3
The Psychology of Thinking: Embedding Artifice in Nature

We watch an ant make his laborious way across a wind- and wave-molded beach. He moves ahead, angles to the right to ease his climb up a steep dune let, detours around a pebble, stops for a moment to exchange information with a compatriot. Thus he makes his weaving, halting way back to his home. So as not to anthropomorphize about his purposes, I sketch the path on a piece of paper. It is a sequence of irregular, angular segments not quite a random walk, for it has an underlying sense of direction, of aiming toward a goal.

I show the unlabelled sketch to a friend. Whose path is it? An expert skier, perhaps, slaloming down a steep and somewhat rocky slope. Or a sloop, beating upwind in a channel dotted with islands or shoals. Perhaps it is a path in a more abstract space: the course of search of a student seeking the proof of a theorem in geometry.

Whoever made the path, and in whatever space, why is it not straight; why does it not aim directly from its starting point to its goal? In the case of the ant (and for that matter the others) we know the answer. He has a general sense of where home lies, but he cannot foresee all the obstacles between. He must adapt his course repeatedly to the difficulties he encounters and often detour uncrossable barriers. His horizons are very close, so that he deals with each obstacle as he comes to it; he probes for ways around or over it, without much thought for future obstacles. It is easy to trap him into deep detours.

Viewed as a geometric figure, the ant's path is irregular, complex, hard to describe. But its complexity is really a complexity in the surface of the beach, not a complexity in the ant. On that same beach another small creature with a home at the same place as the ant might well follow a very similar path.

Many years ago Grey Walter built an electromechanical "turtle," having only tactile sense of its environment but capable of exploring a room, and periodically seeking its nest to recharge its batteries.[1] Today, robots with modest visual sensory capabilities roam about in a number of artificial intelligence alaboratories.[2] Suppose we undertook to design an automaton with the approximate dimensions of an ant, similar means of locomotion, and comparable sensory acuity. Suppose we provided it with a few simple adaptive capabilities: when faced with a steep slope, try climbing it obliquely; when faced with an insuperable obstacle, try detouring; and so on. (Except for problems of miniaturization of components, the present state of the art would readily support such a design.) How different would its behavior be from the behavior of the ant?

These speculations suggest a hypothesis, one that could as well have been derived as corollary from our previous discussion of artificial objects:

An ant, viewed as a behaving system, is quite simple. The apparent complexity of its behavior over time is largely a reflection of the complexity of the environment in which it finds itself.

We may find this hypothesis initially plausible or implausible. It is an empirical hypothesis, to be tested by seeing whether attributing quite simple properties to the ant's adaptive system will permit us to account for its behavior in the given or similar environments. For the reasons developed at length in the first chapter, the truth or falsity of the hypothesis should be independent of whether ants, viewed more microscopically, are simple or complex systems. At the level of cells or molecules ants are demonstrably complex, but these microscopic details of the inner environment may be largely irrelevant to the ant's behavior in relation to the outer

1. W. Grey Walter, "An Imitation of Life," *Scientific American, 185*(1950):42.

2. See, for example, R. Brooks, "A Robust-layered Control System for a Mobile Robot," *IEEE Journal of Robotics and Automation, RA-2*(1986):14 23. And a motor vehicle, NAVLAB, steered itself in the Summer of 1995 on public highways from Washington, D.C., to San Diego, California, and has also demonstrated strong capabilities for off-road navigation.

environment. That is why an automaton, though completely different at the microscopic level, might nevertheless simulate the ant's gross behavior.

In this chapter I should like to explore this hypothesis but with the word "human being" substituted for "ant."

Human beings, viewed as behaving systems, are quite simple. The apparent complexity of our behavior over time is largely a reflection of the complexity of the environment in which we find ourselves.

Now I should like to hedge my bets a little. Instead of trying to consider the "whole person," fully equipped with glands and viscera, I should like to limit the discussion to Homo sapiens, "thinking person." I myself believe that the hypothesis holds even for the whole person, but it may be more prudent to divide the difficulties at the outset, and analyze only cognition rather than behavior in general.[3]

I should also like to hedge my bets in a second way, for a human being can store away in memory a great furniture of information that can be evoked by appropriate stimuli. Hence I would like to view this information-packed memory less as part of the organism than as part of the environment to which it adapts.

The reasons for assigning some a priori probability to the hypothesis of simplicity have already been set forth in the last two chapters. A thinking human being is an adaptive system; men's goals define the interface between their inner and outer environments, including in the latter their memory stores. To the extent that they are effectively adaptive, their behavior will reflect characteristics largely of the outer environment (in the light of their goals) and will reveal only a few limiting properties of the inner environment of the physiological machinery that enables a person to think.

3. I have sketched an extension of this hypothesis to phenomena of emotion and motivation in "Motivational and Emotional Controls of Cognition," *Psychological Review*, *74*(1967):29 39, and to certain aspects of perception in "An Information-Processing Explanation of Some Perceptual Phenomena," *British Journal of Psychology*, *58*(1967):1 12. Both papers are reprinted in my *Models of Thought*, vol. 1 (1979), chapters 1.3 and 6.1. The discussion of these issues is continued in "Bottleneck of Attention: Connecting Thought with Motivation," in W. D. Spaulding (ed.), *Integrative Views of Motivation, Cognition and Emotion*. Lincoln, NE: University of Nebraska Press, 1994.

I do not intend to repeat this theoretical argument at length, but rather I want to seek empirical verification for it in the realm of human thought processes. Specifically I should like to point to evidence that there are only a few "intrinsic" characteristics of the inner environment of thinking beings that limit the adaptation of thought to the shape of the problem environment. All else in thinking and problem-solving behavior is artificial is learned and is subject to improvement through the invention of improved designs and their storage in memory.

Psychology As a Science of the Artificial

Problem solving is often described as a search through a vast maze of possibilities, a maze that describes the environment. Successful problem solving involves searching the maze selectively and reducing it to manageable proportions. Let us take, by way of specific example, a puzzle of the kind known as crypt arithmetic problems:[4]

DONALD
+ GERALD　　　*D = 5*
ROBERT

The task is to replace the letters in this array by numerals, from zero through nine, so that all instances of the same letter are replaced by the same numeral, different letters are replaced by different numerals, and the resulting numerical array is a correctly worked out problem in arithmetic. As an additional hint for this particular problem, the letter *D* is to be replaced by the numeral 5.

One way of viewing this task is to consider all the 10!, ten factorial, ways in which ten numerals can be assigned to ten letters. The number 10! is not so large as to strike awe in the heart of a modern computer; it is only a little more than 3 million (3,628,800, to be exact). A program designed to generate all possible assignments systematically, and requiring

4. The crypt arithmetic task was first used for research on problem solving by F. Bartlett in his *Thinking* (New York: Basic Books, 1958). In the present account I have drawn on his work and on my research with Allen Newell reported in our book, *Human Problem Solving* (Englewood Cliffs, N.J.: Prentice-Hall, 1972), chapters 8 10.

a tenth of a second to generate and test each, would require at most about ten hours to do the job. (With the cue *D* = 5, only an hour would be needed.) I haven't written the program, but a tenth of a second is far longer than a computer would need to examine each possibility.

There is no evidence that a human being could do this. It might take a man as long as a minute to generate and test each assignment, and he would have great difficulty in keeping track of where he was and what assignments he had already tried. He could use paper and pencil to assist him on the latter score, but that would slow him down even more. The task, performed in this way, might call for several man-years of work I assume a forty-hour week.

Notice that in excluding exhaustive, systematic search as a possible way for a human to solve the problem, we are making only very gross assumptions about human capabilities. We are assuming that simple arithmetic operations take times that are of the order of seconds, that the operations are essentially executed serially, rather than in parallel, and that large amounts of memory are not available in which new information can be stored at split-second speeds. These assumptions say something, but not very much, about the physiology of the human central nervous system. For example, modifying the brain by incorporating in it a new subsystem with all the properties of a desk calculator would be a quite remarkable feat of brain surgery or evolution. But even such a radical alteration would change the relevant assumptions only slightly for purposes of explaining or predicting behavior in this problem environment.

Human beings do frequently solve the *DONALD* + *GERALD* = *ROBERT* problem. How do they do it? What are the alternative ways of representing the environment and conducting the search?

Search Strategies

One way to cut down the search drastically is to make the assignments systematically, as before, but to assign numerals to the letters one by one so that inconsistencies can be detected before an assignment is complete, and hence whole classes of possible assignments can be ruled out at one step. Let me illustrate how this works.

Suppose we start from the right, trying assignments successively for the letters *D*, *T*, *L*, *R*, *A*, *E*, *N*, *B*, *O*, and *G*, and substituting numerals in the order 1, 2, 3, 4, 5, 6, 7, 8, 9, 0. We already know that $D = 5$, so we strike 5 from the list of available numerals. We now try $T = 1$. Checking in the right-hand column, we detect a contradiction, for $D + D = T + c$, where c is 10 or 0. Hence, since ($D = 5$, $T = 1$) is not feasible, we can rule out all the remaining 8! assignments of the eight remaining numerals to the eight remaining letters. In the same way all possible assignments for T, except $T = 0$, can be ruled out without considering the assignments for the remaining letters.

The scheme can be improved further by the expedient of calculating directly, by addition, what assignment should be made to the sum of a column whenever the two addends are known. With this improvement we shall not need to search for the assignment for T, for $T = 0$ can be inferred directly from $D = 5$. Using this scheme, the *DONALD* + *GERALD* = *ROBERT* problem can be solved quite readily, with paper and pencil. Ten minutes should suffice. Figure 3 shows the search tree, in slightly simplified form. Each branch is carried to the point where a contradiction is detected. For example, after the assignments ($D = 5$, $T = 0$), the assignment $L = 1$ leads to the inference $R = 3$, which yields a contradiction since from the left-hand column of the problem array $R = 3$ would imply that G is negative.

Figure 3 is oversimplified in one respect. Each of the branches that terminates with a contradiction after assignment of a value to E should actually be branched one step further. For the contradiction in these cases arises from observing that no assignment for the letter O is now consistent. In each case four assignments must be examined to determine this. Thus the full search tree would have 68 branches still a far cry from 10! or even 9!.

An enormous space has been cut down to a quite small space by some relatively small departures from systematic, exhaustive search. It must be confessed that the departures are not all as simple as I have made them appear. One step in the proposed scheme requires finding the contradictions implied by an assignment. This means of course the "relatively direct" contradictions, for if we had a rapid process capable of detecting *all* inconsistent implications, direct or indirect, it would find the problem solution almost at once. In this problem any set of assignments other than the single correct one implies a contradiction.

$D=5$	$T=0$	$L=1$	$R=3$	$G<0$	□			
		$L=2$	$R=5$	$G=0$	□			
		$L=3$	$R=7$	$A=1$	$E=2$	□		
				$A=2$	$E=4$	□		
				$A=4$	$E=8$	□		
				$A=6$	$E=2$	□		
				$A=8$	$E=6$	□		
				$A=9$	$E=8$	□		
		$L=4$	$R=9$	$A=1$	$E=2$	□		
				$A=2$	$E=4$	□		
				$A=3$	$E=6$	□		
				$A=6$	$E=2$	□		
				$A=7$	$E=4$	□		
				$A=8$	$E=6$	□		
		$L=6$	$R=3$	$G<0$	□			
		$L=7$	$R=5$	□				
		$L=8$	$R=7$	$A=1$	$E=3$	□		
				$A=2$	$E=5$	□		
				$A=3$	$E=7$	□		
				$A=4$	$E=9$	$N=1$	$B=8$	□
						$N=2$	$B=9$	□
						$N=3$	$G=0$	□
						$N=6$	$O=2$	$G=1$

Figure 3
Possible search tree for DONALD + GERALD = ROBERT

What is meant by searching for direct contradictions is something like this: after a new assignment has been made, those columns are examined where the newly substituted letter occurs. Each such column is solved, if possible, for a still-unassigned letter, and the solution checked to see whether this numeral remains unassigned. If not, there is a contradiction.

In place of brute-force search we have now substituted a combined system of search and "reason." Can we carry this process further; can we eliminate substantially all trial-and-error search from the solution method? It turns out that we can for this problem, although not for all crypt arithmetic problems.[5]

The basic idea that permits us to eliminate most trial-and-error search in solving the problem before us is to depart from the systematic right-to-left assignment of numerals. Instead we search for columns of the

5. For example, the method to be described does not eliminate as much search from the crypt arithmetic problem *CROSS* + *ROADS* = *DANGER*.

problem array that are sufficiently determinate to allow us to make new assignments, or at least new inferences about the properties of assignments.

Let me go through the process briefly. From $D = 5$, we immediately infer $T = 0$, as before. We also infer that 1 is carried into the second column, hence that $R = 2L + 1$ is odd. On the extreme left, from $D = 5$, we infer that R is greater than 5 (for $R = 5 + G$). Putting together these two inferences, we have $R = 7$ or $R = 9$, but we do not try these assignments. Now we discover that the second column from the left has the peculiar structure $O + E = O$a number plus another equals itself (apart from what is carried into or out of the column). Mathematical knowledge, or experiment, tells us that this can be true only if $E = 0$ or $E = 9$. Since we already have $T = 0$, it follows that $E = 9$. This eliminates one of the alternatives for R, so $R = 7$.

Since $E = 9$, it follows that $A = 4$, and there must be a one carried into the third column from the right; hence $2L + 1 = 17$, or $L = 8$. All that remains now is to assign 1, 2, 3, and 6 in some order to N, B, O, and G. We get $G = 1$ by observing that for any assignment of O there is a number carried into the leftmost column. We are now left with only $3! = 6$ possibilities, which we may be willing to eliminate by trial and error: $N = 6$, $B = 3$, and therefore $O = 2$.

We have traced a solution path through the problem maze on three different assumptions about the search strategy. The more sophisticated, in a certain sense, that strategy became, the less search was required. But it is important to notice that, once the strategy was selected, the course of the search depended only on the structure of the problem, not on any characteristics of the problem solver. By watching a person, or an automaton, perform in this problem environment, what could we learn about him? We might well be able to infer what strategy was followed. By the mistakes made, and the success in recovering from them, we might be able to detect certain limits of the capacity or accuracy of the individual's memory and elementary processes. We might learn something about the speed of these processes. Under favorable circumstances, we might be able to learn which among the thinkable strategies the individual was able actually to acquire and under what circumstances likely to acquire them. We should certainly be unlikely to learn anything specific about the neurological characteristics of the central nervous system, nor would the specifics of that system be relevant to his behavior, beyond placing bounds on the possible.

The Limits on Performance

Let us undertake to state in positive fashion just what we think these bounds and limits are, as revealed by behavior in problem situations like this one. In doing so, we shall draw upon both experimental evidence and evidence derived from computer simulations of human performance. The evidence refers to a variety of cognitive tasks, ranging from relatively complex ones (crypt arithmetic, chess, theorem proving), through an intermediate one (concept attainment), to simple ones that have been favorites of the psychological laboratory (rote verbal learning, short-term memory span). It is important that with this great variety of performance only a small number of limits on the adaptability of the inner system reveal themselves and these are essentially the same limits over all the tasks. Thus the statement of what these limits are purports to provide a single, consistent explanation of human performance over this whole range of heterogeneous task environments.

Limits on Speed of Concept Attainment

Extensive psychological research has been carried out on concept attainment within the following general paradigm.[6] The stimuli are a set of cards bearing simple geometric designs that vary, from card to card, along a number of dimensions: shape (square, triangle, circle), color, size, position of figure on card, and so on. A "concept" is defined extensionally by some set of cards the cards that are instances of that concept. The concept is defined intensionally by a property that all the instances have in common but that is not possessed by any of the remaining cards.

[6] This account of concept attainment is based on the paper with my late colleague Lee Gregg, "Process Models and Stochastic Theories of Simple Concept Formation," *Journal of Mathematical Psychology, 4*(June 1967):246 276. See also A. Newell and H. A. Simon, "Overview: Memory and Process in Concept Formation," chapter 11 in B. Kleinmuntz (ed.), *Concepts and the Structure of Memory* (New York: Wiley, 1967), pp. 241 262. The former paper is reprinted in *Models of Thought*, vol. 1, chapter 5.4.

Examples of concepts are "yellow" or "square" (simple concepts), "green triangle" or "large, red" (conjunctive concepts), "small or yellow" (disjunctive concept), and so on.

In our discussion here I shall refer to experiments using an N-dimensional stimulus, with two possible values on each dimension, and with a single relevant dimension (simple concepts). On each trial an instance (positive or negative) is presented to the subject, who responds "Positive" or "Negative" and is reinforced by "Right" or "Wrong," as the case may be. In typical experiments of this kind, the subject's behavior is reported in terms of number of trials or number of erroneous responses before an error-free performance is attained. Some, but not all, experiments ask the subject also to report periodically the intensional concept (if any) being used as a basis for the responses.

The situation is so simple that, as in the crypt arithmetic problem, we can estimate a priori how many trials, on the average, a subject should need to discover the intended concept provided that the subject used the most efficient discovery strategy. On each trial, regardless of response, the subject can determine from the experimenters reinforcement whether the stimulus was actually an instance of the concept or not. If it was an instance, the subject knows that one of the attribute values of the stimulus its color, size, shape, for example defines the concept. If it was not an instance, the subject knows that the *complement* of one of its attribute values defines the concept. In either case each trial rules out half of the possible simple concepts; and in a random sequence of stimuli each new stimulus rules out, on the average, approximately half of the concepts not previously eliminated. Hence the average number of trials required to find the right concept will vary with the logarithm of the number of dimensions in the stimulus.

If sufficient time were allowed for each trial (a minute, say, to be generous), and if the subject were provided with paper and pencil, any subject of normal intelligence could be taught to follow this most efficient strategy and would do so without much difficulty. As these experiments are actually run, subjects are not instructed in an efficient strategy, are not provided with paper and pencil, and take only a short time typically four seconds, say to respond to each successive stimulus. They also use many more trials to discover the correct concept than the number calculated from the efficient strategy. Although the experiment has not, to my knowledge, been run, it is fairly certain that, even with training, a subject who was required to respond in four seconds and not allowed paper and pencil would be unable to apply the efficient strategy.

What do these experiments tell us about human thinking? First, they tell us that human beings do not always discover for themselves clever strategies that they could readily be taught (watching a chess master play a duffer should also convince us of that). This is hardly a very startling conclusion, although it may be an instructive one. I shall return to it in a moment.

Second, the experiments tell us that human beings do not have sufficient means for storing information in memory to enable them to apply the efficient strategy unless the presentation of stimuli is greatly slowed down or the subjects are permitted external memory aids, or both. Since we know from other evidence that human beings have virtually unlimited semi-permanent storage (as indicated by their ability to continue to store odd facts in memory over most of a lifetime), the bottleneck in the experiment must lie in the small amount of rapid-access storage (so-called short-term memory) available and the time required to move items from the limited short-term store to the large-scale long-term store.[7]

From evidence obtained in other experiments, it has been estimated that only some seven items can be held in the fast, short-term memory and that perhaps as many as five to ten seconds are required to transfer an item from the short-term to the long-term store. To make these statements operational, we shall have to be more precise, presently, about the meaning of "item." For the moment let us assume that a simple concept is an item.

Even without paper and pencil a subject might be expected to apply the efficient strategy if (1) he was instructed in the efficient strategy and

7. The monograph by J. S. Bruner, J. J. Goodnow, and G. A. Austin, *A Study of Thinking* (New York: Wiley, 1956) was perhaps the first work to emphasize the role of short-term memory limits (their term was "cognitive strain") in performance on concept-attainment tasks. That work also provided rather definite descriptions of some of the subjects' strategies.

(2) he was allowed twenty or thirty seconds to respond to and process the stimulus on each trial. Since I have not run the experiment, this-assertion stands as a prediction by which the theory may be tested.

Again the outcome may appear obvious to you, if not trivial. If so, I remind you that it is obvious only if you accept my general hypothesis: that in large part human goal-directed behavior simply reflects the shape of the environment in which it takes place; only a gross knowledge of the characteristics of the human information-processing system is needed to predict it. In this experiment the relevant characteristics appear to be (1) the capacity of short-term memory, measured in terms of number of items (or "chunks," as I shall call them); (2) the time required to fixate an item, or chunk, in long-term memory. In the next section I shall inquire as to how consistent these characteristics appear to be over a range of task environments. Before I do so, I want to make a concluding comment about subjects' knowledge of strategies and the effects of training subjects.

That strategies can be learned is hardly a surprising fact, nor that learned strategies can vastly alter performance and enhance its effectiveness. All educational institutions are erected on these premises. Their full implication has not always been drawn by psychologists who conduct experiments in cognition. Insofar as behavior is a function of learned technique rather than "innate" characteristics of the human information-processing system, our knowledge of behavior must be regarded as sociological in nature rather than psychological that is, as revealing what human beings in fact learn when they grow up in a particular social environment. When and how they learn particular things may be a difficult question, but we must not confuse learned strategies with built-in properties of the underlying biological system.

The data that have been gathered, by Bartlett and in our own laboratory, on the crypt arithmetic task illustrate the same point. Different subjects do indeed apply different strategies in that task both the whole range of strategies I sketched in the previous section and others as well. How they learned these, or how they discover them while performing the task, we do not fully know (see chapter 4), although we know that the sophistication of the strategy varies directly with a subject's previous exposure to and comfort with mathematics. But apart from the strategies the only human characteristic that exhibits itself strongly in the crypt arithmetic task is the limited size of short-term memory. Most of the difficulties the subjects have in executing the more combinatorial strategies (and perhaps their general aversion to these strategies also) stem from the stress that such strategies place on short-term memory. Subjects get into trouble simply because they forget where they are, what assignments they have made previously, and what assumptions are implicit in assignments they have made conditionally. All of these difficulties would necessarily arise in a processor that could hold only a few chunks in short-term memory and that required more time than was available to transfer them to long-term memory.

The Parameters of Memory Eight Seconds per Chunk

If a few parameters of the sort we have been discussing are the main limits of the inner system that reveal themselves in human cognitive behavior, then it becomes an important task for experimental psychology to estimate the values of these parameters and to determine how variable or constant they are among different subjects and over different tasks.

Apart from some areas of sensory psychology, the typical experimental paradigms in psychology are concerned with hypothesis testing rather than parameter estimating. In the reports of experiments one can find many assertions that a particular parameter value is or is not "significantly different" from another but very little comment on the values themselves. As a matter of fact the pernicious practice is sometimes followed of reporting significance levels, or results of the analysis of variance, without reporting at all the numerical values of the parameters that underlie these inferences.

While I am objecting to publication practices in experimental psychology, I shall add another complaint. Typically little care is taken in choosing measures of behavior that are the most relevant to theory. Thus in learning experiments "rate of learning" is reported, almost indifferently, in terms of "number of *trials* to criterion," "total number of *errors*," "total *time* to criterion," and perhaps other measures as well. Specifically the practice of reporting learning rates in terms of trials rather than time, prevalent through the first half of this century, and almost up to the present time, not only hid from view the remarkable constancy of the parameter I am about to discuss but also led to much meaningless dispute over "one-trial" versus "incremental" learning.[8]

Ebbinghaus knew better. In his classic experiments on learning nonsense syllables, with himself as subject, he recorded both the number of repetitions and the amount of time required to learn sequences of syllables of different length. If you take the trouble to calculate it, you find that the *time per syllable* in his experiments works out to about ten to twelve seconds.[9]

I see no point in computing the figure to two decimal places or even to one. The constancy here is a constancy to an order of magnitude, or perhaps to a factor of two more nearly comparable to the constancy of the daily temperature, which in most places stays between 263° and 333° Kelvin, than to the constancy of the speed of light. There is no reason to be disdainful of a constancy to a factor of two. Newton's original estimates of the speed of sound contained a fudge factor of 30 per cent (eliminated only a hundred years later), and today some of the newer physical "constants" for elementary particles are even more vague. Beneath any approximate, even very rough, constancy, we can usually expect to find a genuine parameter whose value can be defined accurately once we know what conditions we must control during measurement.

If the constancy simply reflected a parameter of Ebbinghausalbeit one that held steady over several years it would be more interesting to biography than psychology. But that is not the case. When we examine some of the Hull-Hovland experiments of the 1930s, as reported, for ex-

8. The evidence of the constancy of the fixation parameter is reviewed in L. W. Gregg and H. A. Simon, "An Information-Processing Explanation of One-Trial and Incremental Learning," *Journal of Verbal Learning and Verbal Behavior, 6*(1967):780 787; H. A. Simon and E. A. Feigenbaum, "An Information-Processing Theory of Verbal Learning," *ibid., 3*(1964):385 396; Feigenbaum and Simon, "A Theory of the Serial Position Effect," *British Journal of Psychology, 53*(1962):307 320; E. A. Feigenbaum, "An Information-Processing Theory of Verbal Learning," unpublished doctoral dissertation, Pittsburgh: Carnegie Institute of Technology, 1959; and references cited therein. All these papers save the last are reprinted in *Models of Thought,* vol. 1

9. Herman Ebbinghaus, *Memory* (New York: Dover Publications, 1964), translated from the German edition of 1885, especially pp. 35 36, 40, 51.

ample, in Carl Hovland's chapter in S. S. Stevens's *Handbook*, we find again (after we calculate them, for trials are reported instead of times) times in the neighborhood of ten or fifteen seconds for college sophomores to fixate nonsense syllables of low meaningfulness by the serial anticipation method. When the drum speed increases (say from four seconds per syllable to two seconds per syllable), the number of trials to criterion increase proportionately, but the total learning time remains essentially constant.

There is a great deal of gold in these hills. If past nonsense-syllable experiments are re-examined from this point of view, many are revealed where the basic learning parameter is in the neighborhood of fifteen seconds per *syllable*. You can make the calculation yourself from the experiments reported, for example in J. A. McGeoch's *Psychology of Human Learning*. B. R. Bugelski, however, seems to have been the first to make this parameter constancy a matter of public record and to have run experiments with the direct aim of establishing it.[10]

I have tried not to exaggerate how constant is "constant." On the other hand, efforts to purify the parameter measurement have hardly begun. We do know about several variables that have a major effect on the value, and we have a theoretical explanation of these effects that thus far has held up well.

We know that meaningfulness is a variable of great importance. Nonsense syllables of high association value and unrelated one-syllable words are learned in about one-third the time required for nonsense syllables of low association value. Continuous prose is learned in about one-third the time per word required for sequences of unrelated words. (We can get the latter figure also from Ebbinghaus' experiments in memorizing *Doll Juan*. The times *per symbol* are roughly 10 percent of the corresponding times for nonsense syllables.)

We know that similarity particularly similarity among stimuli has an effect on the fixation parameter somewhat less than the effect of meaningfulness, and we can also estimate its magnitude on theoretical grounds.

10. B. R. Bugelski, "Presentation Time, Total Time, and Mediation in Paired-Associate Learning," *Journal of Experimental Psychology*, *63*(1962):409 412.

The theory that has been most successful in explaining these and other phenomena reported in the literature on rote verbal learning is an information-processing theory, programmed as a computer simulation of human behavior, dubbed EPAM.[11] Since EPAM has been reported at length in the literature, I shall not discuss it here, except for one point that is relevant to our analysis. The EPAM theory gives us a basis for understanding what a "chunk" is. A chunk is a maximal familiar substructure of the stimulus. Thus a nonsense syllable like "QUV" consists of the chunks "Q," "U," "V,"; but the word "CAT" consists of a single chunk, since it is a highly familiar unit. EPAM postulates constancy in the time required to fixate a chunk. Empirically the constant appears to be about eight seconds per chunk, or perhaps a little more. Virtually all the quantitative predictions that EPAM makes about the effects of meaningfulness, familiarity, and similarity upon learning speed follow from this conception of the chunk and of the constancy of the time required to fixate a single chunk.

In fixation of new information, EPAM first adds new branches to its discrimination net then adds information to images at terminal nodes of the branches. There is growing evidence that the eight seconds for fixation in long-term memory is required only for expanding the net, and that information can be added in a second or two to locations (variable-places) in images that are already present in an expert's long-term memory. Such images are called retrieval structures or templates. We will return to this point in discussing expert memory. EPAM's architecture and memory processes are described in H. B. Richman, J. J. Staszewski and H. A. Simon, "Simulation of Expert Memory Using EPAM IV," *Psychological Review, 102* (1995):305 330.

The Parameters of Memory Seven Chunks, or Is It Two?

The second limiting property of the inner system that shows up again and again in learning and problem-solving experiments is the amount of

[11.] For a survey of the range of phenomena for which EPAM has been tested, see E. A. Feigenbaum and H. A. Simon, "EPAM-like Models of Recognition and Learning" *Cognitive Science, 8*(1984): 305 336, reprinted in *Models of Thought,* vol. 2 (1989), chapter 3.4.

information that can be held in short-term memory. Here again the relevant unit appears to be the chunk, where this term has exactly the same meaning as in the definition of the fixation constant.

Attention was attracted to this parameter, known previously from digit span, numerosity-judging, and discrimination tasks, by George Miller's justly celebrated paper on "The Magical Number Seven, Plus or Minus Two."[12] It is no longer as plausible as it was when he wrote his paper that a single parameter is involved in the three kinds of task, rather than three different parameters: we shall consider here only tasks of the digit-span variety. Today we would express the parameter as the amount of information that can be rehearsed in about two seconds, which is, in fact, about seven syllables or short words.

The facts that appear to emerge from recent experiments on short-term memory are these. If asked to read a string of digits or letters and simply to repeat them back, a subject can generally perform correctly on strings up to seven or even ten items in length. If almost any other task, however simple, is interposed between the subject's hearing the items and repeating them, the number retained drops to two. From their familiarity in daily life we could dub these numbers the "telephone directory constants." We can generally retain seven numbers from directory to phone if we are not interrupted in any way not even by our own thoughts.

Where experiments appear to show that more than two chunks are retained across an interruption, the phenomena can almost always be explained parsimoniously by mechanisms we have already discussed in the previous section. In some of these experiments the explanation as already pointed out by Milleris that the subject recodes the stimulus into a smaller number of chunks before storing it in short-term memory. If ten items can be recoded as two chunks, then ten items can be retained. In the other experiments where "too much" appears to be retained in short-term memory, the times allowed the subjects permit them in fact to fixate the excess of items in long-term memory. For experts who have acquired retrieval structures or templates in their domain of expertise into which the new information can be inserted, these times can be quite short a second or two per item.

12. *Psychological Review, 63*(1956):81 97.

Putting aside expert performance for the moment, I shall cite just two examples from the literature. N. C. Waugh and D. A. Norman report experiments, their own and others', that show that only the first two of a sequence of items is retained reliably across interruption, but with some residual retention of the remaining items.[13] Computation of the fixation times available to the subjects in these experiments shows that a transfer rate to long-term memory of one chunk per five seconds would explain most of the residuals. (This explanation is entirely consistent with the theoretical model that Waugh and Norman themselves propose.)

Roger Shepard has reported that subjects shown a very long sequence of photographs mostly landscape scan remember which of these they have seen (when asked to choose from a large set) with high reliability.[14] When we note that the task is a recognition task, requiring storage only of differentiating cues, and that the average time per item was about six seconds, the phenomenon becomes entirely understandable indeed predictable within the framework of the theory that we are proposing.

The Organization of Memory

I have by no means exhausted the list of experiments I could cite in support of the fixation parameter and the short-term capacity parameter and in support of the hypothesis that these parameters are the principal, and almost only, characteristics of the information-processing system that are revealed, or could be revealed, by these standard psychological experiments.

This does not imply that there are not other parameters, and that we cannot find experiments in which they are revealed and from which they can be estimated. What it does imply is that we should not look for great complexity in the laws governing human behavior, in situations where the behavior is truly simple and only its environment is complex.

In our laboratory we have found that mental arithmetic tasks, for instance, provide a useful environment for teasing out other possible pa-

13. N. C. Waugh and D. A. Norman, "Primary Memory," *Psychological Review, 72*(1965):89 104.

14. Roger N. Shepard, "Recognition Memory for Words, Sentences, and Pictures," *Journal of Verbal Learning and Verbal Behavior, 6*(1957):156 163.

rameters. Work that Dansereau has carried forward shows that the times required for elementary arithmetic operations and for fixation of intermediate results account for only part perhaps one-half of the total time for performing mental multiplications of four digits by two. Much of the remaining time appears to be devoted to retrieving numbers from the memory where they have been temporarily fixated, and "placing" them in position in short-term memory where they can be operated upon.[15]

Stimulus Chunking

I should like now to point to another kind of characteristic of the inner system more "structural" and also less quantitative that is revealed in certain experiments. Memory is generally conceived to be organized in an "associative" fashion, but it is less clear just what that term is supposed to mean. One thing it means is revealed by McLean and Gregg. They gave subjects lists to learn specifically 24 letters of the alphabet in scrambled order. They encouraged, or induced, chunking of the lists by presenting the letters either one at a time, or three, four, six, or eight on a single card. In all of the grouped conditions, subjects learned in about half the time required in the one-at-a-time condition.[16]

McLean and Gregg also sought to ascertain whether the learned sequence was stored in memory as a single long list or as a hierarchized list of chunks, each of which was a shorter list. They determined this by measuring how subjects grouped items temporally when they recited the list, and especially when they recited it backwards. The results were clear: the alphabets were stored as sequences of short sub sequences; the sub sequences tended to correspond to chunks presented by the experimenter, or sub lengths of those chunks; left to his own devices, the subject tended to prefer chunks of three or four letters. (Recall the role of chunks of this length in the experiments on effects of meaningfulness in rote learning.)

15. See Donald F. Dansereau and Lee W. Gregg, "An Information Processing Analysis of Mental Multiplication," *Psychonomic Science, 6*(1966):71 72. The parameters of memory are discussed in more detail in *Models of Thought,* vol. 1, chapters 2.2, and 2.3; and vol. 2, chapter 2.4; and in Richman, Staszewski and Simon, *op. cit.*

16. R. S. McLean and L. W. Gregg, "Effects of Induced Chunking on Temporal Aspects of Serial Recitation," *Journal of Experimental Psychology, 74*(1967): 455 459.

Visual Memory

The materials in the McLean-Gregg experiments were strings of symbols. We might raise similar questions regarding the form of storage of information about two-dimensional visual stimuli.[17] In what sense do memory and thinking represent the visual characteristics of stimuli? I do not wish to revive the debate on "imageless thought "certainly not in the original form that debate took. But perhaps the issue can now be made more operational than it was at the turn of the century.

As I enter into this dangerous ground, I am comforted by the thought that even the most fervent opponents of mentalism have preceded me. I quote, for example, from B. F. Skinner's *Science and Human Behavior* (1952, p. 266):

> A man may see or hear "stimuli which are not present" on the pattern of the conditioned reflexes: he may see X, not only when X is present, but when any stimulus which has frequently accompanied X is present. The dinner bell not only makes our mouth water, it makes us see food.

I do not know exactly what Professor Skinner meant by "seeing food," but his statement gives me courage to say what an information-processing theory might mean by it. I shall describe in a simplified form one kind of experiment that has been used to throw light on the question. Suppose we allow a subject to memorize the following visual stimulus a magic square:

4	9	2
3	5	7
8	1	6

Now we remove the stimulus and ask the subject a series of questions about it, timing his or her answers. What numeral lies to the right of 3, to the right of 1? What numeral lies just below 5? What numeral is diagonally above and to the right of 3? The questions are not all of the same difficulty in fact I have arranged them in order of increasing difficulty

17. The letters in the stimuli of the McLean-Gregg experiment are, of course, also two-dimensional visual stimuli. Since they are familiar chunks, however, and can be immediately recognized and recoded, there is no reason to suppose that their two-dimensional character plays any role in the subject's behavior in the experiment. Again this is "obvious" but only if we already have a general theory of how stimuli are processed "inside."

and would expect a subject to take substantially longer to answer the last question than the first.

Why should this be? If the image stored in memory were isomorphic to a photograph of the stimulus, we should expect no large differences in the times required to answer the different questions. We must conclude that the stored image is organized quite differently from a photograph. An alternative hypothesis is that it is a list structure a hypothesis that is consistent, for example, with the data from the McLean-Gregg experiment and that is much in the spirit of information-processing models of cognition.

For example, if what was stored were a list of lists: "TOP," "MIDDLE," "BOTTOM," where "TOP" is 4-9-2, "MIDDLE" is 3-5-7, and "BOTTOM" is 8-1-6; the empirical results would be easy to understand. The question "What numeral lies to the right of 3?" is answered by searching down lists. The question "What numeral lies just below 5?" is answered, on the other hand, by matching two lists, item by item a far more complex process than the previous one.

There is no doubt, of course, that a subject could *learn* the up-down relations or the diagonal relations as well as the left-right relations. An EPAM-like theory would predict that it would take the subject about twice as long to learn both left-right and up-down relations as the former alone. This hypothesis can be easily tested, but, to the best of my knowledge, it has not been.

Evidence about the nature of the storage of "visual" images, pointing in the same direction as the example I have just given, is provided by the well-known experiments of A. de Groot and others on chess perception.[18] De Groot put chess positions taken from actual games before subjects for, say, five seconds; then he removed the positions and asked the subjects to reconstruct them. Chess grandmasters and masters could reconstruct the positions (with perhaps 20 to 24 pieces on the board) almost without error, while duffers were able to locate hardly any of the pieces correctly, and the performance of players of intermediate skill fell somewhere

18. Adriaan D. de Groot, "Perception and Memory versus Thought: Some Old Ideas and Recent Findings," in B. Kleinmuntz (ed.), *Problem Solving* (New York: Wiles; 1966), pp. 19 50. See also the work by Chase and Simon reported in chapters 6.4 and 6.5 of *Models of Thought* vol. 1.

between masters and duffers. But the remarkable fact was that, when masters and grandmasters were shown other chessboards with the same numbers of pieces *arranged at random*, their abilities to reconstruct the boards were only marginally better than the duffers' with the boards from actual games, while the duffers performed as well or poorly as they had before.

What conclusion shall we draw from the experiment? The data are inconsistent with the hypothesis that the chess masters have some special gift of visual imagery or else why the deterioration of their performance? What the data suggest strongly is that the information about the board is stored in the form of *relations* among the pieces, rather than a "television scan" of the 64 squares. It is inconsistent with the parameters proposed earlier seven chunks in short-term memory and five seconds to fixate a chunk to suppose that anyone, even a grandmaster, can store 64 pieces of information (or 24) in ten seconds. It is quite plausible that he can store (in short-term and long-term memory) information about enough relations (supposing each one to be a familiar chunk) to permit him to reproduce the board of figure 4:

1. Black has castled on the K's side, with a fianchettoed K's bishop defending the K's Knight.

2. White has castled on the Q's side, with his Queen standing just before his King.

3. A Black pawn on his K5 and a White pawn on his Q5 are attacked and defended by their respective K's and Q's Knights, the White Queen also attacking the Black pawn on the diagonal.

4. White's Q-Bishop attacks the Knight from KN5.

5. The Black Queen attacks the White K's position from her QN3.

6. A Black pawn stands on its QB4.

7. A White pawn on K3 blocks that advance of the opposing Black pawn.

8. Each side has lost a pawn and a Knight.

9. White's K-Bishop stands on K2.

Pieces not mentioned are assumed to be in their starting positions. Since some of the relations as listed are complex, I shall have to provide reasons for considering them unitary "chunks." I think most strong chess players would regard them as such. Incidentally I wrote down these relations from my own memory of the position, in the order in which they

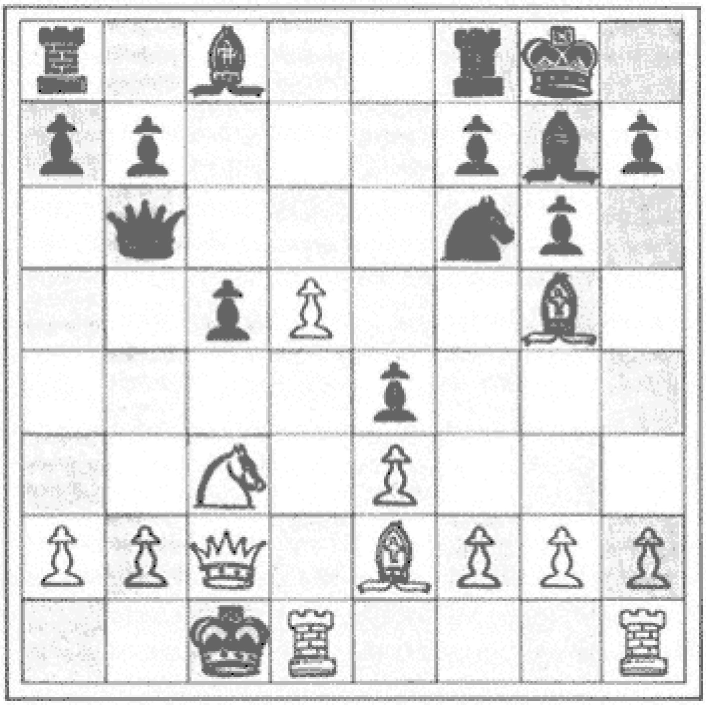

Figure 4
Chess position used in memory experiment

occurred to me. Eye-movement data for an expert chess player looking at this position tend to support this analysis of how the relations are analyzed and stored.[19] The eye-movement data exhibit with especial clarity the relations 3 and 5.

The expert can store the information about the position even more rapidly if he or she recognizes the standard opening to which it belongs in this case, the Gruenfeld Defense there by accessing a familiar template that gives the positions of about a dozen pieces.

The implication of this discussion of visual memory for my main theme is that many of the phenomena of visualization do not depend in any

19. O. K. Tikhomirov and E. D. Poznyanskaya, "An Investigation of Visual Search as a Means of Analysing Heuristics," English translation from *Voprosy psikhologii*, 1966, vol. 12, in *Soviet Psychology*, 2(Winter 1966 1967):3 15. See also *Models of Thought*, vol. 1, chapters 6.2 and 6.3.

detailed way upon underlying neurology but can be explained and predicted on the basis of quite general and abstract features of the organization of memory features which are essentially the same ones that were postulated in order to build information-processing theories of rote learning and of concept attainment phenomena.

Specifically, we are led to the hypothesis that memory is an organization of list structures (lists whose components can also be lists), which include descriptive components (two-termed relations) and short (three-element or four-element) component lists. A memory with this form of organization appears to have the right properties to explain storage phenomena in both visual and auditory modalities, and of pictorial and diagrammatic as well as propositional (verbal and mathematical) information.

The Mind's Eye

The experiments we have been discussing relate not only to visual long-term memory, but also to the Mind's Eye, the short-term memory where we hold and process mental images. In the mind's eye we can often substitute "seeing" for reasoning. Consider the economist's common supply-and-demand diagram, which shows, by one curve, the quantity of a commodity that will be supplied to the market at each price, and by another curve, the quantity that will be demanded at each price. If we notice that the two curves intersect, we can interpret the intersection as the point at which the supply and demand quantities are equal, a point of market equilibrium; and we can read off directly from the x-axis and y-axis of the diagram the equilibrium quantity and price (the x and y coordinates of the intersection). All this processing goes on in the mind's eye, using the information read from the diagram.

Alternatively, we could write down the equations for the two lines and solve them simultaneously to find the same equilibrium quantity and price. Using visual processes and algebraic ones we attain the same knowledge, but by completely different computational paths (and perhaps with vastly different amounts of labor and insight). In many scientific fields, inferences are made with a combination of verbal, mathematical and diagrammatic reasoning certain inferences being reached more easily in one form, others in another. In Alfred Marshall's famous text book, *Principles of Economics*, the text is wholly verbal, the diagrams are provided in footnotes, and the corresponding algebra is given in a mathematical appendix, thus allowing readers full freedom to adopt their preferred representation in each instance.

To understand the interplay of these and other modes of human inference, we need to study the computational processes required to reach conclusions in each representation. Currently, this is a very active area of cognitive research.[20]

Processing Natural Language

A theory of human thinking cannot and should not avoid reference to that most characteristic cognitive skill of human beings the use of language. How does language fit into the general picture of cognitive processes that I have been sketching and into my general thesis that psychology is a science of the artificial?

Historically the modern theory of transformational linguistics and the information-processing theory of cognition were born in the same matrix the matrix of ideas produced by the development of the modern digital computer, and in the realization that, though the computer was embodied in hardware, its soul was a program. One of the initial professional papers on transformational linguistics and one of the initial professional papers on information-processing psychology were presented, the one after the other, at a meeting at MIT in September 1956.[21] Thus the two bodies of theory have had cordial relations from an early date, and quite rightly, for they rest conceptually on the same view of the human mind.

20. J. Larkin and H. A. Simon, "Why a Diagram is (Sometimes) Worth 10,000 Words," *Cognitive Science, 11*(1987):65 100; A. M. Leonardo, H. J. M. Tabachneck and H. A. Simon, "A Computational Model of Diagram Reading and Reasoning" *Proceedings of the 17th Annual Conference of the Cognitive Science Society* (1995); Y. Qin and H. A. Simon, "Imagery and Mental Models of Problem Solving," in J. Glasgow, N. H. Narayanan and B. Chandrasekaran (eds.), *Diagrammatic Reasoning: Computational and Cognitive Perspectives* (Menlo Park, CA: AAAI/The MIT Press, 1995).

21. N. Chomsky, "Three Models for the Description of Language," and A. Newell and H. A. Simon, "The Logic Theory Machine," both in *IRE Transactions on Information Theory,* IT-2, no. 3 (September 1956).

Now some may object that this is not correct and that they rest on almost diametrically opposed views of the human mind. For I have stressed the artificial character of human thinking how it adapts itself, through individual learning and social transmission of knowledge, to the requirements of the task environment. The leading exponents of the formal linguistic theories, on the other hand, have taken what is sometimes called a "nativist" position. They have argued that a child could never acquire any skill so complex as speaking and understanding language if he did not already have built into him at birth the basic machinery for the exercise of these skills.

The issue is reminiscent of the debate on language universals on whether there are some common characteristics shared by all known tongues. We know that the commonalities among languages are not in any sense specific but that they relate instead to very broad structural characteristics that all languages seem to share in some manner. Something like the distinction between noun and verb between object and action or relation appears to be present in all human languages. All languages appear to have the boxes-within-boxes character called phrase structure. All languages appear to derive certain strings from others by transformation.[22]

Now if we accept these as typical of the universals to which the nativist argument appeals, there are still at least two different possible interpretations of that argument. The one is that the language competence is *purely* linguistic, that language is *sui generis,* and that the human faculties it calls upon are not all employed also in other performances.

An alternative interpretation of the nativist position is that producing utterances and understanding the utterances of others depend on some characteristics of the human central nervous system which are common in all languages but also essential to other aspects of human thinking besides speech and listening.

The former interpretation does not, but the latter does, provide an explanation for the remarkable parallelism holding between the underlying

[22.] On language universals see Joseph H. Greenberg (ed.), *Universals of Language* (Cambridge: The MIT Press, 1963), particularly Greenberg's own chapter, pp. 58 90. On the "nativist" position, see Jerrold J. Katz, *The Philosophy of Language* (New York: Harper and Row, 1966), pp. 240 282.

assumptions about human capabilities embedded in modern linguistic theory and the assumptions embedded in information-processing theories of human thinking. The kinds of assumptions that I made earlier about the structure of human memory are just the kinds of assumptions one would want to make for a processing system capable of handling language. Indeed there has been extensive borrowing back and forth between the two fields. Both postulate hierarchically organized list structures as a basic principle of memory organization. Both are concerned with how a serially operating processor can convert strings of symbols into list structures or list structures into strings. In both fields the same general classes of computer-programming languages have proved convenient for modeling and simulating the phenomena.

Semantics in Language Processing

Let me suggest one way in which the relation between linguistic theories and information-processing theories of thinking is going to be even closer in the future than it was in the past. Linguistic theory has thus far been largely a theory of syntax, of grammar. In practical application to such tasks as automatic translation, it has encountered difficulties when translation depended on more than syntactic cues when it depended on context and meaning. It seems pretty clear that one of the major directions that progress in linguistics will have to take is toward development of an adequate semantics to complement syntax.

The theory of thinking I have been outlining can already provide an important part of such a semantic component. The principles of memory organization I have described can be used as a basis for discussing the internal representation of both linguistic strings and two-dimensional visual stimuli, or other non-linguistic stimuli. Given these comparable bases for the organization of the several kinds of stimuli, it becomes easier to conceptualize the cooperation of syntactic and semantic cues in the interpretation of language.

Several research projects have been carried out at Carnegie Mellon University that bear on this point. I should like to mention just two of these, which illustrate how this approach might be used to explain the resolution of syntactic ambiguities by use of semantic cues.

L. Stephen Coles, in a dissertation completed in 1967, described a computer program that uses pictures on a cathode ray tube to resolve syntactic ambiguities.[23] I shall paraphrase his procedure with an example that is easier to visualize than any he actually used. Consider the sentence:

I saw the man on the hill with the telescope.

This sentence has at least three acceptable interpretations; a linguist could, no doubt, discover others. Which of the three obvious ones we pick depends on where we think the telescope is: Do I have it? Does the man on the hill have it? Or is it simply on the hill, not in his hands?

Now suppose that the sentence is accompanied by figure 5. The issue is no longer in doubt. Clearly it is I who have the telescope.

Coles's program is capable of recognizing objects in a picture and relations among objects; and it is capable of representing the picture as a list structure, which, in the example before us, we might describe thus:

SAW ((I, WITH (telescope)), (man, ON (hill))).

I have not tried to reproduce the actual details of the scheme he used, but I have simply shown that a picture, so represented, could readily be matched against alternate parsings of a verbal string and thus used to resolve the ambiguity of the latter.

Another program, completed by Laurent Siklóssy, illustrates how semantic information can aid in the *acquisition* of a language.[24] The reader may be familiar with the "Language through Pictures" books developed by Professor I. A. Richards and his associates. These books have been prepared for a large number of languages. On each page is a picture and beneath it one or more sentences that say something about the picture in the language to be learned. The sequence of pictures and accompanying sentences is arranged to proceed from very simple situations ("I am here," "That is a man") to more complex ones ("The book is on the shelf").

Siklóssy's program takes as its input an analogue to one of the "Language through Pictures" books. The picture is assumed to have already

23. L. Stephen Coles, *Syntax Directed Interpretation of Natural Language,* doctoral dissertation, Carnegie Institute of Technology, 1967. A slightly abridged version is reprinted in H.A. Simon and L. Siklóssy (eds.), *Representation and Meaning* (Englewood Cliffs, N.J.: Prentice-Hall, 1972).

24. Also reprinted in *Representation and Meaning*.

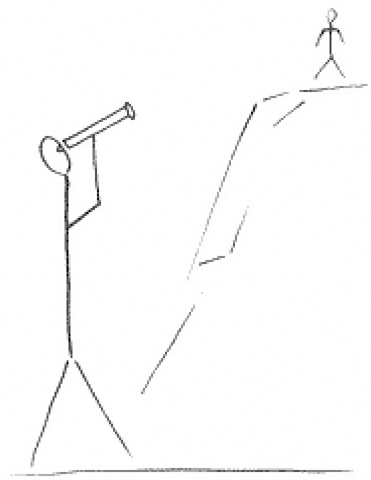

Figure 5
A syntactically ambiguous sentence;
"I saw the man on the hill with the telescope"

been transformed into a list structure (not unlike the one illustrated earlier for Coles's system) as its internal representation. The program's task is to learn, when confronted with such a picture, to utter the appropriate sentence in the natural language it is learning a sentence that says what the picture shows. In the case of the sentence about the telescope (somewhat more complicated than any on which the scheme has actually been tested), one would hope that the program would respond to the picture with "I saw the man on the hill with the telescope," if it were learning English, or *Ich habe den Mann auf dem Berg mit dem Fernglas gesehen*, if it were learning German.

Of course the program could respond correctly only if it had learned earlier, in the context of other sentences, the lexical and syntactical components required for the translation. A child trying to understand the sentence must meet the same requirement. In other cases the program would use the sentence associated with the picture to add to its vocabulary and syntax.[25]

[25.] I may mention in passing that Siklóssy's system refutes John Searle's notorious "Chinese Room Paradox," which purports to prove that a computer cannot understand language. As Siklóssy's program shows, if the room has windows on the world (which Searle's room doesn't) the system matches words, phrases and sentences to their meanings by comparing sentences with the scenes they denote.

I do not wish to expand some pioneering experiments into a comprehensive theory of semantics. The point of these examples is that they show that the kind of memory structure that has been postulated, for other reasons, to explain human behavior in simpler cognitive tasks is suitable for explaining how linguistic strings might be represented internally, how other kinds of stimuli might be similarly represented, and how the communalities in representation the use of hierarchically organized list structures for both may explain how language and "meanings" come together in the human head.

There is no contradiction, then, between the thesis that a human being possesses, at birth, a competence for acquiring and using language and the thesis that language is the most artificial, hence also the most human of all human constructions. The former thesis is an assertion that there *is* an inner environment and that it does place limits on the kinds of information processing of which the organism is capable. The structure of language reveals these limits; and these limits in turn account for such commonality as exists among the Babel of human tongues.

The latter thesis, of the artificiality of language, is an assertion that the limits on adaptation, on possible languages, imposed by the inner environment are very broad limits on organization, not very specific limits on syntax. Moreover, according to the thesis, they are limits imposed not only on language but also on every other mode of representing internally experience received through stimuli from outside.

Such a view of the relation of language and thinking puts a new cast on the "Whorfian" hypothesis that stating it in over strong form only the expressible is thinkable. If the view is valid, it would be as correct to say. "Only the thinkable is expressible "a view that, I suppose, Kant would have found quite congenial.

Conclusion

The thesis with which I began this chapter was the following:

Human beings viewed as behaving systems, are quite simple. The apparent complexity of our behavior over time is largely a reflection of the complexity of the environment in which we find ourselves.

That hypothesis was based in turn on the thesis of the first chapter: that behavior is adapted to goals, hence is artificial, hence reveals only those characteristics of the behaving system that limit the adaptation.

To illustrate how we have begun to test these theses and at the same time to build up a theory of the simple principles that underlie human behavior, I have surveyed some of the evidence from a range of human performances, particularly those that have been studied in the psychological laboratory.

The behavior of human subjects in solving crypt arithmetic problems, in attaining concepts, in memorizing, in holding information in short-term memory, in processing visual stimuli, and in performing tasks that use natural languages provides strong support for these theses. The artificiality hence variability of human behavior hardly calls for evidence beyond our observation of everyday life. The experiments are therefore mostly significant in what they show about the broad commonalities in organizations of the human information-processing system as it engages in different tasks.

The evidence is overwhelming that the system is basically serial in its operation: that it can process only a few symbols at a time and that the symbols being processed must be held in special, limited memory structures whose content can be changed rapidly. The most striking limits on subjects' capacities to employ efficient strategies arise from the very small capacity of the short-term memory structure (seven chunks) and from the relatively long time (eight seconds) required to transfer a chunk of information from short-term to long-term memory.

The claim that the human cognitive system is basically serial has been challenged in recent years by advocates of neural nets and parallel connectionist models of the nervous system. I would make the following cautionary observations. Although there is clearly a lot of parallelism in the sensory organs (especially eyes and ears), after stimuli have been recognized seriality is enforced by the small capacity of the short-term memory that is employed in the subsequent stages of processing. There is also a moderate degree of parallelism in the processing of motor signals, but again, only after the initial signals have passed through the STM bottleneck. Third, seriality of processing at the symbolic level, the level with which we are concerned here, says nothing, one way or the other, about the extent of seriality or parallelism in the neural implementation of the symbolic processing at the next level below. (By an ironic reverse twist, parallel connectionist networks are routinely simulated by programs run on serial computers of standard von Neumann architecture.)

Finally, a large part of the discernible parallel neural activity in the brain may well consist only in passive maintenance of memory, the active processes being largely localized and serial. (Evidence now coming from magnetic resonance imaging [MRI] of the brain is consistent with this view.) The speeds at which people can perform cognitive tasks and the usual limits on the numbers of tasks they can perform concurrently do not provide much evidence for (or need for) parallel processing capacity. Until connectionism has demonstrated, which it has not yet done, that complex thinking and problem-solving processes can be modeled as well with parallel connectionist architectures as they have been with serial architectures, and that the experimentally observed limits on concurrent cognitive activity can be represented in the connectionist models, the case for massive parallelism outside the sensory functions remains dubious.

When we turn from tasks that exercise mainly the short-term memory and serial-processing capabilities of the central nervous system to tasks that involve retrieval of stored information, we encounter new limits of adaptation, and through these limits we acquire new information about the organization of mind and brain. Studies of visual perception and of tasks requiring use of natural language show with growing clarity that memory is indeed organized in associative fashion, but that the "associations" have the properties of what, in the computer trade, are usually called "list structures." I have indicated briefly what those properties are, and more will be said about them in the next chapter.

These are the sorts of generalizations about human thinking that are emerging from the experimental evidence. They are simple things, just as our hypothesis led us to expect. Moreover, though the picture will continue to be enlarged and clarified, we should not expect it to become essentially more complex. Only human pride argues that the apparent intricacies of our path stem from a quite different source than the intricacy of the ant's path.

One of the curious consequences of my approach of my thesis is that I have said almost nothing about physiology. But the mind is usually thought to be located in the brain. I have discussed the organization of the mind without saying anything about the structure of the brain.

The main reason for this disembodiment of mind is of course the thesis that I have just been discussing. The difference between the hardware of a computer and the "hardware" of the brain has not prevented computers from simulating a wide spectrum of kinds of human thinking just because both computer and brain, when engaged in thought, are adaptive systems, seeking to mold themselves to the shape of the task environment.

It would be unfortunate if this conclusion were altered to read that neurophysiology has nothing to contribute to the explanation of human behavior. That would be of course a ridiculous doctrine. But our analysis of the artificial leads us to a particular view of the form that the physiological explanation of behavior must take. Neurophysiology is the study of the inner environment of the adaptive system called Homo sapiens. It is to physiology that we must turn for an explanation of the limits of adaptation: Why is short-term memory limited to seven chunks; what is the physiological structure that corresponds to a "chunk"; what goes on during the eight seconds that a chunk is being fixated; how are associational structures realized in the brain?

As our knowledge increases, the relation between physiological and information-processing explanations will become just like the relation between quantum-mechanical and physiological explanations in biology (or the relation between solid-state physics and programming explanations in computer science). They constitute two linked levels of explanation with (in the case before us) the limiting properties of the inner system showing up at the interface between them.

Finally, we may expect also that, as we link information-processing psychology to physiology on the inner side, we shall also be linking psychology to the general theory of search through large combinatorial spaces on the outer side the side of the task environment. But that is the topic of my fifth chapter, for the theory of design is that general theory of search. Before we take up that topic we must say more about how the large bodies of information used by designers are stored in the human mind and accessed.