Psychological Reports, 1962, 10, 815-827.

AN ARTIFICIAL LANGUAGE AND MEMORY APPROACH TO CONCEPT ATTAINMENT[1]

WAYNE A. WICKELGREN AND DAVID H. COHEN

University of California, Berkeley

Formalization of the concept of information provided cognition psychology with a valuable tool for the analysis of problem solving. Early applications of information theory to psychology treated the human as a communication channel with only simple stochastic noise properties mediating between stimulus input and response output. Such models were of only limited usefulness, but the more general notion of the human problem solver as an information processor has recently led to the development of simulation models for a wide range of symbolic behavior (Newell, Shaw, & Simon, 1958; Feigenbaum, 1959; Feldman, 1959; Hovland & Hunt, 1960). This information processing analogy has far-reaching potential as a language for generating theories of human problem solving. However, any language is of limited predictive value until its concepts are operationally defined in terms of experimental designs. The major purpose of the following study was to devise unambiguous verbal report definitions of information processes in concept attainment. The second purpose was to investigate the variation of these information processes in different concept attainment situations and to correlate these processes with task success.

Information Processes

To illustrate some of the important information processes in conjunctive concept attainment, consider the following list of alternative responses to the presentation of: 23470259 is an example of the concept; followed by: 13470259 is not an example of the concept. Each of the eight places is a stimulus dimension with 10 possible values, 0 through 9. An example of a conjunctive concept is 2 in Place 1, 4 in Place 3, and 5 in Place 7. It will always be the case that in each relevant place one and only one value can be part of the concept.

1. Deterministic Inference.—(a) If an example and non-example differ in only one place, then infer that that place must be relevant to the concept, and the number in that place, *in the example,* is *part* of the concept. Applying this process to the information in the sample situation given above generates the following conclusion: place one is relevant, and 2 in place one is part of the concept.

(b) A variant of process 1.(a) which is valid but makes less use of the information is to infer that the number in the changed place, *in the non-example,*

[1]This paper was written in connection with a research study of creative thinking and auto-instructional methods under a grant from the Carnegie Corporation of New York to Richard S. Crutchfield at the Institute of Personality Assessment and Research, University of California, Berkeley.

is *not part* of the concept. Applied to the above situation, this yields the conclusion that 1 in place one is not part of the concept.

(c) Compare a presented number with a previously given example of the concept, and infer that all places with different values (changed places) are irrelevant, irrespective of whether this number is an example or non-example of the concept. When two examples are compared, this conclusion is correct, but when an example and non-example are compared, the conclusion is false. With the above information this process reaches the false conclusion that place one is irrelevant.

2. *Probabilistic Inference.*—(a) If an example and non-example differ in only one place, then infer that that place is *probably* relevant to the concept, and the number in that place, *in the example,* is *probably part* of the concept. In the above situation, this yields the conclusion that 2 in place one is probably part of the concept.

(b) This is the probabilistic analogue of 1.(b).

(c) This is the probabilistic analogue of 1.(c).

3. *Recording.*—When presented with an example or non-example that differs in at least one place from each previously presented example or non-example, record the instance for later use.

4. *Non-utilization.*—If the present instance is thought to contain no new useful information, ignore it.

Now consider some alternative information processes as applied to *two examples* of a concept: 23470259 is an example of the concept; followed by 23470251 is an example of the concept.

1. *Deterministic inferences.*—(a) If two examples differ in one or more places, then infer that these places are not part of the concept (irrelevant). In the above case this leads to the conclusion that place eight is irrelevant.

(b) If the first example has j in place i, and the second example has k in place i, then infer that j or k in place i is part of the concept. Applied to the above situation, this yields the conclusion (incorrect for conjunctive concepts) that 9 or 1 in place eight is part of the concept.

2. *Probabilistic Inferences.*—Direct probabilistic analogues of 1.(a) and 1.(b).

3. *Recording.*—(as before).

4. *Non-utilization.*—(as before).

Notice that the deterministic and probabilistic inferences discussed above apply only to changed places. Inferences about unchanged places will be referred to as hypotheses. The hypothesis information process might be stated as follows: Given two (or more) examples of the concept, infer that all the unchanged places are (probably) part of the concept. Applied to the above situation, this yields the conclusion that 2 in place one, 3 in place two, . . ., 5 in place seven are (probably) part of the concept. The number of examples that are necessary before a subject will apply the hypothesis process constitutes an important individual difference.

Artificial Language and Memory

Thus far we have listed some important processes for dealing with given information. Solution of a concept attainment problem, however, also requires memory processes to make the information from previous trials available on later trials. Such information includes both inferences from examples and non-examples and also the examples and non-examples themselves. In the case of computer simulation models, memory has a clear definition in terms of the memory subroutines of the simulation program. However, the human memory processes (storage, retention, and retrieval) are unmeasured and uncontrolled intervening variables in ordinary experimental situations.

This formulation of the problem led to the idea of constructing a visible artificial memory in which *S*s are instructed to record all information and inferences. *S*s are told to store information and record conclusions in grammatical English sentences using a small set of English words and numbers. The small vocabulary is more than sufficient to record all information and make all inferences that are necessary in order to gain the concept. Furthermore, the vocabulary consists of relatively simple and unambiguous terms whose exact meaning in the concept attainment task is easily established using a standard set of instructions for all *S*s. Provision is made for expressing a wide range of conclusions with different degrees of confidence. The language is far more flexible than necessary to solve the problem, providing a relatively unrestricted problem-solving situation sensitive to individual differences but yielding clear and unambiguous data. Each word in the vocabulary is written or abbreviated on small paper squares and can be arranged to form sentences in the artificial memory, a large checkerboard pattern of squares of the same size as the word squares. To simulate limited human memory an upper bound is placed on the number of memory board squares *S* can use at one time, but *S* can store any amount of information by making inferences and defining new words to stand for phrases of basic vocabulary words. Each presentation of an example or non-example constitutes a trial, and on each trial *S* is instructed to restructure his artificial memory to accommodate all new conclusions or information he feels he has obtained on that trial.

This study presents operational definitions of trial-to-trial information processing and uses this theoretical framework in an analysis of three problem areas in concept attainment. First, what is the effect on these processes of manipulating size of artificial memory? Second, how do information processes differ with regard to negative and positive instances (non-examples and examples)? Smoke (1933) and Hovland and Weiss (1953) have established the general conclusion that negative instances are not used as effectively as positive instances. Application of our framework leads to a more detailed analysis of the information processes responsible for this difference. Third, what information processes correlate with success or failure in this task?

METHOD

Procedure

S is presented with a series of eight place numbers, one at a time, and is told which are examples of a concept and which are non-examples. The conjunctive form of the concept is carefully explained to *S*. The concept is 6 in place 2, 3 in place 4, 8 in place 5, and 6 in place 8. His task is to identify this concept in as few trials as possible using the presented information. That is, he must discover how many places and what values in these places define the concept. *S* is instructed to guess whenever he thinks he has sufficient information; if he is incorrect, the task continues. After 18 trials, if *S* has not attained the concept, he is instructed to make a final guess, and the session is finished. The schedule of instances is determined in advance by *E* and is the same for all *S*s.

In an attempt to directly measure information storage strategy, we constrain *S* to record information using a restricted vocabulary and a restricted number of memory boxes. This restricted memory is in the form of a large checkerboard pattern of boxes in each of which *S* is permitted to insert a word (from the given vocabulary) representing a place, value, verb, predicate, connective, etc. Following the presentation of an instance, *S* may restructure immediate memory by the rearrangement and replacement of words in the artificial memory registers, allowing complete freedom for trial-to-trial change of both information and recording methods. This process continues until *S* is ready to guess, at which time the contents of the artificial memory board must be sufficient to generate the concept. If they are not sufficient, *S* is so informed at the time of his guess. Hence, *S* is instructed to record all necessary information in the visible artificial memory and make it internally consistent. In both memory conditions it is possible to record sufficient information to generate the concept by recording several basic vocabulary symbols into a new, informationally richer, symbol.

The following words constitute the basic vocabulary of the artificial language available to *S* in this experiment. Notice that this vocabulary consists of symbols representing ordinary English words and numbers.

0, 1, 2, 3, 4, 5, 6, 7, 8, and 9
P1, P2, P3, P4, P5, P6, P7, and P8 = Place 1, . . ., Place 8
AND, OR, NOT, IS, ARE, IN
C = Concept
UT = Untested
EC = Example of Concept

PC = Part of Concept
FC = Frequently in Example of Concept
PPC = Probably Part of Concept
PBC = Probably Concept

Concerning the permissible arrangements of these symbols in the artificial memory, *S* is constrained to record all information in grammatical English sentences. This is done to insure valid interpretation of the information stored on the memory board. The same rationale applies to restricting *S* to a subset of relatively unambiguous English vocabulary words. To further insure unambiguous interpretation of the language, *S* is presented with sample sentences and is then instructed to construct sentences of his own, for example, 8 IN P3 IS PC.

Because there is a limit on the number of memory registers, *S* is permitted to combine several basic vocabulary words into a single new word in order to make it possible to record all necessary information. Since information put on the memory board must form grammatical sentences, new words of 1, 2 or 3 symbols may be defined to represent only grammatical combinations of basic vocabulary words. For example, *S* might grow the vocabulary to include 8P3 = 8 IN P3 and IPC = IS PC, rewriting the above sample sentence, in only two memory registers, as 8P3 IPC. *S* may combine any number of words into a single new word, but each time he does this, he must write the definition on a card and give it to *E*. Since *S* may not consult previous definitions, he must remember what his new symbols mean.

Experimental Groups and Ss

Ss were randomly assigned to one of two independent groups differing in the size of the artificial memory. In one group *Ss* were permitted to use only 12 memory registers in which to record information and inferences and in the other group they were permitted to use 48 registers.

In both groups *Ss* were given the same alternating sequence of positive and negative instances. A within *Ss* comparison was necessary in order to study information processing for the two types of instances in conjunctive concept attainment, since it is impossible to solve the concept given either negative or positive instances alone, and no knowledge of the number of relevant places. *Ss* were 20 University of California (Berkeley) undergraduates recruited from an introductory psychology course.

Schedules of Instances

Table 1 presents the schedule of examples and non-examples for the first 10 trials. The remaining eight trials presented new instances but repeated the information given in the first 10 trials.

TABLE 1

SCHEDULE OF INSTANCES FOR TRIALS 1 TO 10*

Trial	Instance	EC (+) NOT EC (−)	Changed place	Optimal inference
1	76838916	+	Focus card†	Record
2	76837916	−	P5	P5 IS PC
3	26838916	+	P1	P1 IS NOT PC
4	74838916	−	P2	P2 IS PC
5	76838926	+	P7	P7 IS NOT PC
6	73818916	−	P2, P4	Ignore (no information)
7	56838906	+	P1, P7	Ignore (no information)
8	76828916	−	P4	P4 IS PC
9	76038816	+	P3, P6	P3 AND P6 ARE NOT PC
10	76838913	−	P8	P8 IS PC

Concept is 6 in P2, 3 in P4, 8 in P5, 6 in P8.

*The concept is logically derivable after 10 trials.
†The first positive instance is called a Focus card to indicate that *S*s use this example as a standard for making inferences from subsequent instances.

Definition of Analytic Dimensions

Informational analysis of the stimulus situation.—The schedule of instances is specifically designed so that the optimal inference (see Col. 5 in Table 1) can be made from a comparison of a given instance with the positive focus presented on Trial 1. These are divided into two classes, examples and non-examples of the concept (positive and negative cards). Either type of instance may be further partitioned into places whose values are different from the focus card and places whose values are the same, changed and unchanged places. In terms of this analysis it should be pointed out that the informational unit is the place, and logically to solve the concept *S* must obtain two chunks[2] of information about each place, "Is it part of the concept, and, if so, what is its correct value?"

Informational analysis of the response situation.—We classify the sentences *S* writes on the memory board into four types of information processes that can be applied to a given informational unit: deterministic inference, probabilistic inference, recording, and non-utilization. Deterministic inferences are of the form Pi (place *i*) IS (NOT) PC. Probabilistic inferences are of the form Pi IS (NOT) PPC (or FC). Recording sentences on the memory board are of two types: 7 6 8 3 8 9 1 6 IS (NOT) EC, and 7 IN P1 IS IN (NOT) EC. Finally, when an individual neither adds new information nor restructures previous information we call the process non-utilization. Crosscutting the first three information processes is the rather obvious dimension of inference sign. PC, PPC, FC, and EC all have positive inference sign, while NOT PC,

[2]This term is borrowed from Miller (1956).

NOT PPC, NOT FC, and NOT EC have negative inference sign. At some points in the data analysis it is useful to consider PC and PPC as equivalent, and in these situations we denote either PC or PPC as (PC).

Informational analysis of the stimulus-response situation.—Table 2 presents our conception of the different primary stimulus situations from which *S* may make an inference, defines the mode of a deterministic or probabilistic inference in each situation (immediate inference, deferred inference, or hypothesis), and indicates the correct inference sign in each case. It should be noted that an inference based on previously recorded changed place information (deferred inference) is considered to have the same logical status as an inference based on presently occurring changed place information (immediate inference). Both are considered logical inferences as opposed to hypotheses which are based on only unchanged place information. Finally, it should be understood that the distinction between logical inference and hypothesis is *not* based on the type of inference, deterministic or probabilistic. However, there is an obvious implication that *S*s should be less confident of hypotheses than logical inferences, and the data strongly validate this implication since 43 of 49 hypotheses are probabilistic inferences (sign test, $p < .001$, one-tailed), and 85 of 117 logical inferences are deterministic (sign test, $p < .002$, one-tailed).

TABLE 2

DEFINITION OF S-R INFERENCE MODES AND THE CORRECT INFERENCE SIGN FOR EACH STIMULUS SITUATION

History of the place	Positive instance (Trials 3, 5, 9)		Negative instance (Trials 2, 4, 8, 10)	
	Changed place (P_i)	Unchanged place (P_j)	Changed place (P_i)	Unchanged place (P_j)
Place previously changed on + instance, and *S* recorded it.	*	Deferred inference: P_j IS NOT (PC)	*	Deferred inference: P_j IS NOT (PC)
Place previously changed on — instance, and *S* recorded it.	*	Deferred inference: P_j IS (PC)	*	Deferred inference: P_j IS (PC)
Place never previously changed, or it changed but *S* did not record.	Immediate inference: P_i IS NOT (PC)	Hypothesis: P_j IS (PC)	Immediate inference: P_i IS (PC)	Hypothesis: P_j IS NOT (PC)

*Indicates that this situation never occurs.

RESULTS AND DISCUSSION

Effect of Memory Size on Information Processing

Inferences about changed places.—Table 3 compares the two memory size groups with respect to the total number of inferences of each type in the first

10 trials. Since the problem is logically solvable in 10 trials and some *S*s do in fact attain the concept in the minimum number of trials, count data on more than 10 trials would be confounded by successive elimination of *S*s who gain the concept. The results are striking and clear-cut. *S*s in the 12-condition ignore more information, make more probabilistic inferences, make as many deterministic inferences, and record less than do *S*s in the 48-condition. These results suggest the following interpretation. An *S* who has mastered the optimal inference process to handle a given type of instance (positive or negative) makes a deterministic inference regardless of memory size. In the absence of such mastery the safest procedure is direct recording of the instance for interpretation on a later trial, and in the 48-condition this is precisely what happens. However, in the 12-condition restricted memory size precludes extensive recording, resulting in less confident inference (probabilistic) or neglect of information (non-utilization).

TABLE 3

NUMBER OF CHANGED PLACE INFERENCES FOR TRIALS 1 TO 10

Type of inference	Memory size		*p**
	12	48	
Deterministic	53	58	ns
Probabilistic	31	16	.04
Recording	18	46	.002
Non-utilization	18	7	.04
Totals	120	127†	

*Two-tailed sign test.
†In the 48 condition there are 7 cases where an *S* both recorded and made a deterministic inference.

Inferences about unchanged places.—Table 4 compares the two memory size groups with respect to the first trial on which an inference about an unchanged place occurs and the total number of inferences about unchanged places in the first ten trials. *S*s in the 48-condition make significantly more of these inferences significantly sooner. The interpretation of this finding is straight-forward in light of the preceding analysis. Since *S*s in the 48-condition record more instances from previous trials, they have more available data on which to base both deferred inferences and hypotheses. *S*s in the 12-condition more often made the inference previously or failed to record the information in any form. It appears to be a common tendency to generate an hypothesis from two or more instances, and this tendency is accentuated by the visible presence of the instances recorded on the memory board.

Trials to solution.—In the 12-condition, 4 *S*s attain the concept in 10 to 13 trials (early learners); 4 *S*s succeed in 17 to 18 trials (late learners); and 2 *S*s

TABLE 4

UNCHANGED PLACE INFERENCES FOR TRIALS 1 TO 10

Memory size	No. of *Ss* making such inferences	Mean trial of first such inference	No. of deferred inferences	No. of hypotheses	Total no. of unchanged place inferences
12	7	6.3	1	21	22
48	6	3.8	8	28	36
p	ns*	.05†	.02*	ns*	.05*

*Two-tailed sign test. †Two-tailed t test.

fail to solve the problem in 18 trials (failures). In the 48-condition there are 4 early learners and 6 failures. The most significant difference between the two groups is the greater frequency of success within the 18-trial limit in the 12-condition (8/10 vs. 4/10; Fisher's exact test, $p < .08$, two-tailed). There are the same number of *early* learners in both groups, but an equally large number of *late* learners present in the 12-, but not the 48-, condition is responsible for the overall difference. However, $p < .08$ is far from compelling, and a detailed examination of the trial-to-trial data suggested no single explanation for the presence of late learners in the 12-condition. Trials to solution is a rather ambiguous measure of the problem-solving process in that a number of more basic factors jointly determine it. For example, some of the basic factors we have conceptualized are learning not to use positive information incorrectly, learning not to use negative information incorrectly, learning to use positive information correctly, learning to use negative information correctly, recording information, preference for probabilistic or deterministic inferences, and tendency to form hypotheses about unchanged places. One possible explanation is that limited memory precludes extensive recording of instances and thereby induces earlier inference behavior.

Information Processing of Negative and Positive Instances

Inferences about changed places.—A review of Table 1 indicates that the negative instances on Trials 2, 4, 8, and 10 provide the information that Places 2, 4, 5, and 8 are part of the concept and that the positive instances on Trials 3, 5, and 9 provide the information that Places 1, 3, 6, and 7 are not part of the concept. Table 5 compares the negative and positive instances with respect to the type and correctness of inferences about changed places. When presented with negative instances, *Ss* make far fewer deterministic inferences, and those they make are more often incorrect than correct. It is particularly striking to note that of the 17 incorrect inferences made in the experiment, all 17 were made in response to negative instances. Positive instances were never used incorrectly. Instead of making deterministic inferences from negative instances *S*s either neglected the information, recorded the instance, or made probabilistic inferences. Particular attention should be drawn to the following two results: (1) all 14 cases of non-utilization occur with negative instances; (2) the optimal inference process (deterministic correct), which is dominant for positive instances, is very infrequent for negative instances. Classifying all *S*s as to their use of negative and positive instances, we obtain the further results, (1) that 8 of the 20 *S*s used negative instances incorrectly at some point in the trial series while none used positive instances incorrectly (Fisher's exact test, $p < .01$, one-tailed), and (2) that only 1 of the 20 *S*s always used negative instances correctly, while 15 always used positive instances correctly ($\chi^2 = 176$, $df = 1$, $p < .001$, one-tailed). Recording and non-utilization are classified as neither correct nor incorrect.

TABLE 5

NUMBER OF CHANGED PLACE INFERENCES FOR TRIALS 3, 5, 9* VS. 2, 4, 8, 10

Type of inference	Positive instances	Negative instances	*p*
Deterministic	60	25	.001**
Correct	60	10	.001†
Incorrect	0	15	
Probabilistic	9	23	.01**
Correct	9	21	ns‡
Incorrect	0	2	
Recording	11	18	.10**
Non-utilization	0	14	.001**

*Trial 9 is counted twice since two places vary (see Table 1).
**One-tailed sign test.
†One-tailed χ^2.
‡One-tailed Fisher's Exact Test.

Inferences about unchanged places.—Table 6 compares positive and negative instances with respect to the number of hypotheses having positive or negative inference sign. Recall from Table 2 that, for hypotheses, positive inference sign is correct for positive instances and negative inference sign for negative instances. Although the *p* is only .08, Table 6 does indicate a tendency to formulate hypotheses with correct inference sign for the trial on which they are generated. But the most striking results are that 45 of the 49 hypotheses have positive inference sign (sign test, $p < .001$, one-tailed), and 35 of the 49 hypotheses are generated on positive instance trials (sign test, $p < .002$, one-tailed). Drawing from the previously discussed finding that negative instances are used less effectively than positive with respect to changed place inferences, we simply generalize this finding to conclude that negative instances are also used less effectively with respect to hypotheses.

TABLE 6

HYPOTHESES FOR NEGATIVE AND POSITIVE INSTANCES

Inference sign	Positive instances	Negative instances	Total
P_t IS (PC)	34	11	45
P_t IS NOT (PC)	1	3	4
Total	35	14	49

Note.—One-tailed Fisher's Exact Test, $p < .08$.

Experiments by Smoke (1933) and Hovland and Weiss (1953) have established the differential efficacy of positive and negative instances for the traditional success measures in the concept attainment task. The extensive trial-to-trial measures introduced by the artificial language and memory paradigm provide the data for a more molecular analysis of the information process responsible for this difference. *A priori, any* of the significant results reported in Tables 5 and 6 might have accounted for the superiority of positive instances; our experiment demonstrates that *all* play a part for different *S*s on different trials.

Information Processes Associated with Success and Failure

Inferences about change places.—The 20 *S*s were divided into three groups: 8 *S*s who gained the concept early (10 to 13 trials), 4 *S*s who gained the concept late (17 to 18 trials), and 8 *S*s who failed to solve the problem in 18 trials. The three groups were compared with respect to the number of correct, incorrect, and total deterministic inferences; correct, incorrect and total probabilistic inferences; recordings, and non-utilizations. These totals were for only the first 10 trials to avoid effects of systematic elimination of *S*s who solved the concept. There were no significant differences between the solve-early and solve-late groups. But solvers (early and late) made fewer incorrect deterministic inferences in the first 10 trials than the failures (sign test, $p < .01$, $p < .05$, one-tailed). No other comparisons reached significance. There is nothing surprising about the direction of our only positive result. What is surprising is that, of many possible relationships between task success and information processing of changed places, this is the only one that reaches significance.

Consider the following four dimensions of effective information processing: learning not to use positive instances incorrectly, learning to use positive instances correctly, learning not to use negative instances incorrectly, and learning to use negative instances correctly. Learning *not to infer incorrectly* from a class of instances is not equivalent to learning to make *correct* inferences from this class of instances. *S* may record or ignore the information. On each of these dimensions an *S* is classified as having learned or not learned prior to Trial 18. The criterion for learning is consistent performance on a run of at least two trials terminating in Trial 18. For dimensions one and three this means never making incorrect inferences about changed places after some trial, and for dimensions two and four this means always making correct inferences about changed places after some trial. A comparison of the early, late and fail-to-solve groups demonstrated no differences with respect to effective information processing of positive instances (dimensions one and two). However, of the 12 *S*s who solved the problem all learned not to make incorrect inferences from negative cards, while only 4 of the 8 who failed to solve the problem learned on this dimension (Fisher's exact test, $p < .02$, one-tailed). Concerning correct use of negative instances, 4 of 8 learned in the early group, 0 of 4 in the late group, and 1 of 8 in the fail group. There is some suggestion of a relationship here, but no comparison reaches significance. The general conclusion to be drawn from both the above analyses of changed place inferences is that the primary correlate of successful concept attainment is learning not to use negative instances incorrectly. Although correct use of negative instances appears to be unnecessary for success, it is facilitative.

On the basis of this experiment we cannot infer a causal relationship between learning not to use negative instances incorrectly and successful concept attainment. However, if such a causal relation exists, the following testable prediction is generated: more *S*s should solve the problem in 18 trials if *only* positive instances are presented than if *both* positive and negative instances are presented, contrary to what one would predict on the basis of greater information in the mixed series.

Inferences about unchanged places.—Both hypotheses and deferred inferences in the first 10 trials are unexpectedly related to task success. The mean number of hypotheses per *S* for the early, late, and fail groups are respectively 1.2, 2.5, and 3.8. The early and fail group difference is significant at the .02 level ($t = 2.76$, $df = 10$, two-tailed), and the trend is consistent. The mean numbers of deferred inferences (all correct) in the early, late, and fail groups are respectively .88, .25, and .12. The early and fail group difference only reaches the .10 level ($t = 1.85$, $df = 10$, two-tailed) but the trend is again consistent. The data indicate that formulation of hypotheses is negatively correlated with concept attainment, while the frequency of deferred inferences is positively correlated.

SUMMARY

Information processing in a concept attainment task was operationally defined in terms of the manner in which *S*s stored information in an artificial memory using an artificial language. The usefulness of this approach was

demonstrated by using the information processing data to analyze three problems in concept attainment: the effects of memory size, the differences between use of examples and non-examples, and the correlates of task success. Two experimental groups performed the same task, but one group was permitted to use 48 memory registers while the other could use only 12 registers. Smaller memory size resulted in more probabilistic inferences, more non-utilization of information, fewer hypotheses, fewer deferred inferences, and greater likelihood of successful attainment. When presented with non-examples as opposed to examples, *S*s made fewer deterministic inferences and more incorrect inferences, probabilistic inferences, recordings, and non-utilizations, although the amount of information was equal. Interestingly enough, the most important determinant of task success appears to be learning not to use non-examples incorrectly.

REFERENCES

BRUNER, J. S., GOODNOW, J. J., & AUSTIN, G. A. *A study of thinking*. New York: Wiley, 1956.

CHOMSKY, N. *Syntactic structures*. 'S-Gravenhage: Mouton, 1957.

FEIGENBAUM, E. A. *An information processing theory of verbal learning*. Santa Monica, Calif.: Rand, 1959. (Rand Corporation Paper, P-1817)

FELDMAN, J. An analysis of predictive behavior in a two-choice situation. Unpublished doctoral thesis, Graduate School of Business Administration, Carnegie Institute of Technology, 1959.

HOVLAND, C. I., & HUNT, E. B. Computer simulation of concept attainment. *Behav. Sci.*, 1960, 5, 265-267.

HOVLAND, C. I., & WEISS, W. I. Transmission of information concerning concepts through positive and negative instances. *J. exp. Psychol.*, 1953, 45, 175-182.

KENDLER, T. S. Concept formation. *Annu. Rev. Psychol.*, 1961, 12, 447-472.

MILLER, G. A. The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychol. Rev.*, 1956, 63, 81-97.

MILLER, G. A., GALANTER, E., & PRIBRAM, K. H. *Plans and the structure of behavior*. New York: Holt, 1960.

NEWELL, A., SHAW, J. C., & SIMON, H. A. Elements of a theory of human problem solving. *Psychol. Rev.*, 1958, 65, 151-166.

SMOKE, K. L. Negative instances in concept learning. *J. exp. Psychol.*, 1933, 16, 583-588.

Accepted April 30, 1962.