R E P O R T   R E S U M E S

ED 014 916   56   EM 006 022

LEARNING--FROM R-M THEORY TO EDUCATIONAL PLANNING.
BY- CAMPBELL, VINCENT N.
AMERICAN INST. FOR RESEARCH IN BEHAVIORAL SCIENCES
REPORT NUMBER AIR-D10-11-63-TR-B   PUB DATE NOV 63
GRANT OEG-7-48-0000-18
EDRS PRICE MF-$0.25 HC-$1.80   43P.

DESCRIPTORS- *MODELS, *LEARNING THEORIES, *LEARNING MOTIVATION, *COGNITIVE PROCESSES, *EDUCATIONAL STRATEGIES

THIS THEORETICAL PAPER COMBINES FAMILIAR PSYCHOLOGICAL VARIABLES IN A NEW SYSTEM DESIGNED MAINLY FOR PARSIMONY. PRINCIPAL CONSTRUCTS ARE REPRESENTATION, (R) A UNIT OF COGNITIVE ACTIVITY, AND M-VALUE, (M) A MOTIVATIONAL OR HEDONIC DIMENSION, OR THE PLEASANTNESS OF THE ACTIVITY OF AN R. A PROBABILITY-DECISION MODEL RELATES R'S AND THEIR AVERAGE M-VALUES. THE THEORY IMPLIES THAT MANY REPEATED ENCOUNTERS WITH ABOUT THE SAME SITUATION ALLOW R'S TO BE MORE PREDICTABLE, WITH DEVELOPMENT OF SMOOTH BEHAVIOR SEQUENCES. ALSO INFERRABLE ARE THE LAW OF EFFECT, GENERALIZATION, SATIATION, AND CURIOSITY. THIS FLEXIBLE THEORY CAN BE USED IN CLOSE COALITION WITH COMMON SENSE, EMPATHY, AND INTROSPECTION. DIFFERENCES AMONG REALISTIC LEARNING SITUATIONS ARE DISCUSSED IN TERMS OF DEGREE OF ASSOCIATION SOUGHT, SPECIFICITY AND SYMBOLIC CONTROL OF R'S, AND HIERARCHICAL RELATIONS AMONG R'S. TRYING AND MEANINGFULNESS, DEFINED IN R-M TERMS, ARE SUGGESTED AS 2 FACTORS MOST FAVORABLE TO ANY TYPE OF LEARNING. FINALLY, GENERAL IMPLICATIONS FOR EDUCATIONAL STRATEGY, SUCH AS DEGREE OF LEARNER CONTROL OF THE LEARNING SITUATION, ARE NOTED. (LH)

AIR-D10-11/63-TR(b)

ED014916

EM 006022

LEARNING: FROM R-M THEORY TO EDUCATIONAL PLANNING

Vincent N. Campbell

November 1963

Office of Education
U.S. Department of Health, Education, and Welfare

AMERICAN INSTITUTE FOR RESEARCH/PALO ALTO

AMERICAN INSTITUTE FOR RESEARCH

mailing address: Station A, P.O. Box 11487, Palo Alto, Calif.
Telephone 321-2130 (Area Code 415)
office location: 1791 Arastradero Road, Palo Alto, California

LEARNING: FROM R-M THEORY TO EDUCATIONAL PLANNING

Vincent N. Campbell

AIR-D10-11/63-TR(b)

Technical Report

Grant No.: Title VII 7-48-0000-183

Leslie J. Briggs, Principal Investigator

November 1963

Office of Education

U. S. Department of Health, Education, and Welfare

Abstract

Chapter I presents the fundamentals of a general theory of behavior and experience. R-M theory combines a few familiar psychological variables in a new system designed mainly for parsimony. The mind or brain is schematized as a very large set of elements, each active or inactive at any time and each having a fixed hedonic value (M). Elements are activated in functional groups called representations. A representation (R) is like an idea, percept, image, gestalt, set, plan, or any other molar unit of experience or cognition. Elements form into Rs and Rs into associative structures by increments as a result of contiguous activation. The probability of a decision to overtly carry out a represented plan is a function of (a) the probability that the R of that plan is active, and (b) the average M-value of the Rs associated with the plan.

Some implications of the theory: Many repeated encounters with about the same situation causes routinization, or narrowing of an R-structure to a more certain and predictable pattern, and development of smooth behavior sequences. The law of effect, generalization, satiation, and curiosity are some phenomena inferrable from the theory.

Chapter II. The simplicity and molarity of R-M theory permit its use in close coalition with common sense, empathy and introspection. Its flexibility is condoned as a preliminary condition appropriate to improving the general fit of the model to common phenomena.

Chapter III. A parsimonious theoretical description of diverse realistic learning situations is attempted. Some major ways in which learning tasks and objectives differ are : degree of association sought, specificity of Rs, symbolic control of Rs, and hierarchical relations among Rs. Trying and meaningfulness are defined in R-M terms and suggested as two of the factors most favorable to learning of any type. Differences between formal and informal learning situations are discussed.

General implications of the above issues for educational strategy are noted. Consensus as to what objectives are important appears as a critical planning problem, perhaps further aggravated by confusion between objectives and criterion measures. Choice of optimal learning method depends closely on objectives and on frequency of evaluation of individual learning progress needed. The extent to which the learner controls the learning situation is a rich potential source of improvement in educational methods.

Foreword

Chapter I of this paper is a revision of part of the writer's doctoral dissertation completed at the University of Colorado in 1960. There the theory served to generate some specific predictions in the areas of person perception and social interaction, and the thesis reports research testing those predictions.

The theory and its implications for education as discussed in Chapter III provided the rational background for most of the research conducted under the present grant, which is reported elsewhere (Campbell, 1963). The purpose of the grant was to tentatively identify for several prototype school subject matters the most promising ways of improving individualized programed instruction through self-direction and self-evaluation of learning progress. The specific techniques examined were often selected pragmatically according to the particular requirements of an experimental learning task. The research was thus designed more to explore the practical potentiality of various ways of giving the student responsibility for his learning progress than as a rigorous test of parts of the present theory. The main uses of the ideas discussed herein were (a) in classifying types of learning tasks and objectives, (b) as a source of hypotheses and a basis for judging their relative importance, and (c) as a framework for interpretating and interrelating the results of the various experiments.

Table of Contents

CHAPTER I

GENERAL R-M THEORY

One issue dividing psychological theories during the last half century has been the question of how much attention should be devoted to internal states of the person in attempting to relate his behavior to his environment. An apparent drawback of the theories emphasizing internal states has been the generally low consensus on what kinds of concepts should be used in describing internal states. And while the language available for describing behavior and the environment tends to be more generally communicable, some theorists feel that many important aspects of human behavior cannot be adequately explained using these "external" concepts alone. This belief has motivated many to pursue promising leads in the conceptualization of internal states, be they physiological, phenomenological, or quite abstractly schematic.

The rudimentary scheme presented here is the beginning of one such attempt to conceptualize internal states of the person as a means of establishing orderly relationships among behavioral, environmental and experiential variables. This rough-hewn framework (hereafter called a theory for brevity's sake) is not, as a whole, an offshoot or modification of any single existent theory. Rather, each concept introduced may be considered a blend of a large number of older ideas, with any particular antecedent fairly well obscured in the process.

As we watch a person behaving in a natural setting we can, given time, distinguish thousands of details of the environment which might simultaneously have been affecting the person's behavior, and we can anticipate equally as many combinations of specific movements which the person might display in that situation. If we add to this the large number of internal micro-events supposed by some to mediate stimulus-response relationships, the complexity of the system becomes forbidding. A theory stated in terms of such discernible details would fill libraries, and prediction of significant human behavior from these myriad minute events would seem to require too much time and/or expense to be practical for nearly any purpose.

For this reason the present system is aimed primarily at a molar level of explanation or description, and forthcoming references to behavior and the environment are, in general, meant to be interpreted at this level.

R-M Theory

An overview of the theory may be provided by summarizing the principal constructs:

"Representation," or R: the unit of cognitive activity.

"Association," $p(1/2)$: the conditional probability that R_1 will be active given that R_2 is active.

"M-value,": a motivational or hedonic dimension; the pleasantness of the activity of an R.

Representation

In hope of substantially limiting the amount of information about internal states required for behavior prediction, the principal construct proposed is "representation" or R. Consider the domain of representations (approximately the brain) as a fixed and finite set of elements (perhaps neurons). Each element is either active or inactive at any given time. Elements are not independent nor random in their activitations. Rather, there is some degree of functional grouping, even at birth, and through maturation and experience the domain of elements becomes progressively more patterned and elaborately structured in its activity.

A representation (R), then, is a set of elements which tend to be active or inactive as a unit. Recently formed or seldom activated Rs are more like a "core" than a set in that there is likely to be a gradient of diminishing association among elements rather than a sharply defined set of elements. Activation of a given R is a joint function of sensory input and the activity of other Rs. In relation to older terminology a representation may be thought of as a brain structure, an engram, a schema, or a concept. The activity of an R would correspond to the occurrence of an idea, percept or image, the use of a concept, or the arousal of a set or intention.

A most apparent way in which Rs differ is the degree to which their activities correlate with particular environmental events and overt actions. These differences in "specificity" have traditionally been considered so important that separate terminologies and subtheories have developed for sensation-perception, action, and thought. But putting all cognitive events in one class (representations) subject to the same principles seems worth a try.

Especially important for behavior prediction are those Rs which are active longer and more often. These tend somewhat to be the more central or general Rs not highly correlated in activity with particular acts or events,

such as plans, sets, orientations, and perceived characteristics of self. Any one of these Rs ordinarily plays an active role (remains continuously or intermittently active) over a considerably longer interval of time than do the more specific Rs connected with more explicit acts and environmental features, which often change rapidly. General Rs of this type are closely akin to the themata of Murray (1938), who has convincingly argued that any reasonably short sequence of behavior is guided by one thema, or two or three, perhaps, but not a large number. In this sense the general Rs are more stable over time in their activities and allow the slow human observer more time in which to make his predictions. Of course if the observer wishes to predict specific acts rather than general behavior trends with any precision he would likely have to take into acount many of the more specific Rs that are active.

Stimulus generalization and the substitutability of functionally equivalent responses are common phenomena which fit quite naturally into the present paradigm. In the former, the various stimuli reacted to identically all activate a single R (or at least activate the same set of Rs under the conditions studied) rather than being correlated with the activation of separate Rs, and hence the response is the same. Likewise, many specific Rs such as "hitchhiking," "driving," and "taking the Santa Fe Chief" may all be highly related in their activities (in an adult) to the more general R of "going west," so that a person trained to use one such action in a given situation easily and immediately substitutes another if the first is blocked, because of the common mediation of the general R, "going west."

It is apparent that an adult's total repertory of representations is a complex thing, but the simplification hoped for comes with the assumption that a relatively small number of Rs are active during any short span of time. The stability and constancy of perception despite a continuously fluctuating environment seem to support this assumption. Experience tends to occur in chunks or wholes (Rs), but even when immediate perception is quite panoramic and rapidly changing, the <u>remembered</u> experience is simpler and more categorical (Bartlett, 1932; Carmichael et al., 1932; Gibson, 19[illegible]). Gestalt psychology (Koffka, 1935) and the more recent work of Bruner, et al. (1956) on categories of thinking are well known examples of this viewpoint. The way in which culture and language shape such conceptual categories has been treated at length by Korzybski (1951), Cassirer (1944), and Hallowell (1951).

Association, R-process and R-structure

Actual human behavior over an interval of time would involve the activity of certain Rs simultaneously and/or successively during this time. Such a particular sequence or event in real time will be called an R-process. There is some degree of stability and regularity to the relationships of activity among Rs, just as there is some regularity in environmental events and in overt behavior, so that some R-processes are much more likely to occur than others. That is, the activity of a given set of Rs makes it highly probable that certain other Rs will also be active, and highly improbable that yet other Rs will be active at that time. The degree of relationship between the activities of any two Rs may be expressed as a probability of their joint activation. This is the customary meaning of "degree of association," the conditional probability that R_2 is activated (within a short time interval) given that R_1 is active, or $p(2/1)$. R_1 and R_2 may denote either single Rs or sets of Rs. Association, as used here, appears to be more closely related to older notions of association between ideas or cognitive units (Boring, 1950, pp. 168-176, 250-261) than to more recent conceptions (Hull, 1951; Osgood, 1953; Watson, 1930), which have taken elemental stimuli and responses as the entities associated, for the most part.

An "R-structure" is defined by specifiying a set of Rs and the associations among them. Either the total representational domain of a person or some specific subset of Rs may be treated as an R-structure. To summarize, an R-process refers to the activities of Rs as particular events in real time; an R-structure refers to the probabilities of concomitant activation.

A seemingly unavoidable complication in any realistic theory involving association is that there are quite a few complex associations involving sets of several Rs rather than just two or three Rs. Another way of saying this is that any association between two Rs probably varies at least to some extent as a function of which other Rs happen to be active at the time, particularly as a function of the more general Rs referred to as sets or orientations. For example, the association between the Rs of "tree" and "gift" are probably higher when one has an active R of "the Christmas season" than otherwise.

The M-value of an R

The single motivational construct in the theory is a dimension called M. M does not refer to an entity as does R. Rather, M is a characteristic, a

property of an R which is most simply described as the pleasantness to the person of activation of that R, and thus a dimension shared with many other hedonistic theories including those of Bentham (Boring, 1950, p. 705), McDougall (1908) and James (1890), as well as Titchener (Boring, 1950, pp. 410-420) and the other structuralists, for whom "feeling tone" or pleasantness-unpleasantness was an important attribute. The M dimension is at present treated as an equal-unit or interval scale, so that one may meaningfully compare the difference in M-value between two Rs with the M-difference between two other Rs. The M dimension may later be treated as a ratio scale if a meaningful zero point can be established such that Rs having $M = 0$ are affectively neutral, Rs having higher M-values are experienced as pleasant, and Rs having lower M-values as unpleasant.

The M-value of every element in the R domain is assumed to be constant. Then the M-value of a given R would also be considered constant, except over periods of time during which substantial new learning has changed the composition of the R so that it is made up partly of different elements (as described in the learning principle below).

<u>Principles Interrelating R-process, R-structure and Behavior</u>

As noted earlier, people are born with an R-structure of some sort, though it may be crude and simple compared to the adult's, and months or years of maturation may be required before initial R-structures function effectively. The inherited form and function of the sensory and motor parts of the nervous system may innately determine to some degree the organization of the R-structure which will develop. Perceptual organization on the basis of tempero-spatial contiguity, as in the Gestalt laws of organization, may well illustrate this kind of initial organization.

But considering the great variation among individuals and cultures in language and in behavioral and social organization, it seems safe to assume that the representational domain is structured to a large extent through the experience of the person living in his environment. An important way in which this structuring may take place is postulated next as a learning principle.

<u>Learning principle</u>: <u>Temporal contiguity of activation of Rs (or elements) increases the association among them (probability of future contiguous activations) by an increment.</u>

This is closely analogous to Hebb's (1949) physiological model in which transmission of an impulse across a synapse modifies the synaptic connection so as to make future transmission more probable by an increment. An R would correspond to a cell assembly in Hebb's model.

The size of the increment by which association is increased is a key variable needing more theoretical specification than has been achieved so far. Since association has a maximum value of 1.0, the increment might sensibly be e: es ed as a growth function, decreasing as p approaches 1.0, analogous to the formulations of Hull (1951) and Estes (1950) regarding habit strength and the conditioning of stimulus elements respectively.

If all increments in association were appreciable in magnitude one would expect that at an early age a person's entire R domain would come to be almost constantly active as an undifferentiated whole. Lest this be implied, it is proposed that between some Rs or sets of Rs the increments resulting from contiguous activation are negligible in magnitude. If some associations thus remain very low, then this provides a rationale for decreases in association (as explained later). But this again emphasizes the importance of specifying the factors which determine the size of the increment in association. As a start in this direction one might suggest the degree of anatomical "connectedness" of two Rs as determining the size of the increment in association. This would be consistent with Hebb's model or any other based on synaptic or similar neural connections.

The learning principle implies that among sets of elements or Rs which have low associations with each other, these associations are likely to remain low, while high associations are most likely to increase further. In this way differences among associations are magnified as a result of activity of the R-structures. One original source of differential associations is the crude organization of Rs present at birth. Another early source of differential associations is the presence of regularities in the environment which, impinging on the R domain via a relatively immutable sensory input system, tend to promulgate similar regularities (association patterns) in the R-structure. The manner in which these innate and experiential factors reinforce each other has been discussed by Hebb (1958) and by McClelland et al. (1953).

Thus inherited structure and environmental consistency both enhance development of the R-structure in directions begun in earlier organization. But the learning principle also explains the survival of new organizations, in that the first occurrence of a new R-process makes a reoccurrence more

probable. Changes in the environment and changes in other parts of the R-structure are two likely sources of novel R-processes. Viewed in combination, these two implications of the principle help account for the overall continuity of the R-structure over time, both in the persistence of existing patterns and in the directionality of gradual modifications.

The learning principle applies to the formation of new Rs as well as to associations among existing Rs. Repeated contiguous activation of a number of elements eventually increases the associations among these elements to the point where the whole set of elements acts as a unit, i.e., an R, all (or nearly all) being either active or inactive at any given time. To illustrate, consider the formation of the concept (R) of a "dog." Each specific exposure to a dog in the environment is concomitant with the activation of a set of elements, over-lapping but not identical sets being activated on different exposures. Those elements which most often are active when the person is exposed to a dog will become more highly associated than will elements which tend to be activated only during one particular environmental presentation of a dog. These differential increases in association will be greatest when one is exposed to two or more dogs simultaneously or when one watches the same dog do many things within a short span of time. Those subsets of elements which are most consistently active during exposure to a dog may represent, for example, "furry, four-legged, barkingness" plus a certain size range and perhaps the word "dog." These elements, having reached a high degree of association among themselves, tend to be active as a unit (an R) and hence the concept "dog." This formulation is analogous to classical treatments of concept formation, such as those of Heidbreder (1946) and Hull (1920).

Analogous illustrations could be made at a grosser level, where we would speak of the changing associations between Rs rather than between elements. An example might be the association which accrues to living in a home between the R of "getting food" and that of "being in the kitchen."

Decision principle. The R-processes which affect concurrent overt behavior represent either the person's present state (R_0), or a plan of action and its consequences (R_1), where R_0 and R_1 may refer to single Rs or sets of Rs. If we let $m_{1-0} = M_1 - M_0$ (the difference in average M-value between R_1 and R_0), then the probability of a decision to carry out plan R_1, given that R_1 is activated, is

$$P(D_1/R_1) = m_{1\text{-}0} \cdot p(1/0) \cdot C$$

where C is a constant, a personal parameter.

If no assumption is made about whether or not R_1 is active, then the unconditional probability of deciding to carry out plan R_1 in any situation is

$$P(D_1) = m_{1\text{-}0} \cdot p(1/0)^2 \cdot C$$

This formulation resembles several contemporary theories, such as those of Tolman (1932), Hull (1951), and Rotter (1954), in which the probability of an overt act is postulated to be a function of a motivational component and a habit or expectancy, analogous to m and p(1/0) respectively.

If consideration of plan 1 (activation of R_1) does not initiate a decision on plan 1, the person will continue to consider plans (not necessarily aimed to achieve the same goals) until one initiates a decision, presumably upon a plan which promises pleasant results (m is high) and which the person thinks he knows pretty well how to carry out p(X/0) for plan X is not too small.

If p(X/0) is near 1.0 for each plan in a series of rapid decisions (it cannot approach 1.0 for more than one of a set of alternative plans, as discussed later), a person's behavior appears less deliberative and more nearly resembles a smooth purposeful well-learned behavior sequence.

Other Implied Phenomena

The law of effect, as a principle and as a set of phenomena, may be inferred from the learning and decision principles as follows: The decision principle implies that those "plan" Rs which have the highest M-values and which activate other high-M Rs (positive goals etc.) will be the Rs most likely to initiate decisions leading to the overt action represented. R-structure tends to parallel environmental consistency, so that plans activating Rs of anticipated pleasant consequences do, when carried out, usually lead to the pleasant states expected. It is reasonable to assume also that whereas deliberation of a plan may involve only one or two contiguous activations of the plan and the high-M Rs representing hoped-for goals, overtly carrying out a plan probably involves a considerably larger number of such contiguous

activations because the environment, one's behavior, and attainment of the goal would all tend to produce repeated contiguous activation of the plan-R and the Rs representing goals anticipated and achieved by the person. By the learning principle, those Rs most frequently active contiguously have their associations increased most. As a result, in future decisions the consequences of plans carried out are more likely to be anticipated (and to affect the decision) than are consequences of plans not carried out. Hence those plans acted out (which tend to be those leading to pleasant consequences) are more likely to be repeated. This derivation of the law of effect holds to the extent that consequences of plans are anticipated correctly, and they do tend to be, according to the learning principle.

As mentioned earlier, associations among some Rs may remain near zero, just as some actions may be competitive or mutually exclusive. This relative constancy of low association between some Rs may explain why yet other associations even decrease over time. To illustrate, let us depict some of the R-processes of a person learning to drive a car. On his first attempt at driving, when he approaches a stop sign (the R "stop" is active) he may consider two alternatives, "push brake" and "push gas," with about equal probability, say .40. Assume that for reasons analogous to the incompatibility of the responses involved, his Rs of "brake" and "gas" must have low or zero association with each other. According to the decision principle and the law of effect, we might expect that after the first trial or two he would consider "brake" more frequently than "gas." By the learning principle, each additional contiguous activation of "stop" and "brake" would increase the association from .40 toward 1.0. But since the "brake-gas" association must remain near zero, this implies that the "stop-gas" association must decrease toward zero as the "stop-brake" association increases.

This illustration typifies the changes in R-structure which accompany "routinization" of learned behavior with successful practice, i.e., the gradual change in behavior sequences from hesitant, deliberative, trial-and-error to smooth, efficient, coordinated behavior often described as automatic or habitual.

In conventional language we would describe the initial trial-and-error behavior as occurring when (a) the stimulus situation is ambiguous, or (b) given a perception of the situation, no single response has been well learned; under these conditions conflict and vacillation occur. In the present language both (a) and (b) may be paraphrased with the statement that associations

between the Rs activated in sequence in this situation are not high, so that (by the decision principle) at each step several alternative plan Rs may be activated before one leads to an overt decision. As the person is repeatedly confronted with a quite similar situation, he more quickly decides upon the most adaptive (in terms of M-value) plan at each step. Thus, some associations increase while others remain low or even decrease. The result, applied to any behavior sequence, is a more certain and narrowly limited sequence of Rs being active as a person is exposed more and more to similar situations (where situation includes the person's general Rs as well as the environment), less deliberation and fewer errors, hence a smoother, faster sequence of behavior.

When this routinization has evolved to a high degree, as in driving, smoking, or skilled typing, the R-processes involved become so highly specific and stable that the person is able to engage simultaneously in other non-competitive behaviors. In the above car driving example, on his first few attempts the driver may have to devote all his attention to his driving behaviors. When the skill has become well learned (the R-processes routinized), he may easily engage simultaneously in conversation because much of the R-domain is freed (left inactive) with high routinization, and those specific R-processes involved are less amenable to disruption.

In general, routinization in a given kind of situation would probably be accompanied by at least minimal success in that situation, since (by the decision principle) R-processes leading to pleasant states are more likely to reoccur, and thus become routinized, than R-processes leading to unpleasant states. Success from an <u>observer's</u> standpoint cannot be called a necessary covariate of routinization, as evidenced by clinical observations of "fixated neurotic" behavior (Mowrer, 1948) and by animal studies (Maier, 1949) in which maladaptive behavior has become habitual. The "failure" behavior so developed may represent the least of evils for the behaving organism, however, in that any alternative R-processes activated in that situation may have even lower M-value. If such were the case, acquisition of this failure behavior could be considered adaptive in the sense of maximizing M-value, because routinization presumably involves the channeling and narrowing of active R-processes, which would mean that <u>fewer</u> Rs representing the unpleasant situation would remain active. Routinization enables the organism to simultaneously attend to and cognize other more neutral or even pleasant matters so that the <u>average</u>

M-value of all active Rs is higher (less negative). This might explain a good deal of compulsive and ritual behavior. That is, many ritual habits may be ways of maximizing M-value by substituting rather neutral R-processes for the very negative ones which are activated by a repeatedly and inescapably unpleasant situation.

This argument may also provide a fruitful theoretical approach to satiation and curiosity. Satiation may be described as a decrease in the average M-value of the active R-processes. As noted, repeated or prolonged experience in a given situation tends to lead to routinization of the R-processes and a consequent reduction of the number of Rs activated in that situation. If an initially pleasant situation involves activation of Rs predominantly high (positive) in M-value, then routinization of these R-processes should be accompanied by a lowering of overall M-value, since fewer of these pleasant Rs are retained in the person's total R-process as routinization progresses. As satiation becomes greater and the average M-value of the R-processes lower, the person may become more likely to represent his present state as less attractive (lower in M-value) than alternative situations which he anticipates, and according to the decision principle his probability of overtly seeking one of the alternative situations would thereby increase. This quest for new situations, or variation in behavior as a function of satiation, may be thought of as characterizing "curiosity."

In routinization, many Rs which are initially only slightly associated come to be active as a unit so that the whole process can then be treated as a single R for purposes of behavior prediction. Usually a routinized R-process involves correspondingly automatic overt action. The predictability of the outcome of such routines makes practicable a molar theory which ignores the detailed events <u>within</u> a routine. It is an important insight of the rapidly growing cybernetic approach to behavior initiated by Wiener (1948), that in such behavior routines it is typically not a fixed series of responses but rather the person's intended relation between himself and the environment which is the predictable outcome of executing the routine. In other words, when a person decides to carry out an already routinized plan, his R of that plan usually determines the outcome through continuously operating neural feedback loops connecting the R-domain to the musculature and the environment. For example, when a person decides to go to lunch at the cafe down the street, as usual, this entire plan may be carried out routinely while the person thinks

about other matters, even though the person never makes exactly the same sequence of movements on two different trips to lunch. Part of the routine is opening doors, avoiding other people in his path, watching for the caf, entrance, etc. Each of these minor acts is itself a previously learned subroutine which could be further analyzed into more molecular subroutines. There is evidence (Hershberger, 1962) that the more molecular subroutines may be largely innate rather than learned. A plausible discussion of how larger routines or plans of action may be developed and be related to more molecular units has been provided by Miller, et al. (1960).

CHAPTER II

METHODOLOGY

How does one go about inferring the presence or activity of certain Rs in another person in order to better predict his behavior? The problem is fundamentally the same whether one is a behavioral scientist or a layman trying to behave successfully in a social situation. The starting basis for either observer is his assumption that the subject's relevant R-structure is similar to his own. The observer takes into account the subject's surroundings, his current behavior, and the information given to the subject (S) about his present situation. Grossly speaking the observer then pretends he is in S's shoes and infers the latter's R-processes from his own.

Not even the layman behaves this simply in predicting others, of course, unless he knows nothing about the person whose behavior he is predicting. One only assumes the S's R-structure is the same as his own as a baseline from which specific differences in R-structure are inferred on the basis of other information about S. Another major source of information about S's R-processes are his own verbal reports. This source of information is one good reason for defining Rs at a molar level which centers on consciously reportable ideas and images, though it is important to reiterate that unverbalizable, subconscious or unconscious experiences are R-processes as well. The scientist and clinician are trained in their different ways not to assume isomorphism between mental process and verbal report. Perhaps this has been overemphasized. Most of the time most people reflect their thinking (R-processes) pretty accurately with their words, I submit.

In addition to the S's verbal report, his past behavior and his immediate environment, such group membership variables as age, sex, social class, occupation, and religion are often used by both scientific and lay observers in modifying their inferences. In general, however, the scientist is more cautious about relying on such variables just because of cultural acceptance of them as bases of inference, and at the same time he is more optimistic and daring in making inferences from new and esoteric variables which a limited amount of scientific research has shown to be of possible value in behavior prediction.

Another characteristic distinguishing the scientific observer is the importance he places on making his observations and his theoretical basis for

inference explicit and public so that they are verifiable. A great deal of private introspection and intuition is indispensable to the whole process, but this private aspect is generally considered slightly disreputable and so is not given the attention and public scrutiny which might improve its use. R-M theory is meant to explicity incorporate introspection by the scientist and assumed similarity of others to himself as valuable bases of inference. This amounts to the scientist putting himself into the same field of variables as his subjects rather than assuming the role of an objective truth-seer from another world.

Within this theoretical framework there are several ways in which the scientific observer can infer differences between the S's R-processes and his own:

1. He can assume the same R-structure but different M-values. That is, the obersver (O) can assume that the same Rs are activated in the S in his situation as are activated in the O playing the role of S, and further that the associations between the Rs involved are the same. The inferred difference between O and S is in the M-values of certain of these Rs. This is the approach usually implied in value and attitude assessment. The object or goal referred to in the assessment instrument (church, negroes, war, etc.) is assumed to be represented in an equivalent manner by all the respondents (all have an R activated by the label "church," for example, which has certain associations with other consensually validated Rs, such as "place of worship," "reverence for supreme being," etc.). The individual differences which the instrument presumably records then are differences in the M-value of that R.

2. A second kind of inference about differences in R-structure between O and S is made when O assumes that S has essentially the same Rs active in this situation as O does, and that these Rs have the same M-values for O and S, but that the associations between the Rs are different for O and S. This paradigm fits the study of problem-solving behavior where S has to choose between known alternatives. It is assumed by O that he and the Ss have in mind the same general alternative plans, but that there may be disagreement as to which plan is most likely to achieve the goal that all desire.

3. A few systematic techniques assume that S may construe the situation in a different way from O, that O and S may have entirely different Rs activated by the situation. Projective tests work from this assumption, as does Kelly's (1955) Rep test.

To summarize the above discussion, in trying to predict another's behavior we put ourselves in his place and use our own R-processes as a framework for inferring his. But we also use the subject's self-reports, his past behavior, and the behavior of others in this type of situation as a basis for inferring his R-processes even though they be different from our own. Three simple ways of inferring observer-subject differences were illustrated above. More complex combinations of these would probably be required in most practical prediction problems.

As for statistical aids to inference, which are so heavily emphasized in contemporary behavioral science, the present framework would seem to be best facilitated by a Bayesian (Edwards et al., 1963) approach, in which the scientist's subjective probabilities prior to each experiment are taken into account.

This discussion has left considerable room for variation and improvisation in the detailed procedures of making inferences about R-processes and R-structures, partly because the theory is in a primitive stage where it is meant to be tried on loosely and modified freely.

CHAPTER III

LEARNING

The R-M learning principle says that joint activation of Rs increases the association between them, i.e., the probability of future joint activation. Learning is restructuring of the R-domain by changing associations. As discussed earlier, two major factors determine the directions of change in an R-structure:

1. The prior existing structure. R-structure is primarily inherited in the infant, but the effects of experience become more important with increasing age. Whatever their origin, existing R-structures tend to perpetuate themselves, since contiguous activations of those Rs tend to make future contiguous activations even more probable.

2. Environmental consistencies. Because of relatively inflexible sensory systems, the patterns of contiguous activations of Rs will partially parallel environmental patterns, and (by the learning principle) will thus promote R-structures which also tend, in time, to parallel environmental patterns. This parallelism is well described in an information-communication framework. That is, the sensory system "transmits information" from the environment pretty faithfully, despite the findings of the aging "new look" in perception (Bruner and Goodman, 1947) and physiological "gating" (Hernandez-Peon et al., 1956) which are noteworthy as exceptions to the rule.

Lasting and major systematic changes in the environment would, by the same token, imply a roughly corresponding degree of change in R-structure. In general, the R-structure matches itself to changes in the environment (to the degree it does), not suddenly, but by a gradual process, gradual because of the tendency of existing structure to perpetuate itself and thus resist change. Or, more fundamentally, the change is gradual because associations change only by increments as a result of contiguous activation. This means that the speed with which a person can successfully adapt to an environment by forming a matching R-structure should be greater for a more constant environment.

Prior structure and environmental consistency shape the overall course of learning through a lifetime. But this doesn't get us far in predicting or improving learning in specific situations. A first step in this direction is to see whether the diverse phenomena of learning can be described more simply within R-M theory.

Ways in which Learnable R-structures Differ

The variables to be discussed here were chosen for their judged importance in differentiating among real-life learning objectives.

Degree of Association

It might seem at first that if one wishes to establish an association between two Rs, one would always try to maximize the degree of association ($p = 1.0$). Not necessarily. The most apparent reason for limiting p to a moderate value is to permit greater flexibility in future R-processes. In most formal learning areas, knowledge is uncertain enough that the educator may gain by respecting the variability of nature and its important unknowns. If all associations among Rs in a certain knowledge area were near 1.0, there would be a higher degree of routinization of R-processes. Moderate associations on the other hand, increase the probability that the person will consider alternatives to each idea or plan entertained. Expressed another way, breaking an inappropriate mental set, so valuable in science and other creative endeavors, is more difficult the more habitual and routinized are the cognitive associations.

A related and perhaps even more important consideration is the finiteness of the R-domain, which imposes demands of parsimony on the mental apparatus of the behaving organism. Man's brain and his behavior are exceedingly versatile compared to contemporary electronic simulators. Ultimately the R-domain's parsimony depends on how many different R-processes a given number of elements can support. This flexibility is inversely related to the "certainty" of the R-structures, i.e., the extent to which all p approach zero or one. A timely illustration of the difficulty of regrouping elements into a new R-structure when the old one is highly overlearned (routinized) is that adults have more trouble than their third-grade children with modern approaches to arithmetic.

Our species in its evolutionary wisdom may have developed a brain generally resistant to sudden permanent learning in order that only those associations which prove to be important and adaptive in the long run will be retained in more enduring R-structures. Since learning occurs by contiguous activation of the Rs or elements involved, if every environmental change and new R-process dramatically changed the associations so as to ensure repetition of that R-process, the result would be a maladaptive, unstable R-structure shifting

with every fortuitous momentary event. Only if associations change by small increments can the law of effect operate (as explained in Chapter I) so that the more adaptive R-structures survive and the less adaptive R-structures do not.

Viewed in this way, there is no cause for regret when students remember little of any particular reading assignment, if by "remembering" one means literal recall of particulars. It may be that the important things to learn in some topics (e.g., history, art, literature) cannot be acquired by drill or response-oriented instruction, because there are too many specific Rs involved to cite them all as learning objectives and the general Rs cannot be activated at will, but rather depend upon gradual R-formation. It may be that reading and forgetting details and thus building up fairly "uncertain" R-structures over a long period of time results in the most desirable R-structures with respect to these subject matters. In learning the multiplication tables, on the other hand, one might seek to establish associations approaching 1.0, and rote drill may be the best method for doing so.

Since association is defined as a conditional probability, it is a directional relationship, and p(1/2) may differ from p(2/1). The learning of procedures such as assembling or trouble-shooting often means increasing the association primarily in one direction corresponding to the order in which the Rs are activated in the criterion situation. A strong difference in emphasis between p(1/2) and p(2/1) is less common in the "conceptual" or academic learning areas, in which reversibility of thought processes is usually sought.

R-formation vs. R-association

Developing new ideas, forming concepts, abstracting general properties from specific events. These involve R-formation, the grouping or regrouping of elements into new functional units (Rs). R-formation is probably the most difficult and challenging type of learning.

Establishing associations among Rs which already exist in the learner presents quite different instructional problems from those of R-formation. Associations among names, places, facts, and familiar ideas are typical R-associations. Associating such already established Rs is greatly facilitated by the fact that an existing R can be (and usually is) highly associated with a particular symbolic or verbal label. This enables the instructor (man, book, or machine)

to activate such Rs quickly and predictably by symbolic communication. Since learning is effected by contiguous activations, the advantage of symbolic control can be great.

R-formation, on the other hand, is at best only partially and rather unpredictably controlled by symbolic communications, because symbolic labels are associated with Rs rather than elements. In trying to regroup elements into a new R, a symbolic input might activate some of the appropriate elements but it would also activate some inappropriate ones. For example, when an R to be labeled "energy" is first being formed, "light," "heat," motion," and other verbal inputs would each activate some but not <u>all</u> of the appropriate elements, and each of these words would also activate some fairly irrelevant elements.

It may avoid confusion at this point to note that R-formation does <u>not</u> include the kind of experimental concept-formation task (e.g., Bruner et al., 1956) in which the subject searches for the "correct" combination among <u>familiar</u> alternatives such as "black," "white," "square," "triangle," etc. That kind of problem-solving with older children and adults involves R-association rather than R-formation, in that the task requires associating only Rs which already exist and have symbolic labels. Forming the concept "square" (or "black" or "triangle") in the first place during early childhood is more typical of R-formation.

R-formation is probably the predominant type of learning in the preverbal child (typified by the "dog" example in Chapter I), with the predominance diminishing gradually over the years as the person's repertory of Rs grows and enables him to succeed most of the time without forming new Rs.

<u>Hierarchies</u>

The associational patterns which learned R-structures form may be infinitely varied, but hierarchies of Rs represent a pattern which has especially direct implications as to what learning conditions would be most effective (Gagne, 1962). A hierarchical structure in this context means that certain Rs low in the hierarchy must be formed and/or associated before other Rs higher in the hierarchy can be activated. For example, in mathematics it is convenient, if not necessary, to form the Rs "set" and "element" before trying to form the R "intersection." The main implications for method of instruction are the importance of order of establishing associations and Rs, and hence also of evaluation at each major step of learning. If a student fails to

learn something low in the hierarchy and this is not discovered, subsequent instruction may be a waste of time for him.

Ordered Specifics

Another important aspect of a learning requirement is the degree to which the Rs to be associated represent specific acts or events. A good deal of on-the-job technical training, such as for assembly or clerical tasks, involves associating quite specific Rs in chain sequence or in some other orderly pattern.

In formal education learning requirements of this type have diminished during this century as less importance has been attached to verbatim verbal memory and more importance to such educational objectives as informed judgment, creativeness, and practical decision-making. Of course the learning of a language itself involves many highly specific Rs, but this major area of learning is probably best left outside the present category of "ordered specifics," because acquisition of language seems so intimately determined by general meanings and intentions, i.e., by the concurrent activity of much more general Rs.

As discussed earlier, symbolic control is greater with R-association than with R-formation. In associating very specific Rs successful environmental control of the learner's R-process is especially likely in that there is usually virtual isomorphism between the specific R-structure and the overt actions and environmental events (symbolic or otherwise) which the Rs represent. Thus by merely eliciting the proper actions from the learner one can be fairly sure the appropriate R-processes are occurring. It is in the learning of such specifics that ultra-behaviorism which looks only at external events predicts best, but even here such an approach can go far astray if the learner's general orientation to the task is not taken into account. In most real-life learning the activity of more general Rs is an integral part of the R-structure to be learned, so that there is not a dependable isomorphism between R-process and external events.

Perceptual-Motor Skills

That initially rather discrete decisions tend with many repetitions to become a smooth rapid behavior sequence as described in Chapter I, is particularly appropriate for describing the learning of skills, such as driving, skiing, or juggling. As skills of this sort develop beyond initial hesitation,

the coordination of R-process, sensorimotor apparatus and environment becomes so close and continuous that it seems awkward to talk of separate systems. Cybernetic models such as those mentioned at the end of Chapter I would seem better than the present one for studying the acquisition of perceptual-motor skills.

This classification of characteristics of R-structures formed by learning has problems and rough edges, but aspires nevertheless to be a common framework for treating traditionally disparate categories of learning phenomena. The distinction between gradual learning and "insight" should be mentioned explicitly here. Gradual learning corresponds to the incremental increase of any given association as defined in the learning principle. Insight learning takes place when entirely different Rs are activated by the same external situation and this new R-process leads to greater success.

Factors Favoring Learning

Again emphasizing applied learning situations, let us turn to factors which greatly enhance learning for any type of R-structure.

Trying

Probably the most visible of all sources of variance, given a learner and learning task materials, is the extent to which the learner tries, either tries to learn or tries to reach a goal which requires learning. The biggest detractor to learning, it seems, is that the would-be learner decides to do something else. He may overtly leave the field, or, more insidiously, he may maintain the external facade of learning in order to avoid losing his job or his passing grades, while cognitively escaping to greener pastures. The contention here is that while incidental learning is possible, it is inefficient for most formal learning objectives. That is, most learning requires the learner continually to decide to attend to the relevant problem or subject matter. Each time he decides to attend to something else, he activates an irrelevant R-process which interferes with the relevant one, except in advanced stages of overlearning where the relevant R-process is already pretty well routinized. He may also cut off relevant sensory input altogether, e.g., by looking up from his book.

The R-M decision principle suggests two fundamental factors determining which way each decision will be made (to attend to the learning problem vs.

something else): the relative M-values of the Rs activated, and the associations among them. Decisions to attend to the learning task are more likely when the task involves pleasant thoughts and pleasant anticipated consequences (high M-values) and when attending to the task seems to be the easy and natural thing to do next (high association between R-present and R-task). It should be noted here that in talking of "trying" I am not introducing "drives" or "needs." The decision principle should suffice.

Of course satiation enters the picture when the time span considered is enlarged from minutes to hours, and attempts to maximize the frequency of decisions to attend to the learning task would do well to build in breaks (distributed practice). Another important way to avoid satiation and decisions to "defect " is to build variety into the learning task itself.

One typical obstacle to maintaining attention to the learning task seems to be the attention-compelling quality of a marked change in the pattern or sensory input, e.g., a sudden silence, or an unfamiliar voice. In R-M terms, such a change seems innately to activate a plan-R which might be labeled "attending to source of change." The Rs activated by these distracting inputs provide some of the alternative plans for the learner as he repeatedly decides to focus on either the learning task or something else.

Although its role toward the upper extreme of motivational intensity is less certain, "trying" is certainly a factor favoring learning in the regions of lower intensity where the decision to attend to the task or not is involved. The point seems obvious, and for this very reason it may have been understressed in learning experiments.

Meaningfulness

A second major source of variance in rate of learning is meaningfulness of the task to the learner. Learning theorists and educators alike have granted the importance of "meaningfulness of material" while noting also its resistance to definition.

Meaningfulness seems to translate well in R-M theory. The representational correlate of experience (and behavior) was said to be activity of Rs. The more Rs that are active at a given time, then, the more meaningful the experience. Recalling the paradigm in which the total R-domain is a finite set of elements, meaningfulness may be defined as how great a proportion of the the R-domain is active.

Some things seem to be considered meaningful because of the richness of sense imagery involved. Sensory input is a principal source of activation of elements or Rs. The "denser" the sensory input and the more sensory modes utilized, the greater the degree of activity in the R-domain in general, although the degree of R-activity depends also on what particular pattern of R-activity has preceded.

Another aspect of meaningfulness, as ordinarily conceived, has to do with familiarity or degree to which this situation has been cognitively structured before. But the relationship is complex. Moderately familiar situations, all might agree, tend to be more meaningful than kaleidoscopic confusion, for a given complexity of sensory input. A good share of the R-domain acquires its structure through the experiences of the person. If sensory input is to activate such "learned" structures, there must be some commonality of pattern between the new input and the previous input which played a part in developing the existing R-structure. (This familiarity or commonality should include the general R-process of the learner as well as sensory input.)

Yet if the total situation (R-process plus input) remains too constant or is very frequently repeated (e.g., as in walking) routinization gradually reduces the total amount of R-activity, and thus reduces meaningfulness. This curvilinear relation in which meaningfulness is greatest when the situation is only moderately familiar seems consistent with general observations on the relation of familiarity to meaningfulness.

The substantial co-variation between meaningfulness and rate of learning follows from the R-M learning principle. Learning is modification of associations among Rs, and associations change as a function of contiguous activations of the Rs involved. The greater the activity in the R-domain, therefore, the greater in general will be the changes in patterns of associations, i.e., the more learning will take place.

This proposition becomes clearer if we examine it with respect to a pair of elements or Rs. Suppose that each contiguous activation of the two Rs, A and B, increases the association between them by a nonnegligible increment. The more meaningful the learning situation, the greater the number of other Rs which are activated at the same time as A and B. Associations among many of these other Rs will increase as will their associations with A and B. As a result, both A and B are more likely to be activated contiguously in the future whenever any of these other Rs are activated. This means that contiguous

activations of A and B will tend to occur more frequently in a more meaningful context (more other Rs active), and the A-B association will thereby increase faster. Hence, meaningfulness is conducive to faster learning.

Formal and Informal Learning

Although the main intention here is to work toward a useful theoretical framework for research on improvement of formal education, it may lend perspective to consider the major differences between typical formal learning situations, such as in public school, and informal learning situations.

Nature of Environmental Input

First, those human learning situations broadly classed as formal usually present the learner with symbolic material, usually verbal language, as the subject matter to be "learned." Informal learning, on the other hand, usually takes place in a more natural setting and the input to the R-domain tends to be quite complex and varied, often involving several sensory modes at once. In other words, informal learning situations are usually more meaningful, and this permits faster learning, as discussed earlier.

Evolution is helpful again in explaining the greater meaningfulness of informal learning. During most of the era of man, mentally representing and associating concrete events such as food, mother, animals, enemies, etc., were probably more crucial to his survival than were his more abstract symbolic manipulations, although this may not be true at all now. For this reason, our mental equipment is probably primarily designed for efficient processing of ordinary sensory representation more than for symbolic input. On this basis one might expect that more time or trials would be required for symbolic learning, and the handling of review and repetition would therefore take on greater importance in formal learning situations than elsewhere.

A related distinction between formal and informal learning situations has to do with the relevance of environmental input to the learning task. Aside from the symbolic input itself, the formal learning situation usually is set in a physical context which is unrelated to the learning topic. To illustrate, for a person learning to survive in the wilderness a substantial proportion of the environmental input is relevant in some degree to his learning task. But for the schoolroom learner, the tables, chairs, walls, and papers he is exposed to seldom serve directly as clues to the understanding

he seeks. In terms of information theory, not only does the school environment contain less information (over time) than the "real-life" situations, but also there is less transmission of information possible between the schoolroom and the task-relevant R-structure of the learner. This would make learning slower and more difficult in formal than in informal learning situations.

Motivation for the Learning Task

Another difference between formal and informal learning situations is that the formal learning task itself is usually less attractive than the informal situation. The primary reason for this may be that in formal education usually someone other than the learner himself decides specifically what should be learned. To the parent or teacher, the long range benefit of studying Latin may be clearly visualized, but the pupil is less likely to see this. He may not even agree on the merit of the general educational goals. Thus the formal learning situation is often deficient in two aspects essential to the student's decision to try to learn the material: The represented goals may have low M-value for the student, and the association between "studying" and achieving any high-M-valued goals may be low: both of these lower the attractiveness of the task. In terms of the decision principle, M-value is low for the plan, "to study," so the probability of deciding to study is low.

In an informal situation, what is learned must seem more worthwhile to the learner because usually, either (a) he chooses that situation because it intrinsically interests him (M-value for the situation itself is high), or (b) he is trying to solve some "real," immediate problem and he sees his performance in that situation as relevant to the solution (associations between his Rs of plans in that situation and Rs of very positive or very negative M-valued goals are high).

Communication

Formal learning appears to depend more upon communication between persons than does informal learning, in general. This would seem to have both beneficial and detrimental implications for the learning process.

To the extent that the teacher or writer can anticipate and solve the learner's problems, communication may give formal learning an advantage, especially in view of the freedom of arrangement of input in the formal learning situation. On the other hand, most of the material to be "learned" in the

formal situation must come to the learner as symbolic communication from other persons, and whatever obstacles there are to interpersonal communication would tend to impede formal learning.

A basic failure of this sort may be lack of consensus on the goals in formal learning. Aside from differences in M-values of mutually represented goals, as just discussed, the learner may have no Rs of relevant goals corresponding to those Rs in the teacher. The pupil might be more enthusiastic if he could see where he was headed. The significance of this characteristic of much formal learning is that the student, not having clearly represented and valued goals, lacks relevant criteria with which to evaluate his own progress and modify his course when appropriate. For example, his decision as to what to do next may be based on the +M-valued R of "finishing the lesson," rather than on represented deficiencies in his learning.

Another general type of communication problem typifying formal learning is that the teacher or writer has acquired his R-structure slowly over years of study, but he no longer remembers the early stages of his learning nor how he represented the learning task when he first encountered it. In fact, the success of his own learning was usually judged in part by how well he forgot his early errors and retained only a refined end-product of knowledge. As a result, the teacher or expert often tries to directly "implant" his own highly refined R-structure into the naive student. He may be successful to some degree, or he may fail if the learner is overwhelmed by the complexity and meaninglessness of the material in his early encounters with it. (This is an especially common experience in mathematics, it seems.) In informal situations the learner is usually more familiar initially with the issues and specifics involved, so that he may make more substantial use of his existing R-structure from the start.

Communication between teacher and student is facilitated by similarity of their total R-structures, i.e., the extent to which each "knows what the other is talking about." The total R-structure is said to be the communication basis, rather than solely the part representing the learning objectives (which becomes similar only as the student learns), because a crucial function of the teacher seems to be explaining difficult points and removing misunderstandings. To do this, he must "understand the misunderstandings" and use illustrations and analogies which are meaningful to the learner in terms of what the learner already knows. This frequently involves reference to other

knowledge, apart from the learning topic, shared by teacher and learner; hence the importance of similarity of the learner's total R-structure to the teacher's (or some part of the teacher's).

Educational Strategy

Within the present framework, formal education appears to confront the planner with three major phases:

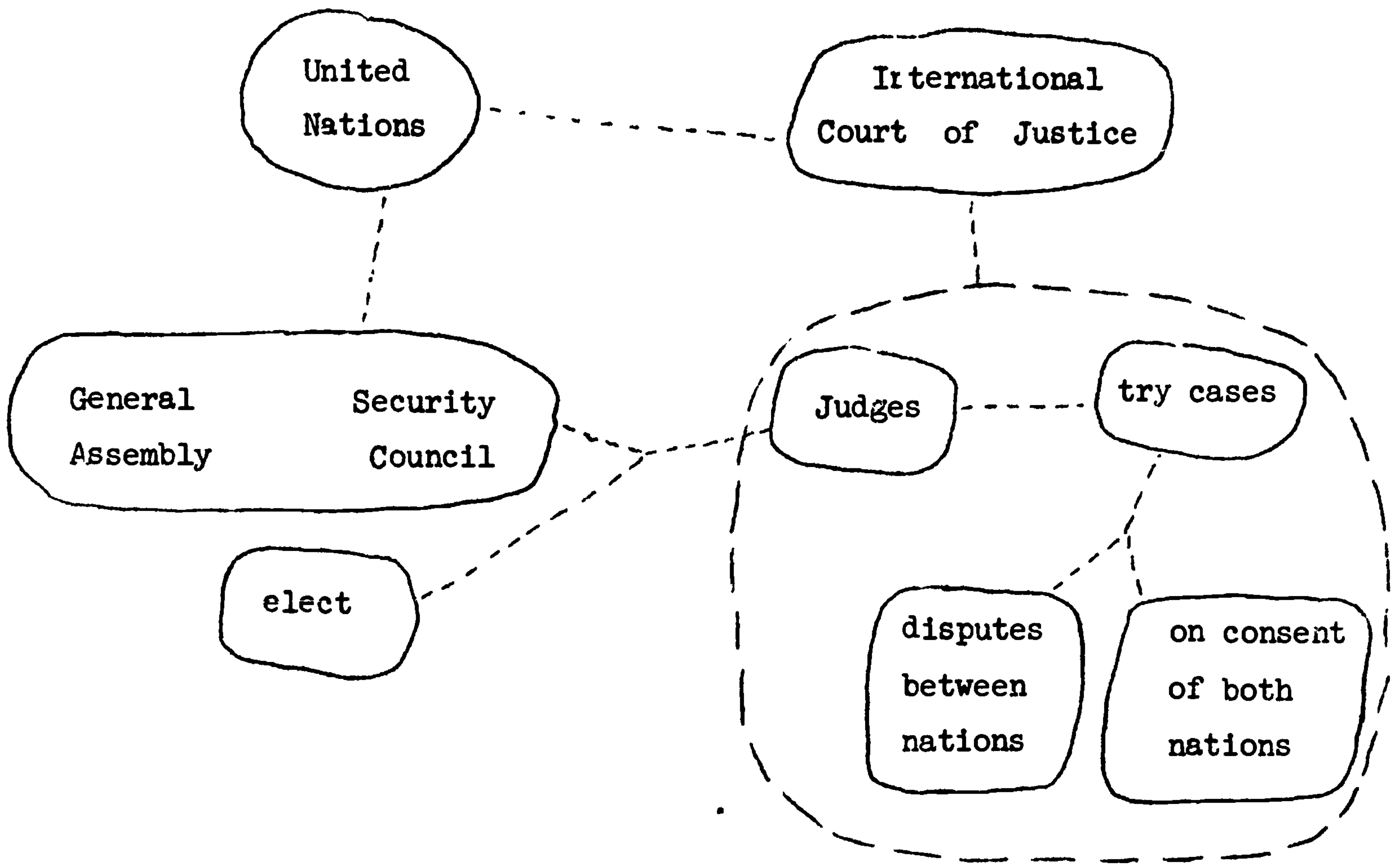
1. Objectives: changes sought in the R-structure of the learner.
2. Learning Methods: efficient ways to activate Rs in combinations and orders most likely to achieve the objectives.
3. Evaluation of learning progress.

Stating objectives is an ideal first step if one is starting from scratch, but short-term educational planning usually necessitates considering all three phases at once; first, because freedom to select learning and evaluation methods is economically restricted by existing commitments to certain types of buildings, equipment, materials and teachers, and second because consensus among planners on learning objectives is often low enough to justify basing choice of objectives partly on availability of resources. For long-term planning beyond existing economic commitments, however, establishing objectives should clearly be the first step, and adopting learning methods most likely to achieve these objectives should follow.

Learning Objectives

For well circumscribed learning tasks, which are usually skills or ordered specifics, stating objectives is simple and straightforward because there is presumed a close correspondence between what is learned and what is done or verbalized. Thus objectives and evaluation criteria are merged and considered essentially identical.

But the great majority of objectives in formal education are not neatly circumscribed nor can they be stated in terms of behavioral specifics. At best one has a core of important Rs and associations with a gradient of associations which diminish gradually in importance as they become more remotely associated with the core. For example, consider the objective of teaching the essentials of the International Court of Justice. The core R-structure to be achieved might be that diagramed below:



The basis for determining "importance" or "coreness" might be either the judgment of one authority or the degree of consensus among planners. In either case the diagram excludes thousands of related associations involving, for example, what the Security Council does, and especially involving the more general Rs such as "disputes between nations," which could itself involve a related R-structure requiring a lifetime of learning. What should the educational planner do about this enormous and bothersome surrounding network of relevant associations? The gradient of importance is seldom so steep that he can afford to ignore it and concentrate solely upon achieving the core R-structure, not even in mathematics, in which the core is perhaps more clearly defined (the importance gradient steeper) than in some other fields. Why not? Because the best way to learn arbitrarily circumscribed cores may not be the best way to learn a whole network of associations. This seems especially critical in view of the importance of transfer and synthesis as general goals of education.

The problem cannot be given the consideration it deserves in this paper, but one approach might be:

(a) concentrate on the core in proportion to its importance relative to the "enrichment," but

(b) give preference to learning methods conducive to enrichment learning as well, and

(c) in evaluating progress, include criteria which give credit _ad hoc_, for unanticipated but relevant associations.

For fields which have rather flat gradients of importance, such as history and literature, it may be difficult to identify a core at all. Since creditable R-structures learned may differ highly from one individual to another, _ad hoc_ judgments of learning progress would seem especially important here, as would the use of learning methods which motivate each learner by allowing him to pursue objectives most interesting to him. In stating objectives it seems worthwhile, then, to indicate the degree to which the core R-structure (the explicitly stated objectives) omits relevant associations.

As for the degree of association sought, it seems artifically precise at this time to state numerical probabilities. "Near zero," "low," "moderate," and "high" would be adequate for most purposes.

Learning Methods

Techniques for activation of Rs in the desired sequence derive from two sources: activation by sensory input and activation as a result of other R-activity. Most actual learning probably involves both sources of activation simultaneously.

The main criteria for choosing among activation (learning) techniques should be the "efficiency" of establishing the desired R-structure and permanence of the new structure (retention). The word "efficiency" implies a dimension of expenditure against which amount learned (degree to which R-structure is formed) must be balanced. These bases of expenditure seem important:

(a) _Time_ expended by the learner. This is probably the most important base, since human learning time is so limited.

(b) _Cost_ (financial) of the technique. I assume that a teacher's time and trouble, as well as development of materials, can be translated into monetary value.

(c) Average M-value of the learner's R-processes during learning (pleasantness). Learning that is fun is certainly to be preferred.

Finding the basic characteristics of learning techniques which relate most highly to learning efficiency seems to be the fundamental problem of education. A comprehensive discussion of the problem is beyond the scope of this paper, but a few guidelines will be proffered.

In discussing ways in which learnable R-structures differ, implications for different learning techniques were touched upon. Where moderately low associations are sought, a single activation of the appropriate Rs may suffice. Reading, attentive listening, watching films or TV, and other passive modes of instruction which may be inadequate for establishing stronger associations may be the most efficient modes for low-association structures because they can cover material so rapidly. A major problem of these passive modes is keeping attention, but for low-association requirements little or no repetition is needed and this may make the task of holding the students' attention easier.

Where higher associations requiring several contiguous activations are sought, one must decide whether it is preferable to activate the same Rs repeatedly until a given association is established, or to activate a number of different R-sets in alternation (part vs. whole learning). Fatigue, satiation, and routinization tend to lower the M-value of an initally pleasant R-process, which in turn makes distraction more probable. Alternation of R-sets might better avoid distraction, then, than many repeated activations of the same R-set. When the objective is to form a new R mainly by activating a number of existing Rs, an additional reason for alternation among the old Rs (rather than repeating each one several times in succession) is that this would help to minimize the number of irrelevant elements associated with the new R.

When newly formed Rs are to be associated in a hierarchical structure it seems quite important to be sure each R is formed before associations higher in the hierarchy are undertaken. In this situation repeated activation of the new R (perhaps initially activating it by alternation among old Rs, as described above) would be better than alternation among entirely different R-sets. Restated, this means building on prior knowledge, or proceeding from what is already known by the learner into new territory rather than plunging him suddenly into the unknown. This is essentially what was advocated by Herbart (Boring, 1950, p. 257) and Dewey (1899). Unfortunately it seems some

interpreters of these pioneers noted only the part about beginning with what is meaningful and left their learners right there rather than leading them into the abstractions of higher learning. The important point of the message was the order of progression for the learner (familiar to unfamiliar), not just where he begins.

Another situation favoring repeated activation of the same R-set is where the educator finds it difficult or time-consuming to get the appropriate Rs activated. For example, when complicated laboratory apparatus is needed as in demonstrating a chemical action, it is probably best to strike while the iron is hot and induce several repeated activations of the particular R-process, perhaps by verbal reiteration of the essential ideas of the demonstrated event. Any structure involving a large number of Rs (e.g., the complex relation of mountains, air currents, moisture, and temperature which cause a Chinook wind) is probably also best repeatedly activated while the learner "has the whole picture in view."

For most actual formal education tasks some compromise between repetition and alternation in terms of the above variables would probably work best. For learning a chain of ordered specifics, however, there seems to be no strong argument for doing anything but activating the specifics in the final order sought. For example, one would probably not try to get students to learn a verbal passage for recitation from memory by engaging them in a discussion of the passage because of the unpredictable direction of a discussion, whereas a group "chant" which fixed the order might work quite well. But even here there is doubt as to how long a chain should be undertaken as a whole, as opposed to dividing it into shorter sequences to be mastered separately (Lumsdaine, 1961).

Most school topics involve associating Rs varying widely in specificity (e.g., "democracy" vs. "a ballot") and associating the general with the specific is a substantial part of the objectives. Unfortunately most instructional methods tend to concentrate upon only a segment of the generality range. Verbal or symbolic presentations tend to slight the very specific Rs (e.g., concrete illustrations), more by custom than by necessity. Laboratories, field trips, and on-the-job training too rarely activate the more general Rs. What is needed are more balanced technique combinations which permit frequent rapid transition between very general and very specific Rs. Only by such contiguous activation can those associations be established. The RULEG system (Evans et al., 1960) for programing instruction is based on the importance of such general-specific transitions.

Another reason for facilitating the general-specific transitions is that usually the Rs having more extreme M-value (positive or negative) are usually either very general Rs, often having to do with self-concept (flunking out of school; being competent in some field), or else quite specific Rs activated by concrete situations (having to kill a frog; getting to build a motor or paint a picture). M-value is an essential variable in whether the student tries to learn, as discussed earlier, and instruction which ignores (does not activate) Rs having more extreme M-value, positive in particular, is missing an important source of leverage on the learner. Such motivators may be especially important in counteracting the effects of frequent failure.

Among those factors (relevant to selecting learning techniques) which are not intrinsic to the R-structure to be learned, perhaps the most important is the degree to which the learner himself controls the learning activity. For example, group learning methods (lectures, films, etc.) allow less learner control of the process than do most individual methods such as reading. It might be guessed that more learner control is generally desirable to the extent that the learner is willing and able to guide himself toward the learning goals. This suggests selecting techniques on the basis of characteristics of the learner population. But do some learning tasks benefit more than others from learner control of the process? If so, learner control is probably best for those tasks in which R-activation depends relatively more on other concurrent R-activity than on sensory input. Why? Because input can be controlled as well by the teacher as by the learner, whereas the learner's R-activity is more accessible to the learner than to the teacher. In learning ordered specifics, such as names for the bones of the human skeleton, the Rs are easily activated by sensory input provided by a teacher, book, or machine. If there is any advantage to giving the learner control over such a task it is for some other reason than convenience of activating appropriate Rs. Perhaps learner control could still serve as a motivational device or a way to optimize part-size and distribution of practice.

An important type of learning in which activation of appropriate Rs depends largely on other concurrent R-activity is that where problem-solving or reasoning is a learning objective in itself. The essence of problem-solving is thinking of questions which if answered might aid solution, seeking answers to such questions, asking new questions in light of new information, and so on toward solution of the problem. In realistic problem-solving what question

the learner will ask himself at any time is relatively variable and unpredictable. But sensory input is helpful only if relevant to the question being asked. In this sense activating appropriate Rs depends on what Rs (questions) are already active. Since the problem-solver knows best what Rs (questions) are active, he is in the best position to determine what kind of input, if any, is needed. Giving the learner considerable control over the situation thus seems essential to learning how to search for information and solve problems. If the teacher could know exactly what Rs are active in the learner he might use his own greater experience effectively to telescope the process, reduce wasted time, and help the learner ascertain important aspects of the problem. Asking the student to speak his thoughts as he works may help the teacher to do this, but it may also interfere with the problem-solving. Furthermore, helping the learner to structure the problem would seem to deprive him of a fundamental part of what is to be learned. A better technique may be to give the learner maximum control while the teacher's role is to make whatever information the learner seeks readily available, as tried experimentally by Suchman (1962).

Techniques which keep the learner trying, and maximize meaningfulness (R-activity), as noted, should be the most efficient for nearly all learning situations. Other variables germane to choice of learning method were noted in comparing formal and informal learning. For example, one way to increase meaningfulness and reduce satiation is to give the learner relevant input through as many different sense modes and different types of overt activity as he typically gets in informal learning situations. Another: Symbolic communication is especially suitable for rapid rearrangement of order of R-activations, as in weighing the pros and cons of a political issue, for example.

Another key issue in choice of learning method is that the method be geared to the type and frequency of evaluation of learning progress needed.

Evaluation of Learning Progress

Evaluation (or fairly accurate estimation) of the learner's relevant R-structure prior to instruction is a logical prerequisite to planning how to change the structure. Periodic evaluation during instruction is important in proportion to the unpredictability of the effects of instruction, and to the extent that the R-structure to be learned is hierarchical.

The purpose of evaluation is to adapt the instruction to the learner's progress, so as to minimize time spent teaching what is already well known or

teaching beyond the learner's grasp. Apart from wasting teaching resources, instruction not keyed closely to the learner's progress either bores or frustrates the learner, which in turn may degrade the whole process further. This seems to be the main advantage of individualized instruction over group methods. Individual differences among students are large, and by individual evaluation of each learner's progress the instruction can be kept appropriate to each learner.

With the exception of tutoring and programed self-instruction, the methods used in formal education seem to have evaluated the learner's progress too rarely. To the extent that evaluation does not disrupt learning, one might suppose the more feedback the better from the standpoint of learning efficiency. In some small-step self-instruction programs, however, it is possible that responses which give feedback to the learner (and to the programer) are required so frequently that they impair learning. This plus whatever cost accrues to evaluation procedures support the notion of an optimum frequency of evaluation rather than a maximum.

The degree of learner control of the learning process enters again as an important factor in that the learner's own judgment may be the best index of learning progress available. To orient students toward taking greater responsibility is a worthy objective in itself, and critical self-evaluation of step-by-step progress would seem to be a <u>sine qua non</u> for this aim. Although some methods and illustrations are more meaningful for nearly everyone than others, meaningfulness is a variable state of each learner's R-domain, and no one can assess the meaningfulness at a particular moment better than the learner himself. Possible disadvantages of self-evaluation are wish-fulfillment and other ulterior motives for distorted evaluation and the teacher's better grasp of the objectives and greater familiarity with other students' progress as a yardstick for evaluation.

What criteria of learning progress are available? For the learner, there is his own feeling of progress or mastery of the task. A criterion which is available to any evaluator is the tangible outcome of a decision which the learner is required to make. If the learner has acquired the proper R-structure, then this should tend to be reflected by his decision in accordance with the decision principle.

Decision outcomes may be the only available source of feedback to the outside evaluator for some kinds of learning, especially those in which the

associations formed cannot be expressed verbally. For example, in R-formation the learning consists of functional grouping of elements, which can seldom be described in words by the learner. Whereas in learning an R-structure involving associations among already existing Rs, for each of which the learner already has a word, the degree of learning may be adequately indicated by the learner's correct and incorrect verbal associations.

Because verbal report is such an efficient, economical source of feedback on learning progress, it has been heavily relied upon in formal education. So strong has this reliance been, that it seems the learning of R-structures which can be verbalized has been overemphasized to the neglect of relatively unverbalized R-structures. An illustration of this is the once conventional requirement that school children memorize short, verbal passages such as the Gettysburg Address, with the result that most achieved flawless recitation without one whit of understanding of the underlying issues. Substantial progress away from meaningless rote memorization seems to have been made in formal education in recent decades, but there remains a more subtle stress on learning only that which can be verbalized, perhaps not because the educator intends it, but because the student knows that feedback and grading criteria are almost wholly verbal in most subject matters. (Math is a major exception.)

Some of the more behavioristic autoinstructional programers also have relied too heavily on the overt (usually verbal) response, not because of expedience but because of confusion between objectives and criteria. As noted earlier, for learning tasks restricted to very specific Rs, R-structure may parallel external events, acts, or words so closely that failing to distinguish objectives from criteria may do no harm. But most learning objectives include more general Rs in a structure causally related to overt action but in no sense equivalent to it. In this case a valid criterion test may consist of a small representative sample from the infinite number of specific performances which could be used to infer properties of the learner's R-structure. However, if the planner then "teaches the test," that is, treats a particular criterion test as the <u>learning objective</u>, he almost certainly invalidates the test as a criterion for his original objectives. To illustrate, suppose the learning objective is forming the R "prime number." A good criterion test of having attained this objective might be for the learner to indicate which of the following numbers are prime: 11, 25, 29, 41, 63, 374. If the planner treats this test as his learning objective, he will find that the easiest way to reach

it is to have the learner memorize the answers (11, 29, 41 are prime; 25, 63, 374 are not), or some similar direct approach. But if he does this, the test is no longer a valid criterion for the original objective (to know what a prime number is). To those who counter that they would reach the general objectives by teaching all important specific criterion behaviors directly, I can only wish them the immortality they and their students will require in order to finish the job. In his lifetime a student can utilize only an infinitesimal fraction of the specific behaviors that might indicate acquisition of any important idea.

REFERENCES

Bartlett, F. Remembering. Cambridge: Cambridge Univ. Press, 1932.

Boring, E. A history of experimental psychology, New York: Appleton-Century-Crofts, 1950.

Bruner, J. and Goodman, C. Value and need as organizing factors in perception. J. abnorm. soc. Psychol., 1947, 42, 33-44.

Bruner, J., Goodnow, J. and Austin, G. A study of thinking. New York: Wiley, 1956.

Campbell, V. Self-direction and programed instruction for five different types of learning objectives. American Institute for Research, Palo Alto, 1963. Technical report No. AIR-D10-12/63-TR(b).

Carmichael, L., Hogan, H., and Walter, A. An experimental study of the effect of language on the reproduction of visually perceived form. J. exp. Psychol., 1932, 15, 73-86.

Cassirer, E. An essay on man. New Haven. Yale Univ. Press, 1944.

Dewey, J. School and Society. University of Chicago Press, 1899.

Edwards, W., Lindman, H., and Savage, L. Bayesian statistical inference for psychological research. Psychol. Rev., 1963, 70, 193-242.

Estes, W. Toward a statistical theory of learning. Psychol. Rev., 1950, 57, 94-107.

Evans, J., Homme, L., and Glaser, R. The ruleg (rule-example) system for the construction of learning. U.S.O.E. Coop. Res. Prog. report, University of Pittsburgh, 1960.

Gagne, R. The acquisition of knowledge. Psychol Rev., 1962, 69, 355-365.

Gibson, J. The reproduction of visually perceived forms, J. exp. Psychol., 1929, 12, 1-39.

Hallowell, A. Cultural factors in structuralization of perception. In Rohrer, J. and Sherif, M. (Eds.), Social psychology at the crossroads. New York: Harper, 1951.

Hebb, D. Organization of behavior. New York: Wiley, 1949.

Hebb, D. Heredity and environment in behavior. In Dulany, D., DeValois, R., Beardslee, D., and Winterbottom, M. (Eds.), Contributions to modern psychology. New York: Oxford Univ. Press, 1958.

Heidbreder, E. The attainment of concepts: II. The problem. J. gen. Psychol., 1946, 35, 191-223.

Hernandez-Peon, R., Scherrer, H., and Jouvet, M. Modification of electric activity in the cochlear nucleus during "attention" in unanesthetized cats. Science, 1956, 123, 331-332.

Hershberger, W. A. Experimentally induced positive feedback in locomotor behavior. Unpublished doctoral dissertation, University of Colorado, 1962.

Hull, C. Quantitative aspects of the evolution of concepts. Psychol. Monogr., 1920, No. 123.

Hull, C. Essentials of behavior. New Haven: Yale Univ. Press, 1951.

James, W. The principles of psychology. New York: Holt, 1890.

Kelly, G. The psychology of personal constructs. New York: Norton, 1955.

Koffka, K. Principles of gestalt psychology. New York: Harcourt, 1935.

Korzybski, A. The role of language in the perceptual processes. In Blake, R. and Ramsey, G. (Eds.), Perception: an approach to personality. New York: Ronald Press, 1951.

Lumsdaine, A. (Ed.) Student response in programed instruction. Washington, D. C.: NAS-NRC, 1961.

Maier, N. Frustration, the study of behavior without a goal. New York: McGraw-Hill, 1949.

McClelland, D., Atkinson, J., Clark, R., and Lowell, E. The achievement motive. New York: Appleton-Century-Crofts, 1953.

McDougall, W. Introduction to social psychology. London: Methuen, 1908.

Miller, G., Galanter, E., and Pribram, K. Plans and the structure of behavior. New York: Holt, 1960.

Mowrer, O. Learning theory and the neurotic paradox. Amer. J. Orthopsychiat., 1948, 18, 571-610.

Murray, H. Explorations in personality. New York: Oxford Univ. Press, 1938.

Osgood, C. Method and theory in experimental psychology. New York: Oxford Univ. Press, 1953.

Rotter, J. Social learning and clinical psychology, New York: Prentice-Hall, 1954.

Suchman, J. The elementary school training program in scientific inquiry. Univ. of Illinois, 1962. (Tech. Rep. under U. S. Office of Education Grant No. Title VII: 7-11-038).

Tolman, E. Purposive behavior in animals and men. New York: Century, 1932.

Watson, J. Behaviorism. Chicago: Univ. of Chicago Press, 1930.

Wiener, N. Cybernetics. New York: Wiley, 1948.