a process by which an individual may obtain new understanding or information about his own codification-evaluation system?

The problem has many branches: (a) It would be desirable to describe the phenomena of self-observation in some way which will not personify a self within the self. (b) The question calls for a definition of what is meant by "new" understanding as distinct from a working out of pre-existing contradictions within the individual's codification-evaluation system. (c) A formal examination is necessary of the actual limits upon self-discovery which arise from the fact that an individual can—of necessity—perceive his own life and actions only in terms of his own system of codification-evaluation. He is, in all cases, unable to perceive the characteristics of the system in terms of which he perceives.

Of these problems it is the third that is of special interest to the present study. The epistemological problem of consciousness and the nature of the self within the self we propose to postpone as being for the present beyond the reach of scientific examination. It has, indeed, been suggested that the subjective experience of consciousness is determined by internal conflict or contradiction. Such a hypothesis would partly remove the problem of consciousness from the epistemological field and place it in the area of the second question proposed above—that of the working out of pre-existing internal contradictions. This second question, in turn, may be postponed in favor of the third, which has the heuristic advantage of greater simplicity. If we can set limits to the possibility of self-perception in an organism not made complex by internal contradiction, these limits will be relevant for any consideration of the more complex cases.

We now consider the possibilities of self-perception in the unconflicted organism. Again, for heuristic reasons, we consider first the case of an organism in an environment such that the premises of codification-evaluation within the organism are true and sufficient for that environment in which it lives. Such a

hypothetical organism will always achieve its goals by means of those sorts of codification and self-correction which are characteristic for the organism, and it will not strive after any impossible goals. The question is whether, from such a sequence of automatic successes, the organism can ever achieve a new insight into its own automatic processes of self-correction.

The answer is surely a flat negative. From all that is known about learning, it would follow that in such a hypothetical case, not only no new insight but even no learning of any kind will occur. Indeed, the hypothetical case here discussed is precisely that of the abstract "player" in a Von Neumannian game.

If, on the other hand, there are contradictions, not within the organism but between the premises of the organism and those which obtain in the environment, then the position is entirely different. We know, as a fact of behavioristic experiment, that in such cases an organism which formerly appeared to act in terms of one premise system may after a period of trial and error gradually or suddenly begin to act as though in terms of another and better-adapted system. Further, we know from introspective reports that such learning may be accompanied by a change in the organism's conscious perception of the environment.

This again is a problem of mistakes and the correction of mistakes, such as was discussed in the previous section. The organism has been "put in the wrong" by the environment, and now the question is what orders of new information the organism can achieve as a result of going through the whole experience of frustration and self-correction, and achieving that new system of codification and evaluation by which the frustration is reduced. Having been put in the wrong, the organism corrects itself, not merely modifying its action but modifying—more or less profoundly—the basic processes and mechanisms by which actions are related to environmental cues. The "mistake" corrected in this sequence is of a very different order from the act of self-correction which characterized the organism at the outset of the experiment. The organism has now modified its system of self-correction.

The following considerations are relevant to the comparison

between the therapeutic process and such modification of system as can be achieved by the isolated organism:

1. The change in the isolated organism, in so far as it is an improvement in adaptation, may be regarded as "therapeutic."

2. It is known from experiment that failure on the part of the organism to adjust its premises to the conditions of the environment may be anti-therapeutic and may lead to experimental neurosis. That is, the organism may be affected by the fact of failure—may, in this sense, have information about failure.

3. It is unlikely that the old premises will be totally obliterated in the course of the change. Rather, they are likely to survive in modified or "repressed" form. It is possible, in fact, that the organism at some stage of the process of learning—and perhaps forever afterwards—will entertain conflicting premises with all the complications which this may entail.

4. It is possible that the organism may achieve new insight into the environment but very doubtful that it will obtain new insight into the self. There may be a deutero-learning—i.e., a learning to learn, or learning about learning (cf. Chapter 8)—such that the organism when again put in the wrong will be, for example, less anxious because of an acquired faith in its own ability to deal with such misfortune. But this is only doubtfully an increased insight into the self. Indeed, it is very doubtful if such an increase can possibly occur as a result of the changes which are here considered.

Before the experiment, the organism perceived its own actions in a certain way determined by the then existing premises; after the experiment it will perceive itself and its actions in terms of the new premises. But this is not a change in the order of self-perception such that we would call it insight. The next step—to see the self as an entity which has characteristically accomplished this change—need not occur. A change in the premises of codification-evaluation need not denote any greater insight into these premises unless the individual can see this change as a contrast, comparing himself with what he formerly was. In such a comparison he is essentially operating as two

persons between whom a contrast can be stated and a comparative method applied leading to greater insight. He is doing something comparable to what normally occurs in a two-person system.

It appears, then, that a two-person system of some sort will always be necessary for insight therapy, but perhaps not for other types of learning. We must, however, also expect that these other types of learning, often themselves therapeutic, will occur in the two-person situation even though the presence of the second person may not be necessary.

The background is now sufficient for the consideration of two-person systems.

COMMUNICATION BETWEEN TWO PERSONS AND METACOMMUNICATION

The next step is to extend what has been said in the previous sections to the phenomena of relationship between two or more persons—anthropomorphic organisms. The primary problem of such communication has been aptly put by Janet Baker (age 10) as follows: "When people thought of a language, how did they think of it if there were no words to think with? After they had thought of it, how did they get other people to understand it? If they went from door to door explaining it, people would think that they had gone crazy because they wouldn't know what the words meant. After the first language was started how were others formed? These are the questions that make me say, 'I wonder how people learned to talk' " (10).

This statement of the matter presents an ultimate problem to which the present inquiry leads; but it must be remembered that between whole human beings, after infancy, there can never be a total lack of mutual understanding. To be sure, there may be misunderstandings, and these may be so profound and dramatic as to seem total, but actually for even misunderstanding to occur, there must be some shared premises of codification-evaluation. Each person must at least have some notions about himself and the other; he must, for example, think of both as alike in being alive and capable of emitting and receiving communication. Indeed, if misunderstanding leads to hostility,

it is immediately clear that there must exist common premises regarding anger and pain. The beginnings of a common codification system are latent in our biological nature, our common anatomy and common experience of bodily functioning and maturation. When two human beings meet, they inevitably share many premises about such matters as limbs, sense organs, hunger, and pain.

As regards external cues, there is considerable evidence that among birds, amphibia, fishes, and invertebrates, the members of a species may share an innate tendency to respond in a complex way to a particular sharply defined cue or sequence of cues —a smell, a shape, a size, a patch of color, and the like—originating in other individuals. Such mutual responsiveness may take on the appearance of ongoing interaction. For example, among the sticklebacks there is such an exchange of behaviors between the sexes leading to reproduction. Each sex has a series of specific differentiated responses, and these are interchanged, each partner's response being a stimulus for a new response on the part of the other until finally the male inseminates and stays with the eggs laid by the female in the nest which he has built (164).

In the case of mammals, and especially man, it seems that such innate tendencies to respond in a complex differentiated manner to highly specific external cues are either poorly developed or are changed and blurred by later learning. Man's instinctual equipment is overlaid by cultural elaborations, but there still remain, common to the species, a number of tendencies to respond in a gross or diffuse manner to certain gross or diffuse stimuli, such as loud sounds, removal of support, heat or cold, pain, and the like.

Further, all human beings, as we know the species today, share the notion that language and gesture are media of communication [2] even though every culture has its specific variants

[2] The late Doctor Stutterheim, Government Archeologist in Java, used to tell the following story: Somewhat before the advent of the white man, there was a storm on the Javanese coast in the neighborhood of one of the capitals. After the storm the people went down to the beach and found, washed up by the waves and almost dead, a large white monkey of unknown species. The religious experts explained that this monkey had been a member of the court of Beroena,

of these media (184). Even within the culture, the poet may have extraordinarily different premises about the use of language from those held by the advertising man. A dancer may have one set of notions about communicative uses of posture, while the catatonic has another; and yet both share the notion that posture is communicative and, at some more abstract level, both systems of communication probably meet in many common premises about the body. If the contrasting persons live in the same culture, they will also share some vague—even distorted—recognition of the points in which they differ.

In this study, we are specifically concerned with communications between persons who have a large measure of common vocabulary and common setting in the American cultural scene, persons who have lived a great part of their lives in the American variant of Occidental culture. And yet, as between patient and therapist there may be a deep gulf of difference in their premises regarding such matters as were discussed earlier. They may have sharply differing notions as to the boundaries of the self, and each may see his own relationship to other human beings in his own idiosyncratic terms. The paranoiac may believe that the environment is all-powerful and set upon his undoing, but it is impossible to predict what phrasing the therapist may entertain as to his own relationship to his environment. Some therapists are willing and others unwilling to see themselves as shaping their human vis-à-vis. The thesis of this book is that only by communication can therapy occur, and communication will depend upon those premises which the two persons have in common and upon the complexities of the two-person system.

Certain special characteristics emerge in the interpersonal

---

the God of the Sea, and that for some offense the monkey had been cast out by the god whose anger was expressed in the storm. The Rajah gave orders that the white monkey from the sea should be kept alive, chained to a certain stone. This was done. Doctor Stutterheim told me that he had seen the stone and that, roughly scratched on it in Latin, Dutch, and English were the name of a man and a statement of his shipwreck. Apparently this trilingual sailor never established verbal communication with his captors. He was surely unaware of the premises in their minds which labeled him as a white monkey and therefore not a potential recipient of verbal messages: it probably never occurred to him that they could doubt his humanity. He may have doubted theirs.

system which were not significantly present in the hypothetical system containing only one organism:

First, each organism is in receipt of cues having a different order of complexity from those emitted by inanimate objects. Indeed, the messages externally exchanged between organisms must be compared with the intra-organismic processes of codification and evaluation rather than with the data which the organism collects from the inanimate environment. The extraordinary complexity of intra-organismic codification was discussed above, and it was then noted that this complexity is, so far as we know, achieved by very simple neural signals traveling in exceedingly complex pathways, in a network with many billions of synaptic nodes. With the aid of this neural network and possibly other parts of the body, the organism achieves the complex units of internal communication which we call Gestalten. The significant fact, for our present purposes, is that in interpersonal communication, the units and aggregate messages reach this same level because words and postures already refer to complex Gestalten corresponding to some of those which the internal system uses. Communication between persons is of course pathetically impoverished compared with the richness of the intrapersonal consciousness, which in its turn is but an impoverished and restricted version of the total psychic life of the person. But still it is important that the external communications are a codification of the internal psychic life and that the recipient of such communication is receiving an already elaborated product from the psychic life of another individual. In this, interpersonal communication differs profoundly from all perception of the inanimate environment. The perceiving individual must synthesize his data about the inanimate environment into appropriate units and has a certain freedom to do this in an idiosyncratic manner, whereas in receiving a verbal or other personal communication he has less freedom because the matter of the message is already synthesized into Gestalten (words and phrases) by the communicator. Even the recipient's understanding of the message is conditional upon his having become habituated to the narrowly confined conventions of codification which culture imposes.

Each individual receives, of course, sense data of the ordinary kind in regard to the other; each sees and hears the other as a physical entity. But in addition each receives verbal and other symbolic matter from the other, and each has therefore the opportunity to combine these two types of data into a single more complex stream, enriching the verbal flow with simultaneous observations of bodily movement and the like. It was suggested above that in intrapersonal processes the body might serve an analogic function supplementary to the more digital processes of neural thought. We now note that the bodily processes of the other person—his postures, tension, flushing, and the like—serve a corresponding function in interpersonal communication. Each person is able to get a multidimensional view of his vis-à-vis, enriching the stream of merely verbal symbols with a recognition of bodily processes in the other, and these are more or less intelligible because of common biological background and cultural conditioning.

In illustration, it is worth mentioning a curious detail in which the strictly Freudian analytic session differs from the majority of two-person systems. When the patient is on the couch and the analyst sits on a chair beyond the patient's head, the analyst gets a fair but perhaps sufficient view of the postures and facial expressions of the patient; but the latter is cut off from seeing his therapist. The asymmetries which this arrangement introduces into the therapeutic situation are undoubtedly very complex and surely vary from therapist to therapist and from patient to patient. From the point of view of the present discussion, it is significant that the patient receives only verbal messages from the analyst and so has maximum freedom to build up a fantasy picture of the affective aspects of the analyst's personality. This picture may be later examined when the transference is analyzed. At first the patient, according to his lifetime habit, attempts to make inferences about the analyst in order to tailor his words to fit that person. Later he discovers, perhaps, that in the therapeutic session such tailoring is difficult and he is then thrown back upon speaking and acting as "himself" with minimal aid from such introjected images.

A further characteristic which emerges in the interpersonal

system but which was almost negligible in the simple relation between organism and environment is the real existence of the group as a determinant of the actions and communications of the separate persons. The relation between organism and environment is already an interaction, and in such dynamic systems as a man driving an automobile or a man walking or dancing, the reality of the interactive whole as a determinant of the functions of the constituent parts is clearly recognizable; but when we deal with two-person systems a new sort of integration occurs. The condition for the existence of a determinative group in this sense seems to be that each participant be aware of the perceptions of the other. If I know that the other person perceives me and he knows that I perceive him, this mutual awareness becomes a part determinant of all our action and interaction. The moment such awareness is established, he and I constitute a determinative group, and the characteristics of ongoing process in this larger entity control both individuals in some degree. Here again the shared cultural premises will become effective.

About the evolution of the "group" in this sense, there is little information, but the question of such evolutionary history is worth considering, if only to stress that the group, as defined in terms of mutual awareness of perception, is something different from groups determined merely by mutual irritability or responsiveness. In the case of the sticklebacks (164) mentioned above, there is a complex mutual responsiveness but no evidence which would indicate that either individual is aware of the other's perception. Similarly in the elaborate communication which von Frisch has demonstrated among the bees, there is no reason to believe that such awareness occurs. Probably this evolutionary step occurred for the first time among mammals, and perhaps the phenomenon occurs only among primates and among animals intimately domesticated by man. The matter needs critical investigation.

Operationally, to determine whether a group is of this higher order, it would be necessary at least to observe whether each participant modifies his emission of signals in a self-corrective manner according to his knowledge of whether the signals are

likely to be audible, visible, or intelligible to the other participants. Among animals, such self-correction is certainly unusual. Among men it is desirable but not always present.

It would also be important to identify among animals any signals of the following types: (a) signals whose only meaning would be the acknowledgment of a signal emitted by another; (b) signals asking for a signal to be repeated; (c) signals indicating failure to receive a signal; (d) signals which punctuate the stream of signals; and so on. With complete awareness of the other's perception an individual should stop repeating a signal after it has been received and acknowledged by the other individual, and this type of self-correction would indicate mutual perceptive awareness. Correspondingly the lack of such adaptation—often observable in people—would denote imperfect awareness of the other's perception except in those cases where some change in meaning or intensity is conveyed by the repetition of the message. Lastly, the motivation for deliberate falsehood can hardly exist without awareness of the other individual's perception, nor is the falsehood likely to be successful. Thus the occurrence of falsehood becomes an evidence that the group is one based upon mutual awareness of perception.[3]

All these criteria for the existence of mutual awareness build together to give a picture of the entirely new order of communication which emerges with this awareness. For this new order of communication, the term "metacommunication" is here introduced and defined as "communication about communication." We shall describe as "metacommunication" all exchanged cues and propositions about (a) codification and (b) relationship between the communicators. We shall assume that a majority of propositions about codification are also implicit or explicit propositions about relationship and vice versa, so that no sharp line can be drawn between these two sorts of metacommunication. Moreover, we shall expect to find that the qualities and characteristics of metacommunication between persons will de-

[3] The falsehoods implicit in animal mimicry, protective coloration, and the like, present a special problem. Here, according to orthodox hypotheses, the self-corrective system is not the individual animal but the larger system of the total ecology within which natural selection operates correctively upon the population.

pend upon the qualities and degree of their mutual awareness of each other's perception.

If we can recognize the existence of such awareness by observing the individual's self-correction of the signals which he emits (and all the criteria are really only special cases of this self-correction), it follows that a variety of characteristics attributed to the other individual have become relevant in shaping and motivating the behavior of the signaler. The signals are being tailored to fit the signaler's ideas about the receiver. From this point onward the evolution of a number of human habits and characteristics—introjection, identification, projection, and empathy—understandably follows. It even becomes possible for one human individual to coerce another in terms of a correct or incorrect understanding of that other's view of the universe.

This discussion of the importance of interpersonal inference introduces a series of other variables significant for the two-person system which did not appear in hypothetical systems involving only one person. When the system consists of two persons, it is possible for these persons to be either similar or dissimilar in their codification characteristics. They may be alike in the way they perceive the universe and act upon their perception; or they may be different in these respects. The new variable which we note is then the statement of similarity or dissimilarity between the two persons.

Another and different variable which emerges only when two persons communicate will state whether or not the premises of the two persons conflict. It is evidently possible that, though two persons are much alike, the very points in which they resemble each other may be a cause of mutual conflict. If, for example, they are alike in their expansionist goals, these goals may well coincide, and rivalry or jealousy may develop. Indeed, as is well known to educators, the establishment of a competitive relationship between persons is one of the most effective methods of training the participants to a similarity or conformity in their perception and evaluation of the common universe in which they live. More formally stated, these are cases in which A's phrasing of the relationship between A's self and part of the environment is superficially the same as B's phrasing of the re-

lationship between the same part of the environment and B's self. Both may say, "It's mine." Such phrasings are really discrepant, because the two selves do not coincide.

Conversely, when the two persons have evidently different phrasings of the universe, there need not necessarily be conflict. It is possible for the phrasings to be complementary, so that a "fit" occurs (149), and the two individuals may be able to cooperate in an asymmetrical relationship. This occurs, for example, in successful relations between persons of opposite sex. And notably, in such cases it is even not necessary for the persons to understand each other's universe, though it may be important that they recognize the fact of difference. Beyond this recognition, the efforts to understand may result in failure to communicate. These, however, are questions which can hardly be considered apart from the role of the cultural matrix, which will be examined in the next chapter.

8 · CONVENTIONS OF COMMUNICATION: Where Validity Depends upon Belief

By Gregory Bateson

In the preceding chapter a theory of communication was built up starting from irritability and adaptive action at the very simplest level, and advancing through the phenomena of codification up to the phenomena of mutual awareness of perception. With this last element the theory begins to describe human relations.

In the present chapter what we have to do is to push on into more human matters. Instead of talking about sticklebacks and abstract entities we begin to talk about beings which, schematically at least, resemble people. The particular step towards humanity which we take in this chapter is to examine the idea that *man lives by those propositions whose validity is a function of his belief in them.*

Two sorts of propositions of this kind were mentioned in the preceding chapter. First, the propositions about codification. Such a statement as "The word 'cat' stands for a certain small mammal" is not either true or false. Its truth depends upon agreement between the speakers that it be true. In terms of such agreement they understand each other; or where disagreement occurs they will meet with misunderstanding. And this statement about the word "cat" is only one of a vast category of statements about codification, which category ranges all the way

from the conventions of local phonetics up through the conventions of vocabulary to the conventions of syntax; and the same category will include the conventions of timing, pitch, emphasis, tone of voice, and all the other modalities of verbal and nonverbal communication, since all communication involves codification and these are the conventions of codification.

In addition, the preceding chapter contained statements about metacommunication; and this category of statements is a larger genus within which the statements about codification are to be included as a subcategory. When A communicates with B, the mere act of communicating can carry the implicit statement "we are communicating." In fact, this may be the most important message that is sent and received. The wisecracks of American adolescents and the smoother but no less stylized conversations of adults are only occasionally concerned with the giving and receiving of objective information; mostly, the conversations of leisure hours exist because people need to know that they are in touch with one another. They may ask questions which superficially seem to be about matters of impersonal fact —"Will it rain?" "What is in today's war news?"—but the speaker's interest is focused on the fact of communication with another human being. With comparative strangers, we "make conversation" rather than accept the message which would be implicit in silence—the message, "We are *not* communicating." It seems that this message would provoke anxiety because it implies rejection; perhaps also because the message itself is explosive with paradox. If two persons exchange this message, are they communicating?

Many sorts of games are of interest in this connection. An implicit message which is exchanged at bridge tables and on tennis courts is the affirmed agreement between the players as to the rules and goals. By participating in the game, they affirm the fact of communication, and by competing, they affirm the fact of shared value premises.

Similarly, every courtesy term between persons, every inflection of voice denoting respect or contempt, condescension or dependency, is a statement about the relationship between the two persons. Such messages are carried on the stream of verbal

communication, and all these messages and their codification determine such matters as role and status, whose truth and stability depend upon implicit or explicit agreement between the persons that the relationship is as indicated. Moreover, all cues which define status and role are metacommunicative, since the recipient of any message is guided in his interpretation of that message and in his resulting action by his view of the relative roles and status between himself and the speaker.

It appears, then, that within the larger genus of metacommunicative propositions, it is possible to recognize at least two subcategories—the propositions about codification and the propositions about interpersonal relationship. It is certain, however, that overlapping frequently occurs between these subcategories and that a very small shift in emphasis or interpretation will cause a given proposition to appear to shift from one subcategory to the other. This shifting character is due to two circumstances: (a) that statements about relationship must still be codified; and (b), that every statement in a given codification is an implicit affirmation of this codification and is therefore in some degree metacommunicative. (When I say, "I see the cat," I am implicitly affirming the proposition that the word "cat" stands for that which I see.) The shifting relation between the propositions of codification and the propositions of human relationship can be illustrated by the following example: The statement "A policeman carries a nightstick as a badge of authority" contains both the statement of status and the statement of how this status is codified; the same example will serve to emphasize that all interpersonal actions are, in some degree, messages. When the policeman uses his nightstick, he is asserting his status in a particular relationship to a particular offender.

The purpose of the present chapter is to examine this whole matter of propositions and implicit premises whose validity depends upon belief.

First, it is necessary to survey, briefly, the occurrence of premises of this order in human life. Broadly, it will be argued that propositions and premises of this kind are scattered through the whole range of life. They are implicit in the phenomena of learning, they recur in the phenomena of character forma-

tion, and finally they determine the phenomena of human relationship and even religious faith.

The matter may best be approached by starting from the experiments on learning (74). In even very simple learning experiments, such as those on rote learning, Hull (79) has demonstrated that a phenomenon appears which is of a higher level of complexity than those which are ordinarily discussed by the psychological experimenters. It is found that an individual learning to recite nonsense syllables by rote not only learns to repeat the nonsense syllables of the given series but also becomes more skilled in learning nonsense syllables. When presented with another series of nonsense syllables, he will learn the second series more rapidly than he learned the first. Similarly, he will learn a third series more rapidly than he learned the second, and so on, up to an asymptotic limit of skill in learning nonsense syllables.

The term *"deutero-learning"* has been coined (18) to describe this higher order of learning, and this word can be regarded as a synonym for "learning to learn."

If we now consider the various sorts of learning experiments, we find that it is possible to classify the various sorts of experiments according to a formal scheme (74). There are the rote experiments, already mentioned; there are the Pavlovian experiments in which the actions of the experimental subject have no influence on the occurrence or timing of the reward or punishment; there are the instrumental reward experiments in which the subject by performing a certain act determines when the reward shall be given; there are instrumental avoidance experiments in which the subject by his own act prevents a punishing event from occurring; there are escape experiments, maze experiments, and so on. In brief, there is a series of types of time sequence and a series of different roles which can be assigned to the experimental subject; and the time sequences and the roles differ from one type of experiment to the next.

We now propose the following hypothesis for which experimental verification is not yet available: [1] if the human subject

[1] Since the above was written, the writer's attention has been called to Harlow's experiment on "learning sets" (72). Harlow's "learning set curves" are precisely "deutero-learning" curves.

shows the capacity for learning to learn in rote experiments, then it is likely that the phenomenon of learning to learn will occur much more widely and is, we suppose, present in all other types of learning experiments. For example, the experimental subject who has experienced a series of instrumental contexts is likely to show added skill in dealing with other instrumental contexts. In fact, there is likely to be a phenomenon of deutero-learning for each type of learning experiment: the experimental subject learning to deal with the particular type of sequential context of which he has had repeated experience.

If that is so, we may go on to ask: What sort of world will the subject with repeated experience in instrumental contexts inhabit? How will he perceive and interpret the world in which he lives? The subject will clearly expect the world to be made up of contexts appropriate for instrumental response; his threshold for the recognition of such contexts will be lowered. Similarly in regard to the Pavlovian subject we may now state that he will learn to expect a world in which he has no control over the good and evil which may befall him; he will try to know when they are coming, and he can take appropriate visceral precautions, readying his body for the food or pain. He can, so to speak, look for omens to tell him when the disaster will come, but it will not occur to him that he can do anything about the disaster, except within his own body. Similarly, the subject with repeated experience of instrumental avoidance will have a different orientation to the world from the subject with repeated experience of instrumental reward; the former looking for avoidance of punishment, the latter looking for positive gain. And so on.

Thus the discussion of the phenomena of learning moves forward from the type of question asked by psychological experimenters—"Under what circumstances will the subject learn to do such and such?"—to a higher-level question concerning the circumstances which will alter the "character structure" of the animal. The Pavlovian experimental subject becomes, as it were, a prototype for a certain species of fatalism. The subject of instrumental experiments becomes a prototype for—if you please—certain themes of American character structure; and so on.

We are, in fact, coining the beginnings of a set of formal categories for describing character structure, and these descriptions are derived not from what the subject has learned in the old simple sense of the word "learning," but from the *context* in which the simple learning occurred.

This is the level at which learning experiments become relevant to psychiatry, and the hypothesis of deutero-learning provides the bridge between simple psychology and psychiatric theory. The psychiatrist is not concerned with the question of whether the patient is able to write, to use a typewriter, to play the piano, to walk, or to do any other thing; but he is concerned with the description of the context in which the patient learned, for example, to typewrite or to control his sphincters. If the patient learned his lesson in a context of threatened punishment, that fact may throw light upon the character structure of the patient, not the mere fact of his having learned the appropriate actions.

Now, we ask, of what order is the conscious or unconscious proposition which guides the subject of instrumental experiments—the proposition which we may crudely verbalize for him as "the world is made up of contexts in which I can act instrumentally"? If we consider this statement, it is at once evident that the instrumental subject will, within certain limits, experience a world in which his propositions are apparently verified. Being an instrumental organism, he will meet the world instrumentally; he will seek out and respond to those contexts which are appropriately structured, and he will thereby reinforce his own belief that the world is an instrumental world. Correspondingly, a fatalist or Pavlovian subject, believing that he can do nothing to further his gain or avoid punishment, will act in the world in such a way that his premise about the nature of the world is demonstrated to be true. These propositions, in fact, about the world in which we live are not true or false in a simple objective sense; they are more true if we believe and act upon them, and more false if we disbelieve them. Their validity is a function of our belief.

The psychiatrist is very familiar with phenomena of this kind. The paranoid, by his action, creates about himself those rela-

tionships to human beings which will in fact reinforce his paranoid premises about the nature of human beings. If he distrusts every man and acts upon his distrust, he will find that people are remarkably untrustworthy. And the same considerations apply to a whole host of aberrational premises.

We note that the premises of which character structure is built are closely related to the contexts in which learning occurs,[2] and further that the premises of character structure are propositions of the general type which we are discussing in this chapter —namely, those whose validity depends upon the subject's belief in them.

We proceed now to discuss human relationships. In order to approach these formally while still maintaining connection with the psychological hypotheses derived from the learning experiments, it is convenient to think of the learning experiment as consisting not of an experimental subject in an inanimate environment, but as a two-person system in which the subject is face to face with another organism. We now proceed, therefore, having personified the experimental subject in the paragraphs above, to personify the experimenter. When it is bluntly stated that the experimenter is an organism, we perceive that he too has placed himself in a context of learning, more complex than that which the subject experiences. The Pavlovian subject's context is one in which he first perceives the conditioned stimulus (for example, a buzzer) and then waits a certain length of time, possibly salivating, and finally experiences the reinforcement (e.g., meat powder). If we state this whole series of events now from the point of view of the experimenter, we thereby define a complementary pattern: the experimenter first acts to give a signal (the buzzer), then remains inactive for a fixed period of time while the animal reacts in one way or another

[2] The precise relationship between learning to perform a given action in a certain context, which we may call *proto-learning*, and the more elaborate learning which we here call *deutero-learning* is still obscure. It is probable that all proto-learning is accompanied by at least some degree of deutero-learning, but the converse is not necessarily true. It is at least conceivable that deutero-learning may occur in entities incapable of proto-learning. In particular, Von Neumann (168) has demonstrated that certain standards and conventions of behavior must logically come into being among hypothetical competitive robots whose total rationality by hypothesis precludes all proto-learning from experience.

under his observation; and finally he administers the reinforcement, regardless of the animal's reactions. Looking at these two sides of the interaction together, we obtain paradigms for such phenomena as dominance and submission, dependency and succoring, and the like. Each of these formerly loose terms can now be sharply defined in terms of some deutero-learned premise, acquired in the learning contexts of human interaction; and these sharper definitions will discriminate a number of species of interaction which were previously confused: for example, the "dominance" of the Pavlovian experimenter is clearly different from the "dominance" of the instrumental reward experimenter.

To illustrate this it is convenient to consider the interaction between two persons, A and B, and to represent the actions of these persons by *a* and *b* respectively. With this symbolism it is possible to substitute for the loose statement "A is dependent upon B" a more precise statement, as follows: "From his past experience of interaction, A has a deutero-learned premise which leads him to expect that, in interaction with B, there will occur frequent sequences of the type:

$$a' \; b \; a''$$

where a' is a signal of weakness or need, *b* is B's helping or succoring response, and a'' is A's acceptance or acknowledgment of this help."

This effort at precise description may seem to the reader to be truistical, but from the study of such paradigms for dependence, succoring, dominance, submission, and the like, there emerges a curious set of paradoxes: the more sharply the paradigms are defined, the more evident it becomes that the persons concerned in the interaction actually have a curious freedom to impose their own interpretations upon the sequences of interaction. It is this freedom and its deterministic limitation by old deutero-learned premises that make it possible for the individual to perceive the sequences of interaction in his own idiosyncratic manner and so to find reinforcement for his own deutero-learned premises.

An example is necessary. The paradigm for the statement "A instrumentally dominates B" will be:

$$a' \; b \; a''$$

where *a'* is A's command, telling B what to do, and possibly defining the conditional reinforcement; *b* is B's obedient act; and *a''* is A's administration of the reinforcement. Now, the paradigm for A's dominance obviously resembles the paradigm given above for A's dependence, both having the form *a' b a''*. The question therefore arises whether "instrumental dominance" is really different from "dependence"; or whether the participants in a given interaction have a freedom of interpretation such that *a' b a''* might appear to some individuals as dominance while to others it would appear as dependence. The answer is that there are certainly many instances in which *a'* can be regarded either as a plea for help or as a command; similarly *b* can often be regarded either as a helping act or as an act of obedience; and *a''*, if it is a statement of acceptance, a "thank you," can be seen either as a condescending reward or as the appropriate response of the dependent. It is up to A and B each to weigh his own interpretation of the events, to determine whether A was dominant or dependent. Finally, it is important to note that A and B need not be in agreement in their perceptions at this level.

The case of Jeeves in the Wodehouse stories will provide a more concrete example. Jeeves is the elderly butler and Bertie Wooster is his scapegrace master. The question which concerns us here is whether Wooster is dependent upon Jeeves (as younger man upon older) or dominates Jeeves (as master to servant). Are Bertie's orders to the butler statements of weakness or commands? Bertie, from his side, is free to see himself as the master; but Jeeves has the freedom, from his side, to dignify his own position by seeing himself as succoring Bertie.

Thus the definition of a relationship depends not merely upon the skeleton of events which make up the interaction but also upon the way the individuals concerned see and interpret those events. This seeing or interpretation can be regarded as the application of a set of propositions about the world or the self whose validity depends upon the subject's belief in them. The individuals are partially free to interpret their world according to the premises of their respective character structure, and their freedom to do this is still further increased by the phenomena of selective awareness and by the fact that the perceiving indi-

vidual plays a part in creating the appropriate sequences of action by contributing his own actions to the sequence.

In the same category with the deutero-definitions of relationship and the premises of character formation go many of the premises of any given culture. In discussing the difference between England and America (Chapter 6) it was suggested that an important fundamental difference can be derived from the fact that the child in America exhibits achievement and self-sufficiency vis-à-vis his parents, who take a spectatorship role; while in England the child takes a preponderantly spectatorship role vis-à-vis his parents, who are models showing him how to act. This American premise, that spectatorship goes along with the other characteristics of parenthood such as succoring and dominance, is not either true or false; it is a convention of the relationship, shaping character structure and owing its only validity to the unconscious or habitual acquiescence of those who participate in the relationship.

Similarly the values of American culture which have been discussed at length (Chapter 4)—puritan morality, success, change, equality, and sociability—reflect premises of this general order. The value set upon morality, success, etc., is continually reinforced by the occurrence of actions and communications in which the more abstract propositions about values are implicit. These value propositions are thus metacommunicative, and their validity depends upon the occurrence of the more concrete actions and words which result from Americans' acceptance of the values. The mechanism by which such value propositions are propagated in a culture is circular (14), (19).

Among the premises of human relationship as culturally defined, we include the premises which define the family constellation and all the premises of role and status, class and caste, which define the processes of interaction. And, in addition to all these, we have to include the conventions of international and cross-cultural conduct (13)—even the tedious and hateful conventions leading up to and ending in international warfare. Not only the premises of smooth interpersonal relationship but also the premises of hostility are carried upon the stream of more objective communication and action; and what is true of persons applies

also to international relations where the gradual breakdown of a *modus vivendi* is slowly documented at a metacommunicative level. This breakdown leads ultimately to the bitter agreement upon the use of force. This agreement, however, has still the same degree of unreality or reality—the same degree of abstractness—that is characteristic of all those truths whose validity is a function of man's belief in them. If, of two nations, each comes to believe in the hostility of the other, that hostility is real to this extent and to the extent that each acts upon its belief. But it is unreal—and there is therefore always some hope for international affairs—in so far as the belief is conceivably reversible. "Tweedledum and Tweedledee *agreed* to have a battle."

This survey of propositions whose validity depends upon belief has now included within the general group of metacommunicative statements the following types: first, the propositions of codification; second, the propositions of character formation; and third, the propositions of human relationship in cultural systems. The survey will naturally continue by examining the premises of man's vast symbolic activity in the fields of play, art, and religion. It is, however, convenient to stop here to consider the bearing of current philosophical thinking upon what has been said so far.

In the preceding chapter, under the heading of contradictions in codification, it was stated that there is always a danger that an individual's lines of thought may get tangled and produce paradoxes of the general type implicit in the statement "I am lying" or in Russell's more formal problem of the "class of classes which are not members of themselves." We now face the peculiar difficulty that the discussion of metacommunication is sure to lead us into paradoxes of this very type. The construction of such paradoxes depends upon a given utterance being simultaneously a statement about itself. Taking as an example the paradox presented by the man who says, "I am lying," we are caught in paradox because he makes a statement, and he makes a statement about this statement, the second being of a different order of abstraction from the first. The paradox arises from the interplay of these two levels of abstraction.

In discussing metacommunicative propositions, we land our-

selves at once in this position because metacommunicative statements are of a different level of abstraction from the simple objective statements upon the stream of which they are carried.

A considerable amount of inquiry in the last twenty years has gone into the attempt to unravel these difficulties, which came to the fore in the twenties. It was then hoped (176) that the whole of mathematics and logic might be made self-contained and unified without recourse to "self-evident" propositions, and Russell and Whitehead labored in the *Principia Mathematica* (177) to establish such a unity between mathematics and logic. It was found, however, that any such attempt involved asking, "What is really meant by the 'self-evident' axioms on which any mathematical system rests?" and that the statements which would define the axioms and give them logical foundation must always be statements of a different order of abstraction from the axioms, as the latter are contained in the theorems which are built upon them. The statements explaining the axioms are in fact metacommunicative as compared with the axioms themselves, and the latter are metacommunicative as compared with the theorems. The status of the axioms therefore becomes ambiguous, since they are used at two levels of abstraction, one relatively metacommunicative and the other relatively "objective"; and the total system of statements thus becomes comparable to the electric buzzer (p. 194) which must oscillate between the "yes" and "no" positions.

Since the days of the *Principia Mathematica* the matter has become even more difficult and more directly relevant to the questions with which we are here dealing. Gödel (63) has now demonstrated with rigorous proof that no system of statements can be self-contained in the sense of explaining its own axioms and not self-contradictory; that always—as a result of the very nature of communication and metacommunication—contradictions of the Russellian type must creep in. This statement of Gödel's—and there is apparently at present no reason to doubt his proof (176)—means in fact that psychology and the study of human communication can never hope to build a self-contained and coherent system which will not be self-contradictory.

In brief, we have to face the fact that when we deal simulta-

neously with both objective communication and metacommunication, contradictions will arise within the very field of our own inquiry.

In practice, this means that we must accept and must expect to find in the great creative fields of human communication—play, art, religion, epistemology, and psychiatric theory—paradoxes of the general type contained in the statement "I am lying."

We are now in a position to examine the nature of play, art, and religion. We have been warned.

In play, the element of "I am lying" is clearly recognizable. The participants in a game set up as fictions the rules of that game, they set up as a fiction (and a fluctuating fiction at that) the convention that the players are opposed to each other or are to compete with each other, and they set up fictional devices of codification to determine how gain and loss are to be symbolized. As we say, "It's only a game."

In art, the matter is more obscure but becomes clear if we consider the difference between art and propaganda (41). The propagandist is concerned with persuading his audience that what he says has more than the truth of man-made conventions. He is concerned with persuading his audience that the propagandic message is an objective statement rather than a metacommunicative message. It is true, of course, that many propagandic forms, films, plays, and the like, have an outward appearance of being honest fiction, but always in the propagandic form the accent is upon the idea that this fiction is in some sense objective truth. The story is presented as "typical," and therefore the audience is urged to act as though the play were a statement of reality. The artist, on the other hand, in contrast to the propagandist, can say honestly, "This is my creation" or "This is how I react to some part of my world"; and in this statement are contained the potentialities for paradox that occur in the statement "I am lying." The truths which the artist expresses contain frankly and honestly the combination of the metacommunicative with the objective. This is perhaps the greatest formal distinction between art and propaganda.

Similarly, Ruskin's "true" and "false" grotesque illustrates the same point. In the case of the "true" grotesque, the artist is honestly presenting some creation of the human imagination, some image either traditional or created in his own mental life, neither true nor false, but human. In the "false" grotesque, the artist tries to persuade his audience at least for a moment that his creation is a reality, is true in an objective sense, and the falsity of the false grotesque consists precisely in this—that no creation of the human imagination has this order of truth. Its only truth is that of being truly a creation—the creation of an honest mind.

In the field of religion, the problem of sorting out the objective, the propagandic, and the artistic elements, and relating these to the general category of propositions whose validity is a function of our belief in them, is exceedingly complex. Indeed, conflicting opinions about the degree of objective truth or "symbolism" contained in religious statements have been a source of strife through the centuries. The Christian religionists have notoriously tended to overstress the position that their mythologies and even parables should be regarded as objective historic truths, while the antireligionists have tended to the equally stupid opposite extreme of denying even metacommunicative or relative truth to any religious document upon which they could throw objective doubt.

Every religion has its central mythological statements. In Christianity, for example, we have the statements defining the omnipotence of God and the relationship of Father and Son to humanity. We are not concerned here with evaluating these statements at the objective or historical level. It is necessary to state, however, that whatever the degree of objective truth in the statements, they carry implicit in their poetry a large number of assertions of the type which we here discuss. We do not ask whether there *is* a Father in heaven; we only state that the words, "Our Father which art in Heaven," in addition to their objective truth or untruth, carry implicit propositions about the brotherhood of man, and we point out that these implicit propositions belong to the category which concerns us here: that, in so far as men can believe and act upon their imputed

brotherhood, this premise will determine their mutual relations; and that in so far as they disbelieve and act upon their disbelief, the implicit contrary proposition becomes true.

The preceding paragraphs raise questions which cannot be answered as yet—especially the question regarding the limits of deutero-truth. We state that the validity of a deutero-proposition *is a function of* belief, and it is clear enough that in many instances there is a range of values for the variables "validity" and "belief" in which an increase of belief will be accompanied by increase of validity. But this is very far from saying that the relationship between these variables is linear or that total belief will be accompanied by total validity. Indeed, it is probable that total validity can only be reached in special instances—if ever. More usually (e.g., in the case of the brotherhood of man) we may expect the validity of the proposition to reach a maximum beyond which future increases of belief will result in frustrating experiences for the believers, some of whom will then doubt the validity of the proposition. Further complications will follow if there is a division of opinion in a population, and that particular species of conflict is likely to occur which Collingwood has described as "eristic"—i.e., conflict about some variable which, if let alone, would settle itself at a value intermediate between the values for which the two sides strive.

The occurrence of such deutero-propositions throughout the fabric of every religious system can only be mentioned. It must, however, be stated that truths of the sort which we discuss are implicit in all religious communication, whether it is mythology or ritual, and that these propositions include not only the ethical implications of religion for human action but also the theories which any particular religion uses to define the relationship between man and the universe. Religion (25) is the great storehouse of those deutero-learned propositions which are summarized in such words as "fatalism," "instrumentalism," "passivity," "acquiescence," "free will," "determinism," "responsibility," "guilt," acceptance of the universe or revolt against it, and so on.

In fact, religion, like science and philosophy and art, is one of the mass agencies which determine our epistemology—our

theories of the nature of the reality in which we live and our theories of the nature of our knowledge of this reality.

This brings us to the conclusion of this chapter. In the next chapter the epistemology implicit in a collection of psychiatric statements will be examined. In contrast, the previous chapter and the present discussion of deutero-propositions form, together, a statement of the epistemology of the authors. This is the definition of our point of view, from which we shall study the psychiatrist's statements. And this statement of the authors' position must be concluded upon a strange negative note. As was said above, it seems that all attempts to build a coherent body of statements at several levels of abstraction must always end in paradox and contradiction. It is evident that statements about the theory of knowing are exceedingly abstract and are members of the class of propositions whose validity depends in part upon belief. This would indicate that the actual processes of knowledge (like the processes of learning discussed above) are surely modified by the knower's theory of the nature of knowledge. If this is so, then there must be a limit beyond which epistemology cannot go—a limit at which our attempt to resolve the contradictions of experience and communication will break down.

At the time of writing, the last word which must be added to the description of the authors' epistemological position is an acknowledgment that we expect our own position, like all others, to be, in the end, either incomplete or self-contradictory.