Cognitive Therapy and Research is a broadly conceived interdisciplinary journal whose main function is to stimulate and communicate research and theory on the role of cognitive processes in human adaptation and adjustment. It attempts to integrate such diverse areas of psychology as clinical, counseling, experimental, cognitive, community, social, personality theory, self-regulation, physiology, biofeedback, philosophy, and the psychology of science. Articles considered for publication include the following: (1) Theoretical and Review Articles, (2) Experimental Studies, (3) Technical and Methodological Articles, (4) Case Studies, and (5) Brief Notes. The majority of publication space is devoted to experimental studies and brief notes.

Cognitive Therapy and Research is published quarterly by Plenum Publishing Corporation, 227 West 17th Street, New York, N.Y. 10011. Subscription orders should be addressed to the publisher.

Subscription rates:
Volume 2, 1978 (4 issues) $36.00. Price for individual subscribers certifying that the journal is for their personal use, $18.00. Prices somewhat higher outside the U.S.

Printed in USA.

Contacts of Cognitive Psychology with Social Learning Theory[1]

Gordon H. Bower[2]
Stanford University

In this paper, I will try to reveal some of the orienting beliefs of cognitive psychology about the organization of the mind and mental skills, and then try to relate that to social learning theory, which is a popular orientation of behavior therapists. A problem here is that cognitive psychologists seldom try to think about how to put together a mental apparatus that would develop a social personality. Therefore, this paper will contain a few speculations about how cognitive psychology views some aspects of social learning, personality, and behavior modification. But first I have to establish a framework within which to cast my remarks, and that is the information-processing framework. I will go through it briefly; for fuller exposition, see Bower (1975) or Rumelhart (1977).

THE INFORMATION-PROCESSING VIEW

The information-processing view of the person takes the computer as its guiding metaphor. The basic components are sensory receptors receiving inputs from the environment, effector units that produce responses, a memory store that holds data structures or programs, and a central processor in which occurs the focus of mental activity—thinking, reasoning, judging, and making decisions.

[1]Based on an invited address to the convention of the American Association of Behavior Therapists, Atlanta, Georgia, December 1977. The author's research is supported by grant MH-13905 from NIMH.
[2]Address all correspondence to Gordon H. Bower, Department of Psychology, Stanford University, Stanford, California 94305.

0147-5916/78/0600-0123$05.00/0

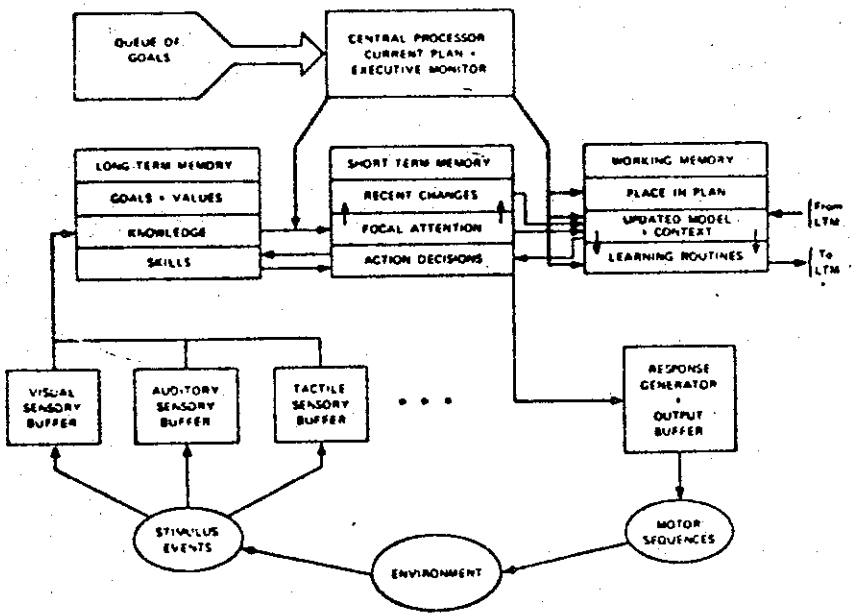

Fig. 1. Diagram of the principal components of the perception and memory system. Arrows indicate the direction of information flow and control among the components. Reprinted from Bower (1975) with permission.

An elaborate diagram of the human cognitive system is shown in Figure 1. The connections between components are specified by the arrows that suggest the flow of information and control. The diagram is best understood by moving up through it as though a stimulus were being evaluated for some response or as though some stimulus pattern were being learned. In this conception, an environmental stimulus event is registered briefly in sensory buffers and then undergoes a recognition or classification procedure which uses information about stimuli of that class in long-term memory. If it is attended to, the classified stimulus event is then entered into an active short-term memory. The contents of short-term memory are roughly what we mean by consciousness. As the stimulus enters short-term memory, the person may call up some associate of it, or may transform it to some other medium (as we do, for example, when we name a visual object). He may decide to respond overtly to it, or may simply try to transfer information about the event into long-term memory.

The working memory on the right in Figure 1 maintains information about the local context, but information that is not in the focus of active memory nor in the distant recesses of long-term memory. Its function is to

build up and maintain an internal model of the immediate environment and what has been happening over the past few minutes. This model is the framework within which the more rapid changes of the perceptual world before us are taking place, causing us to update one or another feature of our current view.

To mention operating principles, let us start by noting that the cognitive system developed through evolution as an instrument of adjustment, as an aid to guiding actions that would be effective in achieving adaptive goals and satisfying motives. It does this by carrying out action plans, which are like programs that reside in memory and which guide the actions taken in response to the contents of short-term memory. I will return later to this plan-following nature of the cognitive system.

The long-term memory is the repository of our more permanent knowledge and skills—it holds everything we know that is not currently in working memory. The information here may be conveniently divided into three types: sensory-perceptual knowledge, procedural-motoric knowledge, and propositional information and beliefs. Perceptual knowledge is represented in analog form in our sensory information store. This perceptual memory is used in classifying sensory patterns, in storing memories of the appearances of things, and in generating and manipulating mental images based on these percepts. Along with images of objects, this store also holds cognitive maps of the layout of objects in space. It is this stored perceptual knowledge that is used in behavior therapy technique using imagery, such as imaginal desensitization, imaginal rehearsal, and so on. Emotional conditioning experiences often operate through this imagery system, and it is heavily involved in fantasies evocative of strong feelings.

The second type of knowledge is procedural, representing information about *how* to do something. These may be motor skills such as how to ride a bicycle or hammer a nail, or more "intellectual" skills such as how to solve linear equations. One approach represents these skills in memory as a hierarchy of productions, much like a detailed cooking recipe. A production is like a generalized stimulus-response rule, only more powerful because it can use variables. A production has one or more conditions that must be satisfied for it to be selected and to be executed. When it is executed, one or more actions occur, and these may be either overt responses or internal ones such as fetching certain beliefs from long-term memory and making them active in short-term memory.

A single production, then, is a situation-action pair and may be thought of as a simple instruction. Sequences of productions are like stored computer programs which accomplish some result, such as searching memory for an answer to a question, solving a linear differential equation, baking a cake, or whatever.

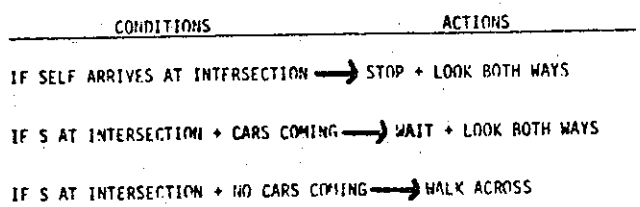

Fig. 2. Example of three productions for crossing a street intersection.

A simple example of a set of productions would be rules for crossing a busy intersection of streets, as illustrated in Figure 2. The condition for the first production is one's arrival at a street intersection, and the action is to stop and look both ways. If cars are coming, the second production says to wait and continue looking. If no cars are coming, the third production says to walk across the intersection. The point is that sets of productions can be learned and strung together to accomplish many different tasks, even very complex ones.

In this example, the conditions are given by sensory information and the actions are overt. However, productions can also apply to thoughts in short-term memory and to transformations of thoughts; in this way, productions can be used to describe the flow of associated ideas through consciousness. Thus, productions provide the control structures needed for cognitive theories, since they carry out successive moves or transformations in thinking, in problem-solving, and in planning actions.

A therapist might notice that productions are a useful vocabulary for describing neurotic thought processes—that is, the idiosyncratic interpretations of social stimuli and the unrealistic inferences the neurotic draws. Some relevant examples are shown in Figure 3.

Figure 3 shows productions for suspicious or paranoid inferences, where a compliment is interpreted as an attempt to manipulate and where a manipulation attempt is answered with a punch if the person in question is smaller and a verbal insult if he or she is larger. Perhaps you can see how to describe more sophisticated paranoid modes of thought with such rules.

The bottom of Figure 3 shows sample productions for depressed thinking. I have borrowed these from a fascinating and provocative article on depressed styles of thinking, published by Beck (1964). In that article he discussed interpretive schemas which channel and distort the thought processes of severely depressed patients. The first example is unwarranted inference, in which the client goes from disagreement to dislike, and then to catastrophizing that having anyone dislike him is a horrible situation he can't stand and which makes him feel sad. You should note that an emotional feeling can be on the action side of a production.

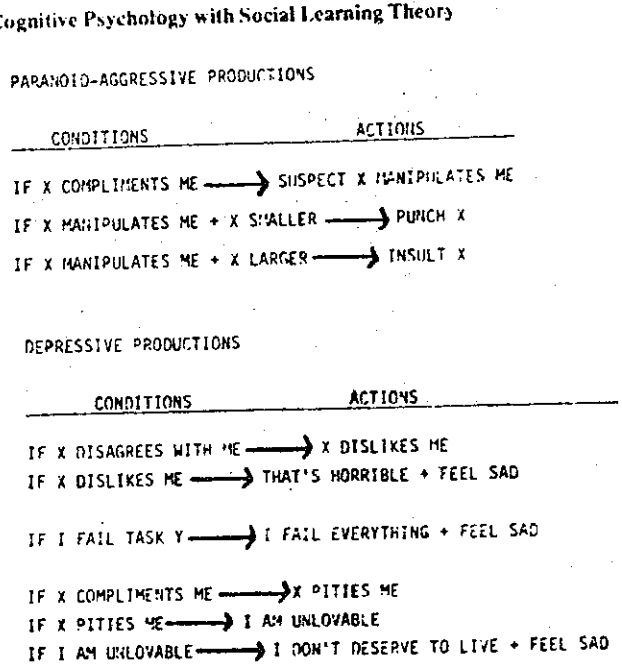

Fig. 3. Productions which drive interpretations and thoughts in a paranoid-aggressive manner (top) or a depressive manner (bottom).

The third production reflects overgeneralization in which, from a single failure, the depressive concludes that he fails at everything and therefore feels sad. The last three productions reflect distortion, whereby a compliment to a person is interpreted as insincere and as really showing that the complimenter pities the recipient. Notice here how a sequence of productions leads the depressed client from a compliment to the interpretation that he is being pitied, thence that he is an unlovable person, and finally to the conclusion that he does not deserve to live. These productions would move the stream of thought along automatically to its inexorable, morbid conclusion.

From the client's point of view these interpretations and inferences are natural, even obvious and inevitable. From the therapist's point of view these productions are unconscious automatisms—that is, automatic habits of thought. They can be broken by challenging them, interrupting them—more specifically, by rehearsing and practicing alternative productions or thinking habits which have more realistic and happier conclusions on the action side. This is one main technique used in rational-emotive therapy and in cognitive therapy with depressed clients.

To return to types of information in memory, the third type mentioned earlier was propositional as opposed to perceptual or procedural

knowledge. This includes our beliefs about ourselves and the world, our knowledge of concepts and word meanings, our knowledge of general facts and of specific objects, events, and episodes. It includes propositions about one's values or goals. A belief is a proposition with a subjective truth value or credibility attached to it, such as "I believe that Oswald shot Kennedy." A value proposition involves a primitive predicate of the form "I value preservation of international peace" or "I value freedom of worship" and such values will have intensity or importance ratings attached to them. In cognitive theories, propositional information is usually represented in terms of semantic networks, so let us examine such networks.

A semantic network represents the meaning of a concept in terms of its relationships to other concepts, much the way a dictionary defines one word in terms of others. Other information about the concept is represented in the same way, as relations this concept holds to others. Now, in order to have our concepts break out of a purely verbal circle, we must also have them point to things that make more direct contact with the world. Thus, we would have concrete concepts such as "hammer" associated to a corresponding sensory image. Also, action concepts such as "pounding with a hammer" would be associated with a production which, when executed, would cause us to pound some object with a hammer. So the conceptual networks are tied down on the sensory side and on the motor and affective output side.

In our theories conceptual networks are usually represented as graph structures in which nodes stand for concepts and a labeled arc between nodes stands for a semantic association between concepts. A simple illustration is shown in Figure 4 where an "introvert" is defined in terms of a general class, namely, it is a kind of human personality; its defining characteristics are traits like being shy, quiet, and a loner; and examples or instances of introverts that this person happens to know are Jack and Jill. Such graph structures can be used in answering questions such as, "Who's an example of an introvert?" or "Is someone who's a loner also likely to be shy?," and so on. Most of the personality stereotypes of our culture would be represented in terms of large graph structures like this in our memory. By the way, I use personality traits in my illustrations because I am representing the folklore psychology of the layman, not because I believe that general traits are useful theoretical tools.

Now, a realistic memory contains many thousands of such concepts, each with many connections, so that the network contains much implicit knowledge. What this means is that we can derive the answer to an enormous number of questions, far beyond the number of direct relations that are explicitly stored. To appreciate this point, consider just a fragment of the knowledge surrounding our concepts of an introvert and of a used-car

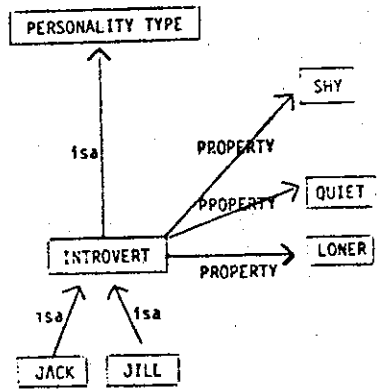

Fig. 4. Example of the concept of "introvert" introduced in terms of its relationships to other concepts.

salesman, as shown in Figure 5. Thus, Jack is an introvert; an introvert is a shy loner and loners are unsociables who probably cultivate private hobbies like reading. On the other hand, John is a used-car salesman, used-car salesmen are untrustworthy but persuasive extraverts, and so forth. Even such small networks permit many questions to be answered, at least approximately. Thus, a person with these beliefs would infer that Jack is probably a shy person who is bashful with women, that John is probably a smooth, fast-talking con man, that Jack would probably never make it as a used-car salesman, and that passive people are seldom very persuasive. These are implications among distant concepts that can be derived by elementary inference rules. Such derivations are evoked by questions which cause productions to be applied to some queried beliefs in order to come up with an answer. Thus, when we ask the network, "Does Jack like to read?," the system will search for a path between these two concepts, and derive the new conjecture that Jack probably likes to read as a hobby. So the model is simulating the sort of probable inferences that stereotypes allow us to make.

Let us consider a few implications of such stereotypes. First, I have just illustrated how the stereotype allows the believer to go from the fact that a person has certain personality traits to the expectation that that person will also have the rest of the traits in that stereotype. These kinds of expectations intrude when we try to remember a trait description of someone. If the trait labels yielded a general impression that the person was, say, an introvert, then when our memory is faint we are likely to believe that some specific introverted trait was ascribed to that person, when in fact it was not mentioned. Cantor and Mischel (1977; in press) found just such results

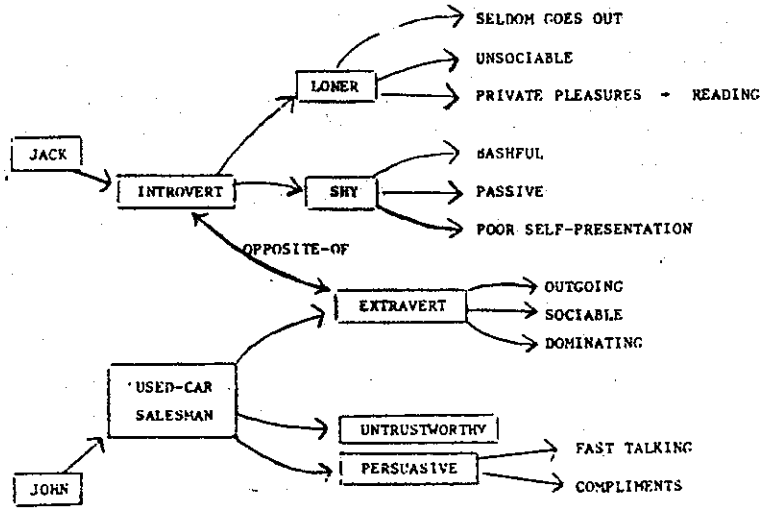

Fig. 5. Part of the network surrounding the concepts of "introvert" and "used-car salesman." Arrows usually represent "has property" relations.

when they studied people's memory for clusters of traits attributed to a fictional character. Sometimes the overall impression of someone is so salient that it overwhelms our faint recollections of conflicting information, leading us to doubt that such inconsistent information was ever told to us. Cantor and Mischel have also found that the distortions here are similar whether the items being recalled are trait names or are descriptions of actions which exemplify those traits.

A second implication here is that stereotypes will direct the process of "person perception," will bias the way we see and interpret someone's behavior. If I believe that Jack is an introvert, then I am prepared to judge him to be unsociable. Those kinds of personality judgments are so poorly constrained by Jack's actual behavior that he would have to go to considerable lengths to convince me that he is not unsociable. We all have ways of discounting any evidence that appears to conflict with our preconceptions. Thus, if I see Jack talking animatedly to a young woman at a party, I am likely to discount that by saying something like, "Well, he's probably drunk and she's probably a friend of his sister's who just asked him about his butterfly collection, which is the only thing he can ever get excited about." So, our preconceptions have a built-in confirmation bias when we are assessing evidence.

This should be an old tune to clinical psychologists familiar with the disturbing studies of Chapman and Chapman (1967, 1969). Those studies

had clinical students of psychodynamic persuasions predicting a patient's psychiatric diagnosis on the basis of personality-test results (Draw-a-Person test, or Rorschach test). Most of the subjects came away from a massive exposure to disconfirming evidence with their preconceptions about symptom-sign relations more firmly entrenched than ever before. This was despite the fact that the actual correlation in the data for their preconceived relation was zero. Moreover, the Chapmans (1969) went on to show that an observer's prior commitment to a false symptom-sign relation actually prevented him from noticing another strong sign-to-symptom correlation that was actually in the data. That is, the observers' causal model of how symptoms should relate to underlying pathology was blinding them to obvious relationships that were in fact in the data. This confirmation bias is one practitioners must beware of, even behavior therapists, since even they infer some category or pigeonhole to put the patient in, be it a classification of simple agoraphobia, or depression, or deficient social skills, or whatever.

A third implication of this view is that our conception about someone determines the way we act towards him. And our behavior towards him will be biased in such a way that he is most likely to respond to us in such manner as to confirm and maintain our initial belief about him. It is as though our stereotype is a prophecy that causes us to behave in such manner as to create a social reality that fulfills our prophecy.

A neat demonstration of this was arranged in an experiment by Snyder, Tanke, and Berscheid (in press). Male college students held a get-acquainted conversation over the telephone with a female student after they had been given a false photograph of her, showing her to be very attractive in some cases or rather unattractive in others. Before the conversation, the men expected the attractive women to be sociable, poised, humorous, and socially adept, whereas they expected the unattractive women to be just the reverse. After the conversation, the men, not surprisingly, judged a woman as more interesting, enthusiastic, and exciting if they thought she was beautiful. But more interestingly, it turned out that what a man expected is exactly what other people said that he got back from the woman he conversed with. Blind judges rated the woman's conversation alone, after the man's voice had been filtered out of the tape recordings. These judges rated the warmth, humor, and social adeptness of the woman to be much higher if she was interacting with a man who believed she was attractive. Thus, to quote the authors, "what had initially been reality in the minds of the men had now become reality in the behavior of the women with whom they interacted." This influence was exerted by the way the men talked. If a man thought he was talking to an attractive woman he was more animated, humorous, sexually suggestive, outgoing, bold, and socially adept—and the woman responded in kind. Obviously a woman can respond in either

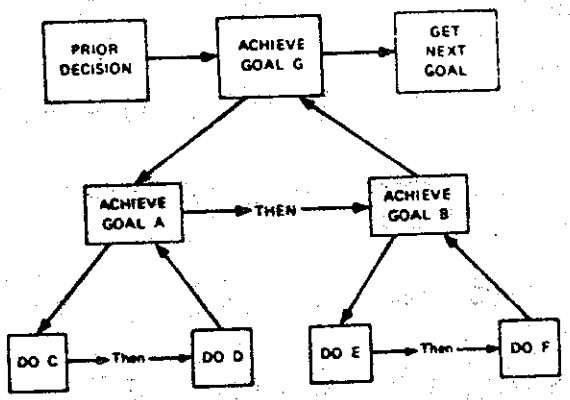

Fig. 6. Diagram of a hierarchical arrangement of subgoals embedded within a top-level goal. Reprinted from Bower (1975) with permission.

fashion, warm or cold, depending on how the male begins the conversation and moves it along. To summarize then, what we expect influences how we behave, and how we behave influences how the other person behaves back to us. We can believe that this is true too of a therapist interviewing a client. Incidentally, if we have a persisting stereotype about someone, that could create a constant biasing of the behavioral samples we see from him, with the result that we believe that he is a very consistent person. This may help explain our beliefs about cross-situational consistency of personality, despite the evidence against consistency reviewed by Mischel (1968) and others.

A prominent cluster of conceptual information we develop from childhood onward refers to ourself as an agent. The self-concept is a synthesis of many sources of evidence, putting together labels applied to us by parents, teachers, and peers, along with observations we have made of our own behavior. Although traditionally the self-concept has referred to relatively global intrapersonal value judgments, I think it would serve us better if we have the self-concept refer mainly to relatively specific assessments of our abilities in many different spheres. This is "meta-knowledge" about what we can do, and it corresponds to Bandura's construct of self-efficacy (1977), or expectations of success in different tasks. Our self-concept influences the way we attack or think about a problem to be solved. If we believe we are inept at a given task, we predict failure and do not enter the situation; or, if we do enter it, we invite failure by presenting ourselves poorly and by being distracted by anticipations of failure. As noted before, our self-concept not only affects our behavior but affects the behavior we get in return from others. If we look, talk, and act like losers, people are likely to

see us that way and treat us that way. As a result, the self-concept will be fulfilled and given another kick downward in a negative spiral.

The aim of some forms of psychotherapy is to enhance the client's self-esteem. I think techniques which teach specific abilities or socially valuable activities are useful, and so is realistic feedback on where the client stands in norms of ability. But merely having him repeat over and over to himself "What a terrific person I am" will have limited usefulness. The problem is that none of us is a credible source for himself in this request, so that the Goody-Twoshoes sentence goes into memory with a credibility near zero. There is no simple way to manufacture self-esteem all inside one's head. One must get the real evidence from behaving successfully in one's environment. That was part of Bandura's (1977) message, that actual behavioral achievements produce more change in self-efficacy than do imagined achievements, because it is more reliable evidence about ability.

PLANS

Having discussed a few components of the memory system, let me backtrack to pick up on the notion of an action plan, and how this relates to the topics of motives, purposes, and intentions. The architecture in Figure 1 indicated that the mental apparatus is usually under the control of one or another action plan. A behavioral episode may be characterized as beginning when a particular goal is selected to be worked on, and this goal together with the situation selects from memory some general plan that in the past has been useful in achieving similar goals in similar situations. The plan is a general procedure, or recipe for achieving a goal in tasks of particular types. In computer simulation models of thinking the plan is a set of productions or a program that is activated or loaded into working memory and which controls the flow of information and actions in short-term memory. Most of our plans are constructed or have been constructed using a heuristic that Newell and Simon (1972) call "means-ends analyses." This is a heuristic wherein a problem-solver repeatedly asks himself: "What's the difference between where I am now and where I want to go? Can I do something to reduce that difference? Is there something I must do first in order to do that? Can I do that first thing now?," and so on.

For most complicated goals a means-ends analysis will generate a hierarchy of subgoals, and these correspond to actions to be carried out as the plan is executed step by step. A simple hierarchical plan is illustrated in Figure 6. Suppose your top goal were to go to a theater play tonight (goal G). To do this, you must first reserve tickets (subgoal A). To reserve tickets, you must first look up the telephone number of the theater (do C), then telephone and reserve tickets (do D). If that succeeds, then this evening you

must get to your seat at the theater on time (subgoal B). This requires leaving on time, getting a taxi to the theater (do E), then paying for your tickets and walking to your seat for the start of the play (do F). By satisfying subgoal F, you also satisfy subgoal B, which is a precondition for you to satisfy the top goal G. After the play, you might begin operating on a new goal, such as finding a coffee house for refreshments after the play.

These hierarchical analyses of plans provide a neat way to segment or chunk significant aspects of behavior, and we can refine and expand the subgoal specification and behavioral description to as much detail as is of interest. Thus, "telephone and reserve tickets" is a primitive (D) in this analysis, but it could be decomposed into action scripts for telephoning, for selecting tickets among those available, and so on. We can give a relatively molar or molecular description of what the person is doing and why he is doing it by focusing on actions and subgoals at different levels of the hierarchical plan he is carrying out. Values and goals propagate down through such a hierarchy, and they provide us with the answers to questions about why the person is doing these things.

Let me consider the topics of motives, purposes, and intentions. In my terms a motive is roughly a valued goal plus a plan to achieve it. When a person commits himself to carrying out a plan, we say that he has an *intention* to do those things in the plan that will bring about the goal. The goal inserted at the top of the plan presumably has a high positive value. Each behavior node in the action plan has some small negative value or cost. As the person thinks over his plan or as he executes it, he is estimating whether the expected total costs will exceed the value gained, and if so he will choose another plan or perhaps alter his goal. A person may start out on a plan and then, because of unexpected costs or delays, revise it and switch to more efficient actions. For example, if I find that taking a taxi to the theater is too expensive, then I might catch the bus to get there.

How is a plan selected? This is based roughly on maximizing expected utility: Of those plans achieving the same goal, select that one you believe will be least costly and most likely to succeed. If two plans conflict with respect to their goals and you must choose, then choose that goal and plan which has the higher expected utility at the moment. The interesting distortions occur, of course, in estimating likelihoods of success. But that is another story.

Often we follow multiple goals. One way to pursue multipl goals is to alternate back and forth between the two plans. For instance, a parent with a very young child may entertain a guest at home for dinner, carrying on a conversation with the guest in alternation with feeding and caring for the child at about the same time. A second common way we handle multiple goals is to allocate certain times of the day when we will do the relevant ac-

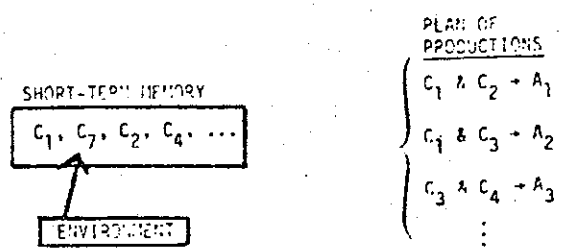

Fig. 7. Schematic relation between the set of productions representing a strategy and the contents of short-term memory (STM). The productions all "look at" the symbols in short-term memory, and any production whose conditions are satisfied by the symbols in STM will "fire" its associated action, which may cause new symbol to be entered into STM.

tivities—we set aside a time to work, eat, play, and so on. Time-management programs get us to decide how to allocate the time of our lives rationally rather than letting outside pressures totally dictate that for us.

These kinds of action plans should be familiar to behavior modifiers. Many behavior-modification treatments are formulated in just these terms. For example, typical programs for helping clients lose weight involve the client's learning subskills and reaching subgoals like how to count calories, when to record calories, how to set daily or weekly subgoals, how to avoid tempting foods, how to alter eating habits, how to exercise more, and so on.

Action plans also figure heavily in cognitive behavior therapy. Much of the work on self-instructional training can be viewed as teaching the client a specific plan which he should run through his working memory to control himself when he is in the problem situation. Figure 7 shows a schematic diagram of a possible relation between a self-instructional plan in working memory and the contents of short-term memory. The contents of short-term memory change moment by moment as the environment sends in signals. Also, the plan causes actions or causes information to be fetched from long-term memory and put into short-term memory. Meichenbaum (1977) has used self-instructional techniques with many different clinical problems—for example, he first used it successfully in training hyperactive or impulsive children to stop, look, and think about all the alternative courses of action rather than pouncing on the first available answer. In Meichenbaum's procedures, the person is first taught the plan by watching a model do the behavior while describing it step by step, then the client performs while describing the steps aloud, then he does it while subvocalizing, and then just thinking of the instructions. The most effective instructions seem to be those that are specific about critical features of the task and what

exactly is to be done to succeed; on the other hand, general pepping-up or esteem-building remarks seem less effective in guiding proficient performance.

Another therapy example of preparing an action plan is in the DESC scripting that my wife, Sharon, teaches her students in assertiveness training. She has described the technique in her book, *Asserting Yourself* (1976). The gist of the method is that the students analyze a recurrent put-down they receive from someone, plan a behavioral contract to rectify that situation, and then write and rehearse a specific script about that contract which they will deliver, perhaps when that put-down occurs again. These assertive DESC scripts are memorized plans, but with many countering side-scripts prepared in advance in case the other person tries to derail the asserter from his assertion goal.

How do these self-instructional programs work?

Meichenbaum (1977) assumes that self-instructional programs work the same way that inner speech works in guiding our thoughts and actions. They act both as discriminative stimuli for overt actions and as motivators for them. These inner commands derive some of their force or controlling power from their resemblance to external commands from authority figures. They are most likely to be followed if the person believes they will help him. It is notable that in spontaneous behavior, inner speech is often evidence that the person is either making up a plan or is executing one. If a plan requires special concentration or novel calculations, as when a person is stymied while solving a problem, then the inner-speech gets louder and breaks out into overt self-talk. Older children, or those more adept at problem-solving, can often be seen talking to themselves while working on puzzles; the younger, poorer solvers do not engage in nearly so much self-talk, certainly not self-talk that is task-oriented.

Self-instructional plans can serve different functions depending on the task. Suppose the task is to endure the mounting pain from ischemia that results from cutting off blood flow to an arm by a pressure cuff. As reported in Meichenbaum (1977, p. 173 ff.), Dennis Turk, one of Meichenbaum's students found that people could nearly double their endurance of ischemic pain if they practiced a coping program of self-talk, telling themselves to relax, to breathe deeply, to think about totally different things, especially pleasant scenes incompatible with suffering, to transform the pain sensation into something nonaversive, and so on. We can describe these strategies as a set of productions in which the actions have the effect either of concentrating attention on neutral events, or occupying short-term memory with some demanding task (like counting backwards by 7s), or loading up short-term memory with words or images evoking feelings incompatible with suffering. These methods are most effective when the subject believes they will be; and the best evidence for their efficacy is a past history of successful endurance of pain.

A plan in working memory can be thought of as a cognitive strategy for dealing with a specific task. Some examples with which I am familiar are deliberate mnemonic devices which are plans for memorizing; that is, they are strategies for improving learning or for retrieving learned materials. We know that the development of mnemonic strategies underlies some of the improvement in a child's memory as he matures. We also know that teaching old people to use mnemonics materially aids their otherwise failing memories.

As another example of strategic plans, Mischel (1974) has investigated children's strategies for delaying gratification, in a situation where they get a small reward immediately or can choose a more favorable reward if they are willing to wait several minutes. The child can be taught various strategies for what to do in order to tolerate the delay. Almost any strategy which diverts attention away from the rewards will increase the child's ability to delay. Such diversions fill up working memory with non-food-related activities, so that the production in memory which says "If a sweet's available, grab it right now" will be blocked and not activated; thus, the distracted child can delay. Mischel and Baker (1975) have also shown that having a child transform a marshmallow mentally into a nonfood object like a cloud or a cotton ball will help him wait longer to get it; on the other hand, if the child really thinks about the chewy, sweet, soft taste of the marshmallow, he can't tolerate the delay—he wants it right now. So it is the reward stimulus that is in the child's head—not on the table in front of him—that determines his ability to delay. In each of these cases the strategy is a program that is dumping into short-term memory thoughts that either interfere with or facilitate the production (in long-term memory) which says "If candy's available, eat it now."

We thus see how productions can mimic psychologically interesting phenomena. Notice that these plans are not in-born; they are learned strategies or software, and they differ across individuals, ages, cultures, and so on. A lot of schooling is involved with teaching plans, with teaching recipes for doing cultural activites in a competent manner.

LEARNING

I have been discussing the architecture of the cognitive system, its alleged parts, and what plans are. I would now like to discuss learning or conditioning, which can be discussed at several different levels. First, I will discuss the learning of new facts in terms of adding new links into pre-

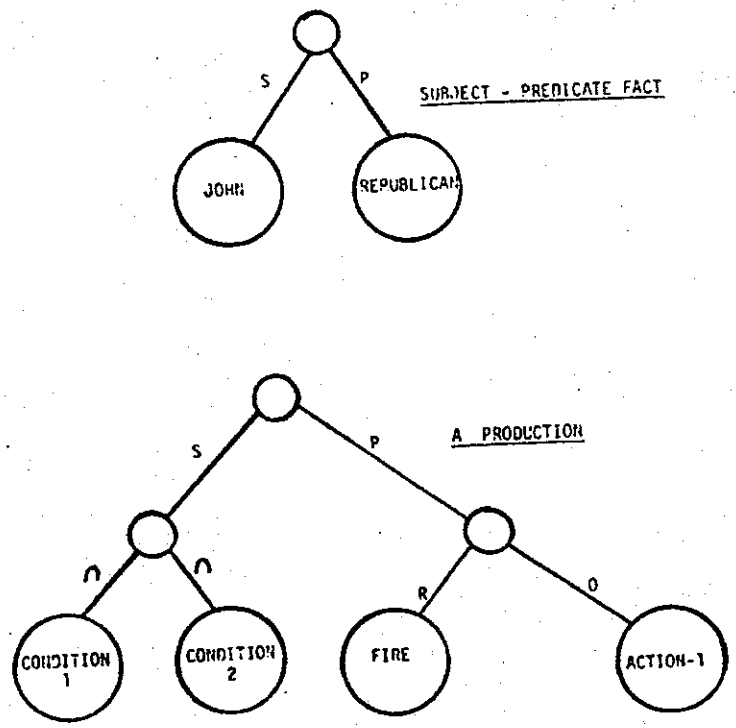

Fig. 8. Examples of recording in memory the content of a simple subject-predicate (S-P) assertion (top). At the bottom is the representation of a production which says that joint presence of condition 1 and 2 is an element of the set of things that fire action 1.

existing semantic networks. Then, I will discuss the cognitive view of standard learning situations such as conditioning and observational learning. I will conclude by noting some commonalities of cognitive psychology and social learning theory.

Learning in Semantic Networks

So to begin, recall that a semantic network consists of concepts connected together by associative relations. In such a system, learning a new fact consists of creating and recording a new configuration of links between concepts that we already know. Some examples are shown in Figure 8. Thus, I might have a concept node in memory for John Jones, and another concept node for Republican, and when I am told "John is a Republican," my linguistic parser adds a new link between the two concepts that is

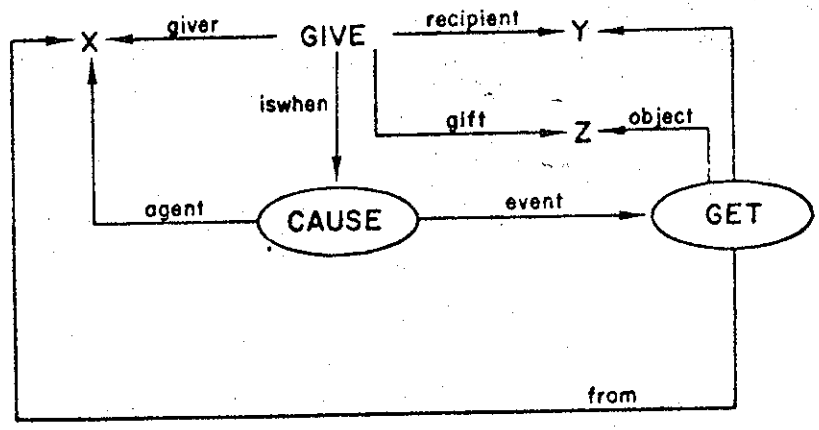

Fig. 9. Schema for GIVE is when a giver X causes an event where a recipient Y GETS a gift Z from X. Reprinted from Rumelhart and Ortony (1977) with permission.

equivalent to a subject-predicate construction. Such a subject-predicate construction is asserting that John is an element of the predicate-set, of Republicans. At the bottom of this figure is the associative configuration for a production—it means that joint presence of conditions 1 and 2 in short-term memory is one of the things that cause Action 1 to fire off. Now, the basic hope is that by using these kinds of simple links we can build arbitrarily complex data structures. But in order to deal with more complex events and assertions a conceptual memory needs several semantic relations and more complex interrelations among its parts. There are several proposals being debated at present, but I will present one particular line here, called schema theory. A "schema" is about the same thing as a stereotype; it refers to a large data structure that represents a generic concept in memory and which specifies the interrelationships that generally hold among parts of that concept. An example of a memory schema is shown in Figure 9. This is the schema for GIVE, as in "John gives Mary a book." GIVE is connected by relations to a set of variables or roles which are to be filled by the actors and objects in any event of GIVING. Thus X denotes a role to be filled by the *giver* who is the agent of an action which causes the *recipient* Y to get a *gift* X. Thus, when a person hears the sentence "John gives Mary a book," the verb calls up the GIVE schema from his memory and he proceeds to fill in the roles X, Y, Z with associations to the concepts of John, Mary, and book.

According to schema theory a perceiver understands a given stimulus event or sentence if it selects from memory a familiar schema which can be fitted onto the event and which seems to account for its significant features. The filling of slots in a schema creates an associative structure in

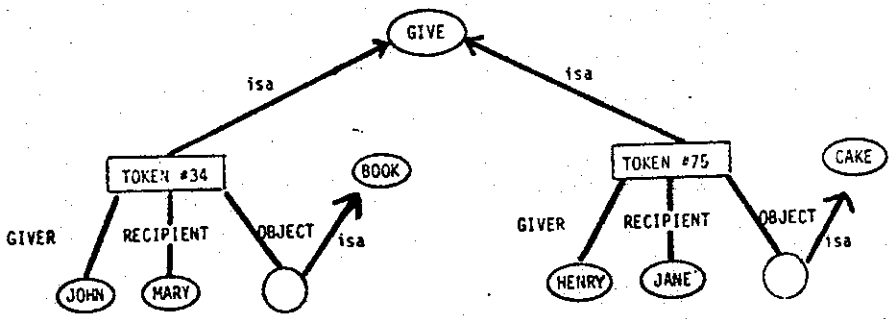

Fig. 10. Examples of two events recorded as tokens of GIVE events with specified givers, recipients, and objects.

the brain that represents the memory trace of that event. Figure 10 illustrates how memory would record a couple of events like "John gave Mary a book" and "Henry gave Jane a cake." Each input causes a new instance of the GIVE schema to be created with labeled links to preexisting concepts. When examining such data structures the interpretive system knows how to insert the designated referents for giver, recipient, and gift into the GIVE schema to reconstruct the several events.

We may think of each of these links as an association that varies in its strength, increasing with the frequency and duration that the person has thought of that relation, and decreasing with the time since he last thought of it. Our memory contains many thousands of such associative configurations, with single concepts connected into many different memories. How then do we retrieve a particular fact? The system retrieves information by undertaking an associative search from entry points into memory specified in the question. A question like "What did John give to Mary?" will set up a production which looks for a match in memory to a data structure having John as agent giving something to Mary as a recipient; if a match is found, the name of the most recent gift fills in the slot in the question, thus causing the production to fire off the answer, "a book." Similarly, if asked "What did Henry do?," the system looks for a match to memory structures in which Henry is the agent, and then recalls the most recent one.

Such a theory predicts fragmentary memories of events, since some associative connections will be available for recollection, whereas others will have been forgotten. Thus, in a few days, this person may remember that John gave Mary something, but he will forget what it was. Such a memory system will also show forgetting due to associative interference: That is, if John gives Mary many different things, then the "book" association will tend to be buried under its competitors and will perhaps not be retrievable.

The schema theory implies that our memory for persons or stereotyped scenes like a fight or a car accident will consist largely of a general schema along with a few details and deviations from the stereotype. As we forget the details, we use the stereotypic values of the schema to guide our recall, so that our recall becomes conventionalized. Thus, when a clinician is remembering the case history of, say, a lesbian, he is likely to distort his recall of details of her early life in a direction making them predictive of impending lesbianism.

These remarks show how schema theory conceives of the accumulation of knowledge in memory, how that knowledge is stored, retrieved, sometimes forgotten, and distorted. I will now move on to my next topic, which is the analysis of conditioning and standard learning experiments.

So far as I can see, social learning theory and cognitive psychology agree on their analysis of learning and behavior change. Both approaches view the human being as a knowledge accumulator who is learning about correlations among events in his environment. Standard laboratory learning experiments are viewed as settings which expose the person to controlled opportunities for acquiring information about correlations between environmental events or between his actions and environmental outcomes.

Classical or Pavlovian conditioning is viewed as the acquisition of an association between the internal representation of the conditioned stimulus, or CS, and that of the unconditioned stimulus, or US. Occurrence of a biologically important US causes the contents of short-term memory to be rehearsed and consolidated so that symbols in short-term memory at the time of the US will normally enter into association with it. This rehearsal process can be initiated in human subjects voluntarily by instructions, which are themselves productions in working memory. It is important to note that the CS need not be close in time to the US; but they must be grouped together in short-term memory at the time the rehearsal routine is set going. After the person abstracts out or discriminates the relevant cue, the occurrence of the CS causes him to think of, expect, or image the US, and that evokes part of the affective and behavioral reactions controlled by the US given the subject's state of deprivation. The associative structure so established is shown in Figure 11, which represents the fact that bell is followed by the arrival of food. The internal representations of the CS and US have corresponding concepts or images attached to them, so that thinking about or imagining the CS will lead to some activation of its conditioned response. This imaginal route is used in imaginal desensitization and in covert conditioning procedures. The imaginal procedure is not fully effective because an image of a situation does not fully reinstate all the requisite features of the real-life scene.

In operant or instrumental conditioning the person acquires a three-term contingency, namely, in antecedent situation A response B

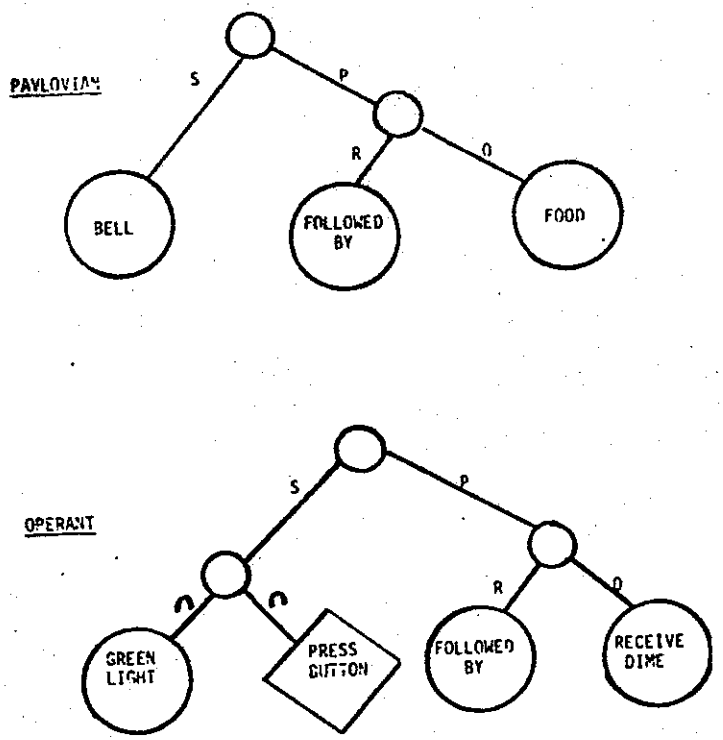

Fig. 11. Subject-predicate linkages encoding the facts of a Pavlovian experiment or an operant-conditioning experiment.

results in consequence C. These situation-action-outcome units are conditional rules of the form "When this green light comes on, if I press this button, I will get a dime." This is shown at the bottom of Figure 11, where the diamond around "Press button" means it is an action command as well as a concept. There are several comments to be made about this view of matters.

First, there is every reason to believe that from the organism's point of view the temporal correlation being detected is a *causal* relation. A person comes to believe that his response is causing the outcome, either directly or indirectly through some responsive agent. As evidence, we may note that people will make up superstitious beliefs to rationalize an apparent causal connection even when their actions and outcomes are in fact not related at all. Conversely, we know that if we tell a person that there is no causal relation whatsoever between his behavior and various outcomes, he will learn practically nothing about any correlation that the devious experimenter may build into the schedule of events. This is part of the message of Edward Thorndike's (1931) early studies on the influence of

what he called the "belongingness" of a reward upon the strengthening of response it followed: If the subject were set so that he did not group the reward as belonging with or being caused by the response, then he learned nothing about the contingency. More recently, the learned helplessness results of Maier and Seligman (1976) suggest that lower animals can be trained to believe that there is no causal connection between their responding in a given situation and termination of electric shocks there. The animal comes to believe that the shocks come on and go off independently of his attempts at escape, and so he appears to become passive and helpless. Similar results occur with human subjects in analogous situations.

I said that this contingency is represented in memory as a conceptual knowledge structure. What does that mean? To call it a knowledge structure is to recognize that there are a large number of ways we could teach that contingency to the person, and a large number of effective actions that that knowledge could mediate in altered circumstances. For example, a person can learn the "button press-yields-dimes" contingency by performing the actions himself, by observing someone else experiencing the contingencies (that is, by modeling), or simply by being told the contingencies. Such instructions would instill the relevant knowledge in symbolic form. Also, once that knowledge structure has been acquired, it can mediate a number of behaviors—not only will I press the button with my right hand when I want dimes, but I can carry out the act in other novel ways; I can answer verbal questions about it; I can hire someone else to do it for me, I can instruct someone else on the contingency, and I can instantly stop responding when told that the dime dispenser is broken. Because there are so many ways to teach the contingency and so many ways to utilize it, it seems most parsimonious to say that the contingency comes to be represented as a knowledge structure rather than as a simple habit relating a specific stimulus to a specific response.

As a human subject learns, he becomes aware of the contingencies and only then does he begin to respond appropriately, provided he values dimes and the experimenter's good will. Brewer (1974) and Dulany (1968) have recently reviewed the critical evidence regarding the role of awareness in human conditioning. The overwhelming conclusion from those reviews is that subjects who appear to condition are those who first figure out the contingencies, either exactly or with a partially valid hypothesis, and then they respond according to their expectations if they want the outcomes available. "Figuring out the contingencies" here means becoming aware of and able to recognize the stimulus-action-outcome correlations programmed into the situation.

I have now discussed Pavlovian and operant conditioning, which involve learning by direct experience. The third main route of learning is vicarious learning where we pick up information by observing the events

happening to someone else, the model. We know a lot about how people learn event correlations this way.

In a series of studies, Bandura (1971) has extensively analyzed conditions influencing observational learning and imitative performance. Bandura divides the processes conceptually into, first, those affecting attention to the model, second, the encoding of the modeled response sequence into memory and its retention over a delay interval, third, the skillfulness in executing the behavior and fourth, reinforcement and motivational factors that influence the whole chain.

Each of these theoretical processes corresponds to large bodies of theory and research in cognitive psychology. To single out just one for discussion, the encoding of a modeled behavior sequence into memory involves two symbolic representational systems—a verbal one and a sensory-image one. To elaborate on this, during exposure to the modeled response sequence, corresponding sensory experiences occur and will become interassociated or integrated by temporal contiguity. Later the revival of this sensory imagery can guide the observer's behavior during imitation. The second representational system is verbal coding, and we know it can be very effective. If the observer knows how to categorize single actions and has a label for each, then he can describe the model's acts as they occur; he can then rehearse and learn this verbal translation of the action sequence. Revival of the verbal sequence later as inner speech will cue successive response to be performed in imitation.

We know that the learning of a response sequence will be easier if the observer has already learned how to segment perceptually the series into meaningful chunks which he can classify and label verbally. For example, if one watches a model perform a series of hand signs in the American Sign Language for the deaf, he will be better able to code and reproduce that series if he knows American Sign Language (see Gerst, 1971). The same would be true for imitating ballet dance steps, or replaying an endgame in a chess match, or reduplicating a molecular structure out of tinker-toy atoms.

In one sense, a modeled sequence of events is not too much different from a laboratory study in verbal learning in which a subject observes a series of items exposed by a memory drum and later recites them back. As in the American Sign Language example, items that are familiar and meaningful will be better learned than unfamiliar nonsense because, for the knowledgeable person, they contain less novel information to be learned.

We can predict that variables effective in laboratory verbal learning experiments would also be effective in observational learning. Let me suggest some which, so far as I know, have not been researched. First, one should observe a serial position effect in reproduction of a long series of unrelated acts. Second, a series of unrelated arbitrary acts should be learned more poorly than an action series in which each act enables and naturally flows into the next action, giving it a natural coherence. Third, coherent action series which have a conspicuous goal or end-state towards which each action is directed will be learned better than will action series which are coherent but not conspicuously goal-directed. Of particular interest would be how the imitator responds when the model appears to be endorsing one goal in his verbal behavior but is obviously pursuing an entirely different goal in his overt behavior. Do observers do as we do, or do as we say? Is any advice nullified if it appears two-faced? Fourth, we should be able to disrupt retention of a given modeled sequence by having a second model perform other sequences which the observer must also attend to and learn before he tries to recall the first one. This associative interference among several modeled sequences would increase the greater the similarity of the models, the settings, and the action sequences themselves. Also, an action series that is seen should be interfered with somewhat by a similar action series that is merely named verbally for the subject but not seen. Fifth, a learner should be able to recognize a modeled sequence even though he could not recall or reproduce it himself. Sixth, if the action sequence is shown from the subject's point of view, as though the observer were sitting inside the model's head, the observer should be able to reproduce the actions more faithfully than if he watches the model as an onlooker. Seventh, if there is a cultural stereotype for the action sequence (e.g., a standard sequence for eating at McDonald's restaurant) and the model deviates from it slightly, the observer will assimilate the modeled actions to the cultural script, he will forget the deviations, and will conventionalize his reproduction of the actions.

Those are a few suggestions of variables that ought to affect observational learning and performance. I mention them to show that experimental studies of imitative learning make close contact with work in cognitive psychology. Also, in theoretical discussions Bandura talks about observational learning in terms of coding, symbolic representations, imagery, planning, and action-evaluations—all of which are cognitive concepts. It seems to me that social learning theory is a form of cognitive psychology that is being applied ingeniously to issues of socialization, to personality development, psychopathology, and behavior modification. I think it is time for people to see that behavior-modification technology could just as well rest on cognitive psychology as on S-R conditioning theory. Certainly, in comparison to S-R conditioning theory, cognitive psychology provides a more facile account of many recent therapeutic techniques such as covert modeling, thought control, self-instructional training, cognitive restructuring, persuasion, and—as Brewer (1974) argues—perhaps even the more direct operant conditioning procedures themselves. Since social learning theory and cognitive psychology seem to have been sleeping together for years, I think it is

high time that we announce their up-coming wedding. I don't think we'll need a shotgun to get them to the church on time.

REFERENCES

Bandura, A. *Psychological modeling: Conflicting theories.* New York: Aldine-Atherton, 1971.

Bandura, A. Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review,* 1977, *84,* 191-215.

Beck, A. T. Thinking and depression. II. Theory of therapy. *Archives of General Psychiatry,* 1964, *10,* 561-571.

Bower, G. H. Cognitive psychology: An introduction. In W. K. Estes (Ed.), *Handbook of learning and cognitive processes* (Vol. 1). Hillsdale, N.J.: Lawrence Erlbaum Associates, 1975.

Bower, S. A., & Bower, G. H. *Asserting yourself: A guide for positive change.* Reading, Mass.: Addison-Wesley, 1976.

Brewer, W. F. There is no convincing evidence for operant and classical conditioning in humans. In W. B. Weimer & D. S. Palermo (Eds.), *Cognition and symbolic processes.* Hillsdale, N.J.: Lawrence Erlbaum Associates, 1974.

Cantor, N. E., & Mischel, W. Traits as prototypes: Effects on recognition memory. *Journal of Personality and Social Psychology,* 1977, *35,* 38-48.

Cantor, N. E., & Mischel, W. Prototypicality and personality: Effects on free recall and personality impressions. *Journal of Research in Personality,* in press.

Chapman, L. J., & Chapman, J. P. The genesis of popular but erroneous psychodiagnostic observations. *Journal of Abnormal Psychology,* 1967, *72,* 193-204.

Chapman, L. J., & Chapman, J. P. Illusory correlation as an obstacle to the use of valid psychodiagnostic signs. *Journal of Abnormal Psychology,* 1969, *74,* 272-280.

Dulany, D. E. Awareness, rules, and propositional control: A confrontation with S-R behavior theory. In T. R. Dixon & D. L. Horton (Eds.), *Verbal behavior and general behavior theory.* Englewood Cliffs, N.J.: Prentice-Hall, 1968.

Gerst, M. S. Symbolic coding processes in observational learning. *Journal of Personality and Social Psychology,* 1971, *19,* 7-17.

Maier, S. F., & Seligman, M. E. P. Learned helplessness: Theory and evidence. *Journal of Experimental Psychology: General,* 1976, *105,* 3-46.

Meichenbaum, D. *Cognitive behavior modification.* New York: Plenum Press, 1977.

Mischel, W. *Personality and assessment.* New York: Wiley, 1968.

Mischel, W. Processes in delay of gratification. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 7). New York: Academic Press, 1974.

Mischel, W., & Baker, N. Cognitive transformations of reward objects through instructions. *Journal of Personality and Social Psychology,* 1975, *31,* 254-261.

Newell, A., & Simon, H. A. *Human problem solving.* Englewood Cliffs, N.J.: Prentice-Hall, 1972.

Rumelhart, D. E. *Introduction to human information processing.* New York: Wiley, 1977.

Snyder, M., Tanki, E. D., & Berscheid, E. Social perception and interpersonal behavior: On the self-fulfilling nature of social stereotypes. *Journal of Personality and Social Psychology,* in press.

Thorndike, E. L. *Human learning.* New York: Century Co., 1931. (Paperback edition, Cambridge, Mass.: MIT Press, 1966.)

Cognitive Therapy and Research, Vol. 2, No. 2, 1978, pp. 147–164

Type of Cognitive Modeling, Imitation of Modeled Tactics, and Modification of Test Anxiety[1]

Monroe A. Bruch[2]
Bradley University

Effects of two types of cognitive modeling (problem-solving rules vs. coping self-instructions) on laboratory test anxiety were examined in two experiments. Both studies included assessment of observers' cognitive processes while working anagrams to evaluate modeling effects on cognitive tactics and to determine how they related to subsequent behavior. Experiment 1 investigated the influence of type of modeling and model reward on performance. Experiment 2 served to replicate any modeling effects and used a pre-post assessment of observers' strategies to clarify how modeling operated. Results from both studies offered some support for the multiple influence of cognitive modeling procedures over a model control, but were equivocal regarding the relative superiority of the two types in facilitating anagram performance. Modeling effects were limited by unequal difficulty among anagrams and perceived differences in relevance of particular tactics providing mixed support of Sarason's research. Assessment of observers' cognitive processes suggested that negative self-talk was task interfering and imitation of solution rules but not coping instructions was associated with quicker solution. Pre-post analysis revealed that college female observers possessed a repertoire of solution rules similar to modeled rules and that modeling acted to increase subsequent use of tactics in contrast to controls.

There is increasing awareness that modeling and observational learning should not be equated with simple matching of a model's overt behavior. When designing modeling treatments workers often directed their attention to the specific overt behaviors to be demonstrated, and overlooked the

[1]This research was sponsored in part by a small grant from the Bradley Board of Research, Creative Production, and Sponsored Programs, Bradley University.
[2]Address all correspondence to Monroe A. Bruch, Bradley University, 134 Bradley Hall, Peoria, Illinois 61625.