CHAPTER 3

Gordon H. Bower
Stanford University

MENTAL IMAGERY AND ASSOCIATIVE LEARNING*

I wish to discuss some of our research on mental imagery in associative learning. I think there are some fascinating but difficult intellectual and scientific puzzles connected with imagery, problems whose solution would significantly advance our understanding of mental life.

My discussion will be divided into two parts—the first predominantly philosophical and phenomenological in tone, and the second within the proper confines of experimental psychology. Although I am an amateur at it, I still think that some philosophical discussion of imagery is necessary, and mainly for therapeutic reasons, because many experimental psychologists cannot entertain thoughts about imagery without some deep sense of guilt associated with forbidden taboos. Our fraternal indoctrination that imagery is the forbidden fruit has been handed down to us, of course, from the heydays of radical behaviorism, which consigned it to the flames along with other cognitive concepts. Because speaking and writing were observable behaviors of a person, words or utterances were let pass as mere motor responses. Today, the domain of studies of human memory is almost coextensive with verbal learning and motor-skills learning.

Now, whether or not psychologists wish to admit it, the radical behaviorism of John Watson was a philosophical doctrine. More specifically, it contained an implicit epistemology about how we acquire reliable knowledge of the world—in particular, knowledge about other persons. So I think the defense of imagery also has to be partly philosophical, with a healthy sprinkling of subjective experience. In much of this discussion, I am drawing from arguments in a book by Brian Smith (1966).

THE NATURE OF MEMORY IMAGERY

The primary distinction I wish to draw between types of cognitive memories is remembering in imagery versus remembering in propositions. Remembering is

*This research was supported by a grant, MH-13950, to the author from the National Institutes of Mental Health. Thanks are due several students who helped conduct these experiments. They include Michal Clark, Samuel Bobrow, David Tieman, David Wintenz, Alan Leggold, Laura Bolton, and Michael Fehling. Their assistance was indispensable in the preparation of this chapter.
a performance term, and many performances involve one or another of our motor skills (e.g., speaking, writing). When we are reporting or acting upon what we are remembering some of our learned motor-output skills are likely to be engaged. But we just as frequently remember something without reporting it, so the overt motor-output skills need not be engaged. So with this nod to motor-skills, let us turn to remembering in imagery and how it differs from remembering in propositions.

How Versus What

The function of memory imagery is to put us in direct contact with how things looked, or sounded, or felt, or tasted, as distinct from what they resembled, what they sounded like, looked like, felt or tasted like. This "how something looks" is called by philosophers its appearance. And the claim is that imagery is our only way of remembering appearances. Said differently, memory imagery is our means of representing to ourselves the appearances of past events we have witnessed.

This distinction between how something looked and what it looked like runs parallel to the distinction between images and propositional memory. That is, we remember appearances in imagery, and we also remember propositions, propositions sometimes containing higher-order perceptual inferences about events witnessed.

The difference between how something looked and what it looked like is the same as the difference between a sighted and a blind person's knowledge of the visual world. In the auditory domain, it is the difference between how an orchestral symphony sounds and how one might try to describe it to a deaf person. William James (1890) put it as follows: "The best taught blind pupil...yet lacks a knowledge which the least instructed seeing baby has. They can never show him what light is in its 'first intention'; and the loss of that sensible knowledge no book learning can replace." This "first intention" is what is meant by how something looked or sounded.

The unfortunate difficulty with this distinction is that our descriptive terms are adequate only for describing what something looks like, not for how it appears. One resorts to citing cases which exemplify the difference between how something looks or sounds and a catalogue of descriptive propositions concerning what it looks or sounds like. For example, most of us know how an oboe or a clarinet sounds—that is, we can recognize their sounds, we can have auditory imagery in remembering them, and we can compare them in memory—but our verbal descriptions of such sounds are woefully meager, impoverished, and metaphorical. A second example is that one's chances of apprehending a wanted criminal are very much better given a few photographs of him compared to a book-length, verbal description of him. Part of

the problem here is to describe the color of the light (e.g., Gibson, 1965) or the very broad category of unique perceptual content is a many-one mapping from stimulus. For example, color identifies, yet it makes it hard to judge, whether the Christmas tree is real or in memory imagery. The problem needs to be considered further, that it is structure of the image.

Roger Shepard (1970) completed experiments on U.S. states (e.g., when he was planned, multiple operations of instruments, but actually heard when he remembered). Given a matrix, there exist for extrastriate areas, Euclidean space, approximately the same space as the imaging condition. Shepard showed that imaged shapes, locations will also represent events. The ability of auditory imagery to represent past perceptions of the subject has pronounced similarities to judgments and mental report.
the problem here is simply an absence of appropriate verbal terms to label or describe the vast array of stimulus variables which we can differentiate (e.g., Gibson, 1966). But even when we have verbal labels, they are usually very broad categories, so that much stimulus information is lost when some unique perceptual event is replaced by a class label. A descriptive category is a many-one mapping or compression of the information available in the stimulus. For instance, The Science of Color (Optical Society, 1953) estimates that the color solid is divisible into about 7,500,000 just-noticeable differences, yet English speakers commonly use only about eight color names. But the many-one coding imposed by our color naming does not force us to judge, when remembering, that green peas are the same color as green Christmas trees. There need be no claim that our discriminations from memory images are as sharp as those from present appearances. All that needs to be claimed is that memory images contain some information, and that it is structured in ways resembling the structure of our present perceptions.

Roger Shepard at Stanford has done some work on this question. In one completed experiment, similarity ratings were obtained for the shapes of pairs of U.S. states either when the person is shown outlined pictures of the states, or when he merely remembers what they look like. In another study being planned, musically-knowledge subjects will give similarity ratings for pairs of instruments (e.g., oboe, bassoon, clarinet, etc.) either when the person actually hears the two instruments alternately playing a short passage, or when he remembers how they sounded when playing the same passage. Given a matrix of similarity measures for a set of stimuli, analytic techniques exist for extracting their underlying pattern, representing stimuli as points in Euclidean space (Shepard, 1962). The question is whether one gets approximately the same spatial configuration of the stimuli under the sensing and imaging conditions. Rapaport and Fillenbaum (1968) have shown that the configurational representations are about the same for sensed versus imaged colors. Shepard showed that the configurations are about the same for sensed versus imaged shapes of states. It is a reasonable presumption that the representations will also be about the same for sensed versus imaged sounds of instruments. If so, then one could assert that a set of memory images has a similarity structure isomorphic to the structure of the set of original perceptual events. The alternate interpretation is that the subject is not really comparing auditory images but is remembering the verbal judgment he made during his past perceptions. But this requires the implausible assumption that the subject has perceived all \( n(n-1)/2 \) instrument-pair comparisons in the past, pronounced similarity judgments at that time, and now remembers those judgments and transforms them to the appropriate scale for the experimental report.
Now strictly speaking, we do not remember events in the world; rather, we remember our autobiography—the appearances presented to us by events, our perceptual inferences, and propositions we formulated about past events as they happened or as we later revive them and rethink them. As a result, our memory for some event ordinarily includes both some imagery of it and some propositions about it. This makes operational separation of the two modes of remembering difficult, but I think there are sufficient instances where we remember predominantly in one way without the other. For example, much of my memory of my preschool childhood is propositional, remembering either my own judgments or propositions my parents told me about my early childhood. I remember that my father’s elderly aunt had a small wrinkled moonface, and that she once told a terribly ribald joke in my presence, but I have no remembrance or imagery of her. I am only recalling a proposition I formulated about her when I was very young. My knowledge in this case is about on a par with my knowledge that Brutus stabbed Caesar, even though I was a witness to the former but not the latter event. I could easily be argued out of either proposition by contrary evidence, because the memory is not supported in either case by remembered appearances.

Contrariwise, we can remember in imagery without propositions. Examples would be our memories of paintings, or ballet choreography, or photographs (Shepard, 1967), or the imagery of a person after a good trip on LSD, which he finds to be such an ineffable and noncommunicable experience. Certainly, our memory for musical melodies can hardly be called propositional.

Problems With a Purely Verbal Approach

In the visual cases above, it should be clear that I am distinguishing between the imagery in which a person is remembering, and the propositions or descriptive statements he might be presently formulating about what he is remembering. Radical behaviorism would not allow that distinction, insisting in fact that my present verbal utterances are my remembering, that one does not remember images but only earlier utterances or other responses made at the time a particular stimulus event happened.

I think this view is indefensible for several reasons. First, there is sometimes a strong introspective feel to the difference between remembering an appearance and remembering propositions about an appearance. Second, even though we might remember the appearance of some event and some propositions we formulated about that event, we occasionally do reinterpret that experience and alter our propositional memory about it in the light of new evidence or a new frame of reference. What we took to be a motorcycle policeman by the side of the road we later agree must have been a painted, wooden dummy set up to frighten speeding motorists. The banging night-time noises I take to be a prowler turn out to be a cat rummaging in the garden. Third, against the verbal doctrine of relating propositions when percepts fail to be present, it does not prevent at least some laboratory can learn a serial list, continuously engaged, say, in the verbal theorist to such categorization, categorization, and naming central mechanisms of linguistics is inhibited. At this point the advantage it might have had over any other hypothesis is operationally no better. The theorist is to say that whether in some scene, verbal description (Hovland, 1953). As a matter of fact, there are even those cases where it is completely alternate codings of an event may be employed (see Wallach & Averbah, 1958).

A final criticism of the strict propositional approach can only rarely be decisive. Although laboratory tasks are a feature of our day-to-day discourse, our ability to generate and recognize evidence suggests that we remember to the sentences by which they are indifferent even to the language (1966) examples is of a German who learns English and forgets most of his mother tongue, who forgets all the scientific propositions which he would still be able to, only making trite that we primarily remember elementary facts about linguistics.

The Imagist Theory of Cognition

A good deal of philosophical theorizing, cognition and thinking that words are derivative tokens associated with imagery. There are
Mental Imagery and Associative Learning

be a cat rummaging in the garbage cans. Such reinterpretations, double-takes on
our perceptual inferences, would be pointless if we remembered were prop-
sitions that we formulated at the time of the event. The validity of such reinterpre-
tations presupposes that we remember something else, which constitutes the
evidence to be reinterpreted. This "something else" is what I call the appearance
of the event.

Third, against the verbal descriptive doctrine, one is seldom aware of formulat-
ing propositions when perceiving, for example, a painting, but this nevertheless
does not prevent at least some recall of these appearances. Or a subject in the
laboratory can learn a serial list of pictures even though his vocal apparatus is
continuously engaged, say, in counting backwards by 3's. A standard defense of
the verbal theorist to such counterexamples is to postulate that implicit verbaliz-
ation, categorization, and naming go on in all such instances. That is, the same
central mechanisms of linguistic description are activated but the motor outflow
is inhibited. At this point the verbal approach loses whatever operational advantage
it might have had over an imagery approach. Unconscious, inhibited verbaliz-
ations are operationally no better than images. A final move by the verbal
theorist is to say that whether or not verbalizing is necessary for remembering
some scene, verbal description surely improves one's memory for it (cf. Kurtz &
Hovland, 1953). As a matter of fact, I doubt whether this is generally true; but
even those cases where it is correct raise no problems for my position, since two
alternate codings of an event may make it more memorable than one coding does
(see Wallach & Averbach, 1955).

A final criticism of the strictly verbal approach is that our memory of a propo-
osition can only rarely be described as a memory of a verbal utterance or a sen-
tence. Although laboratory tasks often demand verbatim reproduction, a salient
feature of our day-to-day discourse is our competence at paraphrasing, our
ability to generate and recognize roughly synonymous expressions. Such casual
evidence suggests that we remember propositions that are reasonably indifferent
to the sentences by which they are expressed. For bilinguals, the propositions
are indifferent even to the language in which they are expressed. One of Smith's
(1966) examples is of a German scientist who immigrates to America, learns
English and forgets most of his German. He would not thereby be expected to
forget all the scientific propositions he had learned in German. Most of them
would still be available, only now they would be in English. The behaviorist doc-
trine that we primarily remember verbal utterances has not faced up to the most
elementary facts about linguistic competence.

The Imagist Theory of Cognition

A good deal of philosophical analysis has gone into the view that all our concep-
tualizing, cognition and thinking is in mental imagery. For example, one tenet is
that words are derivative tokens whose "cash values" are a capacity for arousing
associated imagery. There are a number of difficulties with this imagist view of
cognition. One problem is that many concepts (e.g., animals, flowers) are names for heterogeneous instances, so that an image of any particular instance could not possibly represent the concept in all its applications. Another problem is that there are no criteria for deciding which of several superordinate classes an image of an instance may be representing. Does an image of a dachshund represent the class of dachshunds, dogs, mammals, animals, or living things? A third problem is raised by abstract concepts like infinity, interpretation, and so on, where imagery is either absent, or is at best metaphorical and nonreferential.

One avoids all such problems by admitting that we have concepts which have nothing to do with images. We infer that a person possesses a concept, for example, of dogs, when he can recognize instances. But we cannot have an image of a generalized DOG, only of numerous individual dogs. Similarly, I have no image of a generalized Lee Gregg, but only several images of his appearances in several contacts we have had.

The old imagist view that a concept is an image bears a strong resemblance to the template-matching view of pattern recognition, which we know to be weak for several reasons. I would say that images can exemplify but not substitute for concepts. For example, I can have no imagery of relational concepts like above or between; but I can image a spatial organization of objects exemplifying such relations.

I am proposing that we can assign an important role to imagery in cognition and memory, without subscribing to the entire imagist view of such matters. Let us now consider some of the functions of imagery in our thinking.

Imagination Imagery

One of the primary uses of memory images is in productive imagination. Such imagery may be characterized as a new assemblage of memory images; that is, the recombining of memory images into scenes which the person has not witnessed. For example, I can remember Al Newell in imagery and I can remember an elephant in imagery, so I can also image a scene in which Al Newell is riding on the back of an elephant, even though I know that this does not correspond to any actual scene I have witnessed. We construct these novel assemblages by using our memory images as the basic components.

This imagination game resembles a picture-blocks game my children have, in which one independently selects a picture-block with a mouth, another for the nose, another for the eyes, ears, hair, neck and so on, all put together to make a composite but very bizarre face. Our imaginative images differ from the blocks in being less determinate but more manipulable in shape, size, color and perspective. The wings we imagine on a flea are a different subjective size than those we imagine on an elephant. In some people's visual imagery, geometric forms (e.g., a parabolid or a box) can be made to undergo deformations and perspective transformations like rotation, translation, etc., as if the observer were walking around an indeterminate object.

In such instances, imaging is connected to a computer program or Green's (1961) or Roberts' (1965, 1967). The manipulability of imagery is often used to manipulate diagrams, graph structures, bit strings, etc., and we manipulate these within the context of imagery in this respect is part of our mental memory. Our imaginal scenes or simulations at any one time.

To mention one further role of imagery, it can serve as a supplementary aid in comprehending a statement is true, or to make it true. This view is not Tractatus, that a proposition of Pictorialism, or imagining it is not a necessary condition for formulativeness that requires no imagery. Furthermore, although imagery can be indeterminate whether the conclusion after, or was noncausally connected to the ability to generate appropriate sentences that require no imagery.

Characteristics of Imagery

Our memory imagery, like our imagination, can be more or less precise, detailed, and elusive. Moreover, even when the imagery and associated context; that is, the way in which the imagery is seen, can be the principal way in which we be the given memory. Old crones will often reproduce details associated with some of these reconstructions.

Introspectively, the memory imagery is in its perspective, in the relative details may be colored. Frequently we are in the indeterminate ground until we...
Mental Imagery and Associative Learning

transformations like rotation, reflection, and shearing, as though the observer were walking around and/or changing his distance from a three-dimensional object.

In such instances, imaging is being used like the flexible, on-line CRT displays connected to a computer program such as Sutherland's (1963) "Sketch Pad" or Green's (1961) or Roberts' (1965) programs for rigid 3-D rotation of figures. The manipulability of imagery is one of its chief advantages in thinking. We often use tangible replicas or models as an aid to our thinking. We draw diagrams, graph structures, benzene rings, build solid models of DNA molecules, etc., and we manipulate these tangible tokens in thinking. The principal limitation of imagery in this respect arises from the limited capacity of our short-term memory. Our imaginal sketch pad can be focused on just a few symbols or manipulations at any one time.

To mention one further role of imagery in thinking: it can sometimes serve as a supplementary aid in understanding novel sentences. One way to comprehend a statement is to imagine some situation or state of affairs which would make it true. This view is the picture-theory of Wittgenstein's (1922) Tractatus, that a proposition corresponds to a picture (true or false) of reality. Pictoriality or imaginability of a sentence may be a sufficient but is clearly not a necessary condition for comprehension, since we use millions of meaningful sentences that require no imaging to understand them (e.g., this sentence).

Furthermore, although imaging may ensure comprehension, it would be indeterminate whether the click of semantic comprehension came before, after, or was coincident with the imaginal construction. That is, the ability to generate appropriate imagery may sometimes be just one of many behavioral dispositions that issue from comprehension of a linguistic proposition.

Characteristics of Imagery

Our memory imagery, like any picture or verbal description of an object or event, can be more or less precise. Faint memory images are sketchy, lacking in detail, and elusive. Moreover, faint memories cannot readily be expanded into an associated context; that is, it is difficult to fill in other objects and happenings. This expansion of a memory image into its associated context is the principal way by which we build up subjective confidence in the validity of a given memory. Old cronies who meet at class reunions may pride themselves on reproducing details associated with some football game. It is notorious that some of these reconstructions are imagined.

Introspectively, the memory image of a visual scene is usually fairly faithful in its perspective, in the relative sizes and locations of parts of the scene, and it may be colored. Frequently only part of the scene is figure, the rest being indeterminate ground until we generate another part of the scene.
To say that a person is remembering an event in imagery is to say that some central mechanisms are generating a (probably sequential) pattern of information which corresponds more or less to the structural information in the original perception. Imaging is a generative process and a performance. There is a very strong tendency in our everyday language to "substantiate" our activities, to talk about a performance as though it resulted in a tangible entity. Performance verbs tend to become nouns. We run a long run, breathe a deep breath, perceive a clear percept, and think a bright thought. Imaging is likewise a performance verb: one "has" images. This use is quite innocent and no one need be seriously misled into believing that there is a palpable entity located in some definite place, that it can be weighed, and so on. To say that we are remembering in imagery is to say that there is some structural isomorphism between the information presently available to us and the information picked up from the stimulus event we are remembering. Needless to say, this structure need not be apparent in the person's overt response, which may simply consist of his answering "yes" to the question "Are you remembering X now?"

Problematical Issues With Imagery

The common man says that imaging is very much like constructing mental pictures. His critics are fond of pointing out how unlike photographs or tape recording images are. Images surely do have a large degree of indeterminateness to them, much like impressionist paintings. People can visualize a zebra but not be able to count its stripes; or rehearse the twangy, nasal voice quality of an orator without recalling any particular part of his oration; or visualize the front of a building but not know how many stories it has. This is all quite true; but to disparage imagery in this way is unfair because most of our perceptions are similarly vague, indeterminate, and lacking in detail. A tachistoscopic flash of a building or zebra would produce similar inaccuracies. When the object is continuously before us, we move our eyes successively over it and count. But this we cannot do with our memory imagery.

Another issue concerns individual differences. The classic questionnaire study of Sir Francis Galton (1883) turned up some people who reported having no imagery. A more recent and adequate survey by McKellar (1965) of 500 British adults from a variety of occupations turned up none who reported no imagery; 97% reported availability of visual imagery, 92% had auditory imagery, and over half had a variety of sensory imagery available, including movement, touch, taste, smell, pain, and temperature. We may still wonder what to make of those occasional respondents who report no imagery whatsoever.

Smith (1966) lists three alternatives: they are liars, they have only propositional memory, or they have misunderstood the reference of the question. The first may be discounted; of the last two, I prefer the "misunderstanding"

Indicators of Imagery

Because imaging is customarily associated with the periphery of conscious awareness, psychologists are fond of mentioning moments in imagery. Psychologists are fond of mentioning moments in imagery. Certain events strike a listener as "imagery" because associated with minimal eye movement and following moving objects in imagination.
Mental Imagery and Associative Learning

Account. Psychologists are familiar with respondents' misunderstandings of self-descriptive terms. Respectable society matrons will deny giving vent to erotic impulses although their behavior and their husbands speak otherwise.

To indict misunderstanding is to say that the problem is one of semantics. Having imagery is a private affair, and Skinner (1953) has told us why it is difficult to teach people to have the same referent for names of private events. Since, by definition, I am not aware of my son's private mental event, I have to infer it either from an instigating stimulus or from his behavior. Either of these provides me, as the trainer, with an SD for applying the label or saying “You’re experiencing X now”.

Thus, I can teach him to label pain by saying “Ouch, that’s painful, isn’t it?” when I see him crack his shin, double up, and howl. Remembering in imagery has neither unique instigating stimuli nor unique behavioral signs. Learning to use imagery terms would therefore derive mainly from induction of stimulus generalization. Verbal labels acquired with respect to public events (where the reinforcing community can train us to sharp discriminations) are then transferred to private events on the basis of common properties. Thus, children learn the referent of words like picture and learn how to copy, or make up and draw pictures, and games are played with them where they close their eyes and try to see an imaginary picture. I presume this is how most of us came to learn imagery terms and it may account for the prevalence of “mental picture” descriptions. Most of our subjective terms are probably learned by induction from public terms, which may be why our subjective descriptions (e.g., of emotions) are predominantly metaphorical. In our metaphor talk, passion is a consuming fire, anger is a raging storm, and intellect or wit is a sparkling instrument that is piercing or dull.

Indicators of Imagery

Because imaging is customarily a private event, the scientist's first inclination is to try to achieve some sort of public measurement of it. Whatever nervous mechanisms generate imagery, they conceivably might also send efferent messages to the periphery where they can be recorded. The analytic leverage provided by such a reliable, public indicant should not be underestimated. For example, the tremendous surge of recent research on dreaming must in part be attributable to Dement and Kleitman’s (1957) discovery that rapid-eye-movement sleep is temporally correlated with subjective reports of dreaming. For such reasons, one searches for indicators of imagery.

The obvious first candidate is spontaneous eye movement during visual imagery. This has been studied in detail by Singer (1966), and Singer and Antrobus (1965). Their results are most curious: generally, visual imaging was associated with minimal eye movements except when the person’s task was to follow moving objects in imagery (e.g., an imaginary tennis game). Instructions
to engage in “hard thinking” (e.g., adding three-digit numbers) increased eye movements. Although there were minimal eye movements during passive daydreaming or fantasizing, numerous eye movements occurred when subjects were instructed to break off and stop their daydreaming. A possible interpretation is that one breaks off a given daydream by deliberately replacing that theme by a rapid succession of unrelated images, with corresponding eye movements. Although the technique requires further exploration, these preliminary results do not suggest that eye movements are a perfect indicant of imaging. Corresponding eye movements seem to be neither a necessary nor a sufficient condition for visual imaging.

A learning experiment of ours may be added here. College students learned pairs of concrete nouns, operating under instructions to generate visual images linking the objects denoted by the concrete nouns of each pair. Some subjects were told to close their eyes and scan the mental picture as they constructed it. Other imagery-instructed subjects were required to maintain a steady gaze fixated upon a small spot on the wall opposite them. The items were spoken to the subject, and eye-condition during study and immediate paired-associate recall were varied. This experiment found no recall differences whatsoever due to eye-condition either at study or recall test. Fixation subjects reported no difficulty in having images. This is not implausible; a notorious characteristic of the daydreamer is the far-away look in his eyes and the glassy stare.

Another possible indicator of imaging is eye-pupil size, which may index cognitive events. Pupils during imaging have been measured by Paivio and Simpson (1966) and Simpson and Paivio (1968). The pupil dilates when the person is told to image the object denoted by a concrete noun. The pupil dilates even more when the person is asked to get an image for an abstract noun. In fact, the pupil dilates whenever the person is given any sort of mental task (e.g., adding numbers), and the dilation seems correlated with the amount of mental effort involved. The results also depend on how the subject has been instructed to report accomplishing the cognitive task. Such results imply that pupil dilation may not discriminate imaging from other types of cognitive processing.

Further attempts to track down imagery have involved the electroencephalogram (EEG). This literature (Oswald, 1962; Simpson, Paivio and Rogers, 1967) is distinctly unencouraging. No one has yet turned up a distinctive EEG pattern that correlates with subjective reports of imaging. Kamiya (1969) taught people to discriminate and then to control alpha rhythm in EEG recorded from scalp electrodes. Blocking of alpha activity can be induced by imaging a familiar object. Alpha activity can be induced by letting go of deliberate conscious engagement, and maintaining a placid, receptive state of suspended cognition. However a variety of cognitive activities other than imaging will also block alpha rhythms.

In contrast to the unpromising outlook for physiological measures of imagery,
research on behavioral indicants of imagery (particularly visual imagery) has been more successful.

Behavioral Evidence of Imagery

If we have a person describe from memory a picture or some spatial arrangement of information, he will usually have the subjective feeling of generating an internal spatial representation (an image) and deriving his verbal description from that. If, on the other hand, he is asked to give verbatim recall of a sentence he has just heard, his performance seems to involve predominantly the speech mode with little or no internal spatial representations. Lee Brooks (1968) has performed some highly ingenious experiments to support this idea that visual (or spatial) and verbal information are dealt with in distinct, modality-specific manners.

Brooks' method hinges on the conflicting processing that may occur if the modality of the remembered material is the same as the modality in which the subject has to report his judgments. In one condition, the subject learned a short sentence, and then at a signal reported some particular categorization of the successive words in the sentence. For example, the sentence “a bird in the hand is not in the bush” might have to be categorized into noun versus nonnoun words, with the subject reporting “no, yes, no, no, yes, no, yes, no, no, yes.” Since different categorizations were requested, the subject could not anticipate the judgments to be required. There were two methods of report: either saying “yes” or “no,” or pointing to a “y” or “n” printed in a column of y-n pairs before the subject. When the remembered material was a sentence, the visually-guided pointing responses were much faster than the speech reports. In the other condition, the subject studied and learned a simple visual diagram, such as the block letter F. He was later asked for a particular categorization of each corner of the diagram. For example, the categorization might be whether each point is on the periphery or is an interior point. In this case, the speech reports were very much faster than the visually-guided pointing responses.

In several experiments, Brooks has found a similar interaction between modality of the memory and modality of the reports. His interpretation is that spoken sentences and spatial diagrams are processed in memory by separate systems (speech versus visual imagery). Requiring judgments in the same modality produces conflict or strains the capacity of that system. In another work, Brooks (1967) showed that visual reading conflicted with a task for which visual imaging was used, but not with a task for which visualizing was not used.

A recent experiment of ours followed Brooks' ideas. The basic task was paired associate learning of 30 concrete noun pairs read to the subject. Half of the subjects learned these by overt rote rehearsal of each pair, while the remaining subjects were instructed to connect the words of each pair by visualizing a scene of imagery,
of some vivid interaction linking the two objects denoted by the words. During study and recall testing of all pairs, the subject was continuously engaged in a subsidiary tracking task, which either did or did not engage the visual analyzing system. The tracking task involved an irregular wavy line which moved rapidly and continuously past the 1-inch window of a memory drum. The subject was required to track the horizontal location of the waving line by keeping it between the extended index and middle fingers of his right hand. In the visual tracking task, a clear plastic shield covered the window and only visual cues were available for locating and tracking the target. In the tactile tracking task, the wavy line was composed of string glued to the paper in the same pattern. The subject was instructed to close his eyes and to keep the raised string touching the sides of his extended index and middle fingers. Each subject had one study and one recall test trial on four different lists of 30 noun pairs. Two lists were learned during the visual task, and two during the tactile tracking task. (Task order was counterbalanced.) Subjects learning with visual imagery were expected to learn and/or recall more poorly with the visual than with the tactile tracking task. Subjects learning by rote rehearsal were expected to recall equally well under the two tracking tasks. The paired associate recall scores were in perfect agreement with these expectations. For visual imagery subjects, recall percentages were 66% under tactile tracking and 55% under visual tracking, a difference significant at the 2% level. For rote rehearsal subjects, recall percentages were 23% under both visual and tactile tracking. Although the effect here is not as large as one would like it to be, it was quite consistent over subjects and does agree with expectations.

Interference between imaging and perceiving in the same modality has been demonstrated by Segal and Fusella (1969). In a signal detection and identification task, their subjects were presented with a near-threshold visual or auditory signal (or neither signal) and had to report which event occurred. In different series of detection trials, the subjects were instructed to generate either a visual image (e.g., of a tree) or an auditory image (e.g., a ringing telephone) with the detection trial presented only after the subject had the image clearly. A control detection series was also run without subsidiary imaging. Measuring sensitivity in terms of the d' parameter of signal detection theory (which takes into account possible criterion shifts), Segal and Fusella found that although either type of imaging reduced d', the reduction was about twice as large when the signal was in the modality of the person's imagery. Averaging together their visual and auditory d' values, the control d' was 2.70. It dropped to 2.18 with imaging in the opposite modality from the signal, and dropped to 1.74 with imaging in the same modality as the signal. They also found a significant effect of image familiarity. Imaging of relatively unfamiliar events produced greater reductions in sensitivity. It is not yet clear whether the effect is central or peripheral in origin. For example, we know that visual imaging produced pupil dilation and possibly misfocusing on the visual target, which might reduce visual sensitivity.

Posner, Boies, Eichelman visual images and long-term subjects more quickly judge same name if the two letters oppositely to B and b). But this was a delay of about 2 sec presumably due to a rapidly available for effecting a phy- template-matching at lower ever, they obtained a similar auditory prewarning occurs person can respond "same" if the stimuli condition at those implicit icons of the pre-war test flash. A fast name-identity be achieved either by flashing the subject generate a first R.

A more definitive experiment between verbal and pictorial schematic faces, the names are presented. These stimuli were then used by the stimulus (name or picture) was followed by a second stimulus to the "same" if the two were (or different." Over many experiments were run corresponding to the first stimulus, (2) the subjects heard a stimulus, and (3) the modality of the second stimulus was manipulated by the relative modality of the second stimulus over all, the time was that reaction time for either stimulus was in the expected modality of the first stimulus faster to a second picture than if the subject expected a word (upright or slanted) were run. Tversky interprets her findings. Specifically, she suggests that the modality corresponding to the presumably because template repre- sentatives of both stimuli are
Posner, Boies, Eichelman, and Taylor (1969) have some recent evidence on visual images and long-term memory. In initial experiments, they found that subjects more quickly judged whether two successive alphabetic letters had the same name if the two letters were physically identical (for example, B and B as opposed to B and b). But the faster physical-identity match disappeared if there was a delay of about 2 seconds between the two flashed letters. The effect is presumably due to a rapidly fading visual icon. The first letter becomes less available for effecting a physical-identity match to the second flash (much as in template-matching at lower afferent stations). In subsequent experiments, however, they obtained a similar effect with verbal (auditory) prewarings. If the auditory prewarning occurs 0.5-1.0 seconds or more before the test flash, the person can respond "same name" as quickly as in the two-flash, physical-identity condition at those interstimulus intervals. This 0.5-1.0 seconds is presumably the time the subject requires to generate from long-term memory explicit icons of the pre-warned letter, in readiness for matching against the test flash. A fast name-identity match of a first B to a second B can apparently be achieved either by flashing the first shortly before the second, or by letting the subject generate a first B from his long-term visual memory.

A more definitive experiment by Barbara Tversky (1968) distinguishes between verbal and pictorial encoding. Her subjects first learned names for eight schematic faces, the names and faces differing along three binary dimensions. These stimuli were then used in a task like that of Posner et al (1968). A first stimulus (name or picture) was shown, followed by a 1-second blank period, followed by a second stimulus (name or picture), and the subject was to respond "same" if the two were (or had) the same name, and otherwise to respond "different." Over many experimental sessions, a variety of intratrial event frequencies were run corresponding to independent variation in (1) the modality of the first stimulus, (2) the subject's expectation of the modality of the second stimulus, and (3) the modality of the second stimulus. The expectations were manipulated by the relative frequencies (80% versus 20%) of the modality of the second stimulus over large blocks of trials. The important result was that reaction time for either response was very much faster if the second stimulus was in the expected (more frequent) modality. Regardless of the modality of the first stimulus, if the subject expected a picture he responded faster to a second picture than to a second word, but the converse occurred if the subject expected a word. Within-modality controls (words presented upright or slanted) were run to rule out a simple surprise explanation.

Tversky interprets her findings as indicating modality-specific encoding. Specifically, she suggests that a subject will encode the first stimulus into a modality corresponding to the expected modality of the second stimulus, presumably because template matching is possible only when the representatives of both stimuli are in the same modality. Wallach & Averbach (1955) suggested a similar view of recognition memory. According to this
viewpoint, a first name will be converted into an internal facial picture if the subject expects to match it to a picture. This view was supported also by the fact that “different” responses were faster when there were more differing visual features between the expected (imaginal) picture and the actual picture.

There are also indications of specific brain-localization of verbal versus pictorial encoding systems in the work of Kimura (1963) and Milner (1968) with lobectomized patients. The coding systems would appear to be predominantly represented on different sides of the brain. Specifically, removal of the left temporal lobe selectively impairs verbal memory, whereas removal of the right temporal lobe selectively impairs visual (pictorial) memory. Excisions of the parietal or frontal lobes appear not to produce either deficit. Visual memory was tested by recall or recognition of pictures shown earlier. Milner used recall of geometric designs and recognition of photographs of people’s faces, while Kimura used recognition of abstract nonsense figures. Verbal memory was tested by recall or recognition of numbers, nonsense syllables, words, word pairs, and stories. The selectivity of the deficit is the significant fact. Patients with right-temporal lobectomies were deficient in visual but not verbal memory; patients with left-temporal lobectomies were deficient in verbal but not visual memory. There is some evidence too that the right-temporal patients are deficient in mnemonic recognition of nonverbal auditory patterns, for example, they fail to recognize snatches of familiar melodies. Milner argues that the right-temporal deficit is specifically mnemonic rather than perceptual since such patients perform almost normally on a variety of simple perceptual tasks.

Natadzo (1960) has demonstrated some curious biases produced by prior imaginary experiences of a kinaesthetic or tactile sort. In a typical experiment, a subject is asked to imagine metal balls being hefted in each hand, with the left ball being far heavier than the right. If two equally heavy balls are then actually placed in his hands, he is likely to report a contrast illusion. The ball in his left hand is said to be lighter. A size-contrast illusion was produced by similar imaginings. Since imaginal lifting of a weight often produces recordable muscle-action potentials in that arm (Jacobson, 1932), the effects in such experiments might be attributable to peripheral differences in muscle-tone just before the equal-weight tests.

Another line of evidence concerns the effect of imaginal or mental practice on the development of perceptual-motor skills. A variety of evidence is available (cf. Richardson, 1963) to show that imaginal practice improves some skills (e.g., basketball shooting, mirror-reflected star tracing). Although a mental-practice effect does not necessarily implicate imagery versus rehearsal of verbal descriptions or principles, Start and Richardson (1965) have evidence that larger improvements occur for subjects who have vivid but “well-controlled” imagery.

Mental Imagery and Assoc

Another bit of evidence by Laura Phillips (1958) words with five gray patches gray then underwent GSR other nonsense words yield visual similarity of their arousing sensory imagery with that reason the GSR generated equally distinct.

The last bit of evidence We know that a really go practically anything the h is over whether he is mere by Graham (1969) appear subjects (high school students) gray circles when a buzzer sounded the buzzer and h upon actual backgrounds the the left side of the project Graham asked his subjects hallucinated circles. A pro the basis of brightness control background appeared bright white background. When the hallucinated gray cirl

The foregoing brief review some conventionally accepted effects of the process we do most readily explained by which has the ability to rely of previous perceptual experience experimental work on imp be seen, this too provides processes.

EX

The Paired-Associate Task

My research group has li always visual) upon ease of most elementary associative
Another bit of evidence is provided in a mediated generalization study by Laura Phillips (1958). Her subjects first learned to associate five nonsense words with five gray patches varying in brightness. The word for the lightest gray then underwent GSR conditioning. Subsequent tests for GSR to the other nonsense words yielded a smooth generalization gradient according to the visual similarity of their associated gray patches. Presumably the nonsense words arouse sensory imagery which is ordered along the brightness dimension, and for that reason the GSR generalizes differentially to words which are otherwise equally distinct.

The last bit of evidence on imagery is from hypnotic hallucinations. We know that a really good hypnotic subject will report seeing or hearing practically anything the hypnotist tells him to see or hear. The controversy is over whether he is merely role playing. A recent demonstration at Stanford by Graham (1969) appears relatively sham-proof. Graham conditioned his subjects (high school students) under hypnosis to visually hallucinate two gray circles when a buzzer sounded. Then, while they were awake, he sounded the buzzer and had them hallucinate the gray circles as projected upon actual backgrounds of different brightness. For instance, in some tests the left side of the projection card would be white and the right side black. Graham asked his subjects to judge the relative brightness of the two hallucinated circles. A preponderance of their judgments were predictable on the basis of brightness contrast. The gray circle hallucinated against a black background appeared brighter than the other circle hallucinated against a white background. When the two sides of the projection card were white, the hallucinated gray circles were judged equally bright.

The foregoing brief review of behavioral evidence of imagery indicates some conventionally acceptable (i.e., nonintrospective) evidence for the effects of the process we call imaging. At the present time, such results are most readily explained by postulating an information-processing system which has the ability to regenerate a structural representation or analogue of previous perceptual experiences. I will now turn to an account of our experimental work on imaginal mediators in associative learning. As shall be seen, this too provides evidence for the behavioral reality of imaginal processes.

EXPERIMENTAL STUDIES

The Paired-Associate Task

My research group has been concerned with the influence of imagery (always visual) upon ease of associating disparate items of experience. The most elementary associative task is paired-associate learning (PAL) which
we have used almost exclusively in our studies of imaginal mediation. Rather than using perceptual events or real objects as the terms to be associated, we use noun words. This is an experimental convenience but it also keeps our work in contact with the verbal-learning mainstream. The use of nouns introduces some special problems, however, because we depend upon the noun, together with appropriate instructions, to arouse some imagery in the subject. Nouns surely vary in their probability of evoking imagery.

How do nouns come to evoke their associated imagery? Psychologists since Aristotle have given roughly the same account: by association (conditioning) of the word to the perceptual imagery (sensory responses) elicited by the referent of that word. Sheffield (1961) or Staats (1968) can be read for recent accounts. Semantically ambiguous nouns (there are thousands) are also imaginatively ambiguous, in the sense that bag may arouse imagery having to do with containers or with old women. Abstract nouns presumably evoke little or no imagery directly, but may do so indirectly through associated concrete words (e.g., church or priest as associates to religion).

In our typical task, the subject (a university student) is presented with arbitrary pairs of unrelated concrete nouns, like dog-bicycle, and instructed to associate them by imaging a visual scene or mental picture in which these two objects are interacting in some way. The cognitive situation here is unconstrained in several respects: first, we neither control nor measure the specific instances the subject imagines to exemplify the concepts; second, we do not control the exact scene in which the two objects are placed. After illustrating the types of interactions we have in mind, the subject is left to generate his own imaginative scene. All of the college students we have tested carried out such instructions without the slightest difficulty.

The first results illustrate the imagery effect we shall be discussing throughout the remainder of this paper. Each subject had one 5-second exposure to each of 20 concrete-noun pairs, followed by a cued recall test of the 20 pairs immediately after the study trial. The left-hand word was the cue for literal recall of the right-hand word. Five seconds were allowed for recall, and the subject was informed of the correct response at the end of this test trial. Five successive lists of 20 pairs were tested in this manner. At the end of the session all 100 pairs were tested again. Subjects given imagery instructions were compared to control subjects who received standard PAL instructions: to learn the pairs so the left member could cue recall of the right member.

Recall percentages are shown in Figure 3.1. The salient feature is the clear superiority of the imagery subjects: they recall about one and a half times as much as the Control subjects, and the difference is statistically significant on each list. The delayed recall scores exceed the immediate recall scores, presumably the result of the information provided at the end of

Figure 3.1 Immediate Recall

This elementary experiment yields a clear imagery effect. In fact, the effect is even more impressive because interviews with Control subjects using imaginal and/or verbal strategies and Snyder (1966) found no imagery effect. Control subjects were given the following instruction: That is, they were told to picture the pairs themselves.

Similar differences arise when the pairs being learned concurrently were shown in one of two locations. John Schnoor at Stanford has shown that these were to be learned by imagery by repeating the words aloud 8 seconds per pair. Within 12 by repetition. Immediately afterwards, imagery pairs and 33% for the subjects were very accurate. All the subjects were very accurate. All the subjects were very accurate.

Recall differences like these are due to differences in processing style. Imaginal processing appears to suit some people very well, but not others. The evidence suggests that imagery is an accessible and useful memory tool for some people.
the first test trial. There is a small practice effect over the five lists that is of no interest in this context.

![Graph showing immediate and delayed recall over 5 lists for imagery versus control subjects.](image)

**Figure 3.1** Immediate and delayed recall over 5 lists for imagery versus control subjects.

This elementary experiment illustrates the magnitude of the imagery effect. In fact, the effect is probably underestimated in Figure 3.1 because interviews with Control subjects turned up some who were spontaneously using imaginal and/or verbal (sentence-linking) mediators. Paivio, Yuille, and Snythe (1966) found similar control reports. In our later experiments, Control subjects were given explicit repetition-learning instructions. That is, they were told to repeat the word pairs over and over to themselves.

Similar differences arise when imagery versus repetition is varied between pairs being learned concurrently by the same subject. A PAL experiment by John Schon at Stanford had the following format: concrete noun pairs were shown in one of two locations on a table top. Those shown to the left were to be learned by imagery. Those shown to the right were to be learned by repeating the words aloud three times. Exposure time was controlled at 8 seconds per pair. Within each list, 12 pairs were studied by imagery, and 12 by repetition. Immediate recall tests showed about 80% recall for the imagery pairs and 33% for the repetition pairs. Moreover, on recalled pairs, the subjects were very accurate in remembering whether the pair had been studied by imagery or verbal repetition.

Recall differences like those illustrated above could arise for a number of reasons. Imaginal processing may raise the general availability of the response term (cf. Horowitz, Norman, and Day, 1966). This hypothesis is discounted
by an experiment comparing imagery versus repetition instructions, wherein a random half of the pairs was tested for recognition and half for recall. The recognition test was a five-alternative multiple-choice display. The left-hand member (stimulus) of the pair was shown along with five right-hand members (responses) from the prior study list. The subject selected the one response word which he thought had been paired with the stimulus word. This procedure was replicated over three lists of 30 pairs, studied and tested at a 5-second rate. Recall was 87% for imagery and 37% for repetition. Recognition was 97% for imagery and 71% for repetition. The latter difference is significant. This recognition outcome discounts the response availability conjecture, since the imagery-repetition difference still appears strongly when the responses are directly available in the multiple-choice test. Thus an associative effect of imagery is implicated.

Incidental Learning

A further question is whether the superior learning displayed under imaginal elaboration depends upon the intention to learn. Accordingly, intentional versus incidental learning was compared within groups given imagery instructions. The cover story for the incidental learners was that we were collecting norms for the English Literature department on the vividness of word pictures. The subjects were instructed to imagine some interactive scene, and to rate its vividness on a five-point scale, doing this within the 5-second exposure of each pair. Subjects in the intentional condition were similarly instructed, but were also told they had to remember the pairs for an upcoming recall test. After one study trial of 20 pairs (List 1 of Experiment 1), cued recall was 77% for the intentional-imagers and 71% for the incidental-learners. These do not differ significantly. Both scores may be contrasted with that of the intentional controls learning this list in Exp. 1. They recalled 35%, less than half that of either imagery condition.

This outcome for the incidental-intentional comparison is not surprising in the light of the recent literature on incidental learning (Postman, 1964). That literature has shown that the intent to learn is superfluous if the cover task orient the subject to the relevant material and requires him to make differential responses to it. For later recall, the cognitive or imaginal elaboration itself is the important ingredient, not his motivation to remember.

Recall that these subjects rated the vividness of their imaginary scenes at the time of construction. Figure 3.2 shows the relationship between the vividness rating of a pair and its probability of later recall. More vivid scenes are better remembered for both groups. This relation held within each individual: each of the 36 subjects gave a higher mean vividness rating to pairs he later recalled than he did to pairs that the subject has available presentation predictive of whether it might include (a) the story of the speed and uniformity required to search for and the two components. Vividness to fix, and likely to be retained. This may reduce effective word pair.

Multiple Associates to Each Peg

In another experiment, we integrated multiple items into the pegs to see if the pegs were as strong as the pegs used in the original investigation. The pegs used in the original investigation were used in the present experiment. The pegs were the pegwords, which the subject learned in the first list. He then uses these as imaginal pegs to recall the list. Thus, for list 1 (peg #1), the second list-word is the first list-word. The subject is asked to remember all the items on the list. In this context, our null hypothesis was that recall suffers with the use of each peg. Accordingly, diff
Mental Imagery and Associative Learning

he later recalled than he did to nonrecalled pairs. This relation means that the subject has available, at the time of imaginal elaboration, information predictive of whether he will remember the pair. Suggested factors might include (a) the strength of the word-image associations, determining the speed and uniformity of imagery aroused by each word, and (b) the time required to search for and generate a sensible relation or interactive scene for the two components. Vivid imagery is stable; faint imagery is elusive, difficult to fix, and likely to be rapidly replaced by alternative imagery for the words. This may reduce effective inspection time for any given pictorial coding of the word pair.

![Graph](image)

**Figure 5.2** Relation between recall of a pair and its rating of imaginal vividness during study for intentional and incidental learners.

Multiple Associates to Each Cue

In another experiment, we were interested in whether the person could integrate multiple items in an imaginal scene, and whether the associations so achieved were as strong as those involving only two items. The first investigation used a “pegword” mnemonic — a system for converting a non-cued (free) recall task into a paired-associate task using implicit stimuli (the pegwords) which the subject can generate from memory. The subject first learns the pegwords such as one-gun, two-shoe, three-tree, and so on. He then uses these as imaginal pegs for the successive items of any new list he is to learn. Thus, the first list-word is imagined in a scene with a gun (#1), the second list-word with a shoe (#2), and so on. This strategy solves one of the major problems in free recall: finding a way to remind oneself of all the things one is supposed to recall.

In this context, our multiple-associates question can be phrased by asking whether recall suffers with fewer pegwords but more list-words attached to each peg. Accordingly, different subjects were instructed to use either
1, 2, 5, 10, or 20 pegs to learn a 20-item list. The words were presented serially at a 5-second rate. Subjects learning with k pegs were instructed to sequentially elaborate a grand imaginative scene involving the first peg and the first 20/k items, the second peg and the second set of 20/k items, and so on. Recall was scored as free recall, although subjects were instructed to try to generate their recall from the k successive pegs they had learned and used at storage. The subjects had one study and recall of each of five lists of 20 concrete nouns, followed by a final recall test for all five lists at the end of the session.

The results of this experiment are shown in Table 3.1 which also contains the scores for a control group given standard free recall instructions. In brief, the number of items per peg has no differential effect, and all imagery subjects recall two to three times more than the controls on the delayed test. With one peg (the 1-20 condition of Table 3.1), it can hardly be said that we have any pegword retrieval cues at all. So the effect in the 1-20 and 2-10 conditions must arise from sequential associations—consecutive chunks of, say, three to five items integrated into progressively developed but overlapping scenes. The process may resemble episodes in a movie that slowly unfolds. For example, the list “dog, cigar, hat, bicycle, policeman, lady, fence, basket,...” can be integrated into scenes (or sentences) like “A dog smoking a cigar and wearing a hat, is riding a bicycle which collides with a policeman; the policeman chases the lady (who owns the dog) and corners her near a fence, but she throws a basket at him, and ...”. Subjects can easily generate composite imaginal scenes containing a large number of separate items without very much interference from successive items.

<table>
<thead>
<tr>
<th>Number of Pegs</th>
<th>Words/Peg</th>
<th>Immediate</th>
<th>Delayed</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>1</td>
<td>81</td>
<td>69</td>
</tr>
<tr>
<td>10</td>
<td>2</td>
<td>85</td>
<td>66</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
<td>87</td>
<td>71</td>
</tr>
<tr>
<td>2</td>
<td>10</td>
<td>89</td>
<td>77</td>
</tr>
<tr>
<td>1</td>
<td>20</td>
<td>87</td>
<td>75</td>
</tr>
<tr>
<td>Controls</td>
<td>-</td>
<td>52</td>
<td>28</td>
</tr>
</tbody>
</table>

A later paired-associate experiment modified this conclusion in one respect. The recall probabilities from one cue-term to N response-terms were all high if the N responses were given simultaneously (massed) or in consecutive serial order, so that the subject could integrate them with the cue-term. The recall probabilities were lowered if the N pairs involving the same cue-term were presented singly, well separated in time by intervening pairs involving other stimuli, giving average single response-term percentages. These results can be interpreted in terms of pre-existing associations among response terms, the recall of which is facilitated by separate attention to the cue-term.
pairs involving other stimuli in the list. These results are shown in Figure 3.3, giving average single response recall when the cue-term had 1, 2, or 5 response-terms. These results came from separate lists involving 20 response-terms where the number of stimuli (cue-terms) was 20, 10, or 4 for the different lists. Presentation times were 5 seconds per S-R unit, and subjects received imagery instructions. In the separated condition, recall probability declines with the number of response-terms hooked singly to the cue-term. Introspectively, the massed presentations permit the subject to place the cue and all N response-terms together in one imaginal scene, whereas the singly-presented, separated pairs are more likely to be treated in isolation, with the N response-terms imaged in separate scenes with the cue-term. In conventional terminology, the former arrangement allows for associative mediation among response terms, whereas the latter permits unlearning of earlier pairs by separate attention to the later pairs involving the same stimulus.

![Graph showing immediate and delayed recall related to the number of response-words attached to each stimulus-word. "Massed" denotes simultaneous presentation of the several responses to each stimulus; "separate" denotes temporally distributed presentations of the several S-R pairs with the same stimulus.](image)

**WORD-ATTRIBUTE EFFECTS**

Since words are the materials in our experiments on imagery, one is led into a concern for the image-arousing capabilities of the words used, how this varies with other measurable attributes of the words, and how PA learning varies with both. Within the traditional S-R formulation, we would think of the word as a stimulus associated to a number of mediating responses: verbal, imaginal, and perhaps evaluative (or emotional) responses. Verbal associates might include synonymy, opposites, subordinate and superordinate
Cognition in Learning and Memory

categories, common members of categories to which the word belongs, and syntagmatic associates. Through instruction, the imaginal associates of a word can be primed selectively. The vividness of imagery aroused by a given word would be an index of the strength of the association between the word and some imaginal mediator. Paivio, Yuille, and Madigan (1968) have collected imagery ratings for 925 nouns. They find, of course, that concrete nouns have higher imagery value than abstract nouns. And as the strength-theory expects, latency for image-production is shorter for words rated high in imagery (Paivio, 1966).

The effects of imagery value upon PA learning can be traced with the help of the flow diagram in Figure 3.4. In brief, at the top, words A and B are first given a semantic interpretation which primes or selects particular object images. These are placed into some interactive picture, and a program for generating that imaginal scene is recorded in memory. At the bottom, when cued for recall, word A is interpreted, again leading to an imaginal code for A. The code may retrieve the imaginal composite, from which the subject selects the imaginal code for B. He then decodes this into a report of word B. The multiple arrows represent alternative branches (or competing mediating response) which can be followed, and which may thereby interfere with successful flow of events in the chain.

![Flow diagram of mediating events during the study trial and recall test.](image)

**Figure 3.4** Flow diagram of mediating events during the study trial and recall test.

Ambiguity and Similarity of Meaning

Consider first ambiguous nouns as cues. Since the semantic interpretation precedes the imaginal production, recall will be poor if the semantic coding of the noun differs between the study trial and the recall trial. The semantic interpretation of nouns can lead to different adjectives. Thus, a nylon stocking can be a nylon stocking or another. In experiments where nouns are repeated exactly, or the adjective meaning (e.g., lamb chop versus karate chop). Recall is about double the recall of the word with the same meaning. This is consistent with imagery associations; and (b) there is a word to an image. A semantic rule controls whatever imagery is produced.

Now let us consider synonyms. May arouse similar or identical pictures retrievable from memory. To the stimulus CAR, a stimulus AUTOMOBILE may arouse synonyms on the response terms for differences when he should have said SHOES.

One of our experiments has pairs to subjects learning a list of objects (with unique responses) (with unique stimulus terms) with visual imagery instruction, and by auditory rehearsal--that is, they said the word pair over and over.

The results of this experiment show that all recall failures (errors plus imagery subjects versus only imagery subjects) was largely due to synonym confusion. The experimental condition illustrates how the coding of different types of material.
interpretation of nouns can be altered by context, in particular by qualifying adjectives. Thus, a nylon slip is one thing, while a Freudian slip is quite another. In experiments conducted in my lab by Samuel Bobrow, PA conditions have been studied in which the cue-term (adjective noun) is either repeated exactly, or the adjective is altered either to preserve similarity of meaning (e.g., lamb chop versus meat chop), or to alter meaning (e.g., lamb chop versus karate chop). Recall percentages in the first two cases are equal and are about double the recall in the latter case, when the critical noun cue has its meaning altered. This occurred despite the fact that subjects knew that the noun was the critical recall cue. Such findings, of course, do not implicate imagery in any way. But they do make two important points: (a) what is stored in most such cases are the meaning relations, not word-word associations; and (b) there is probably no automatic or direct connection from a word to an image. A semantic interpretation is probably interposed, and it controls whatever imagery will be aroused.

Now let us consider synonymous confusions. Synonyms are words which may arouse similar or identical imagery. If synonymous stimulus words are paired with different response words, there will then be two composite pictures retrievable from memory by the single image code of either stimulus word. To the stimulus CAR, subjects may recall the response appropriate to the stimulus AUTOMOBILE. Similar confusions may be expected from priming synonyms on the response side. If the words BOAT and SHIP appear as response terms for different stimuli, the subject may respond BOAT when he should have said SHIP.

One of our experiments compared subjects learning a unique list of concrete pairs to subjects learning a list involving half synonym pairs on the stimulus side (with unique responses) and half synonym pairs on the response side (with unique stimulus terms). Two groups were run: one group received the usual visual imagery instructions, the other subjects were instructed to learn by auditory rehearsal—that is, by listening to the composite sound as they said the word pair over and over during the study interval.

The results of this experiment are shown in Table 3.2. The salient features of these data are first, that imagery subjects recall more, and second, that list-type and learning method interact upon recall. Specifically, the shift from the unique list to the synonym list caused an appreciable drop in recall of imagery subjects, but little change in recall of the auditory-rehearsal subjects. Moreover, the increase in errors on the synonym list by imagery subjects was largely due to synonym confusions of the types mentioned above. Of all recall failures (errors plus omissions), 31% were synonym intrusions for imagery subjects versus only 2% for auditory rehearsal subjects. This experiment illustrates how the coding set used by a subject may affect the learning of different types of material.
TABLE 3.2 Recall Percentages for Different Lists for Subjects Given Different Learning Sets

<table>
<thead>
<tr>
<th>Condition</th>
<th>Visual Imagery</th>
<th>Auditory Rehearsal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unique list</td>
<td>.83</td>
<td>.39</td>
</tr>
<tr>
<td>Synonym list</td>
<td>.66</td>
<td>.36</td>
</tr>
<tr>
<td>Synonym errors/</td>
<td>.31</td>
<td>.02</td>
</tr>
<tr>
<td>all errors</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Returning just for a moment to the results of Table 3.2 in relation to the flow chart of Figure 3.4, it should be mentioned that just because a response term has a synonym does not imply that the synonym will intrude. Rather, the synonym has to be primed into availability. The response words in the unique list above have synonyms, yet these never intruded in recall of that list. As has been often noted (e.g., Underwood and Schultz, 1960), some selector mechanism or editor confines recalled responses to the list. In our flow chart, this means that the response-term image is selected from the primed or recently tagged verbal labels for that image.

One final point here: with multiple trials on a synonym list, imagery subjects probably overcome confusions by elaborating differential imagery for the two terms. Thus, BOAT comes to be imaged as a dinghy, and SHIP as an ocean liner. Because no two words are really identical, a selection of differential elements for elaboration can always be done.

Imagery Value

Consider next in Figure 3.4 the effect of variation in the imagery value of the cue-term of the paired associate items. If imagery rating is an index of word-image association strength, then we may expect high imagery cues to be more effective for two reasons: first, the imaginal code occurs quicker, so there is more time to search for and generate the relational scene, or to study the cue code-to-response association. Second, recall depends upon reactivating the same imaginal code (for the cue-term) as was activated during the study trial. If the cue does not activate the same code, then recall will surely fail (see Martini, 1968). The strong associations of high imagery cues increase the reliability of the imaginal code. For this reason, the imagery value of the cue-term should be especially important.

A number of experiments could be used to illustrate these effects. We used people's names as cues for a noun paired-associate response. Each subject answered a questionnaire which enabled us to identify three different types of names for him: (a) names of 21 of his personal friends, whom he could picture clearly in his mind; (b) names of 21 historical characters, about whom he knew some facts b (e.g., had not seen pictures of known people randomly selected and were presented in three mixed paired with concrete nouns i instructions. On the test trial, cues was .93 for personal friends and .86 for unknown names. Subjects re historical characters because they helped them imagine a prototype of his famous thing (e.g., Eli Whitney).

Although I would like to add some of the names, frequency of exposure was confounded along with imagery. The treatment was done with photographs of people as cues for paired associates of national and campus celebrities by our subjects. In this experiment, known persons shrank to 0.8 of the difference is of small mac images and unknown names. The average recall differences here are small, so recall differences here are small, but not very large.

Controlled variation of word detail by Paivio and his coworkers. Important finding is that the imagery in PAL is a critically important in PAL is a critically important factor in PAL is a critically important factor in the imagery rating and response sides of pairs, I-value cases, but much more so on the other side.

The results of one of our experiments. Subjects received one study trial on three lists, each condition shown in Table 3.3. At the end of the study, a backward recall test were given, and their final tests are shown in Table 3.3. The large effect of stimulus I-value on test I (85 versus 71). The other factor between forward and backward recall; with
Mental Imagery and Associative Learning

whom he knew some facts but did not know exactly what they looked like (e.g., had not seen pictures or paintings of them); and (c) names of 21 unknown people randomly selected from the telephone directory. These were presented in three mixed lists, seven cues of each type in each list, paired with concrete nouns for one 5-second presentation with imagery instructions. On the test trial following each list, noun recall to the name cues was .93 for personal friends, .66 for historical characters, and .43 for unknown names. Subjects reported being able to do moderately well on historical characters because the few facts they knew about the characters helped them imagine a prototype character, dressed in period clothes, doing his famous thing (e.g., Eli Whitney stuffing his cotton gin).

Although I would like to attribute this effect to differential imagery value of the names, frequency of experience, and verbal associations may be confounded along with imagery. Let me therefore add that a subsequent experiment was done with photographs of front-faces of known versus unknown people as cues for paired associate noun responses. The known faces were of national and campus celebrities, all of whom could be immediately named by our subjects. In this experiment, the recall difference between known and unknown persons shrank to only 10% (65% versus 55% recall after one trial). The difference is of small magnitude relative to the difference between known and unknown names. The argument is that with the picture, the visual (imaginal) cue is immediately available for unknown as well as known characters, so recall differences here are slight. With the names as cues, however, imagery is available for known but not for unknown names, so the recall differences are very large.

Controlled variation of word-attributes in PAL has been researched in great detail by Paivio and his coworkers (cf., Paivio, 1969, for a review). A first important finding is that the image-evoking value (I) of the stimulus term in PAL is a critically important factor in learning rate, even with nonmeaningful response terms. When I-values are simultaneously varied on the stimulus and response sides of pairs, I-value is positively correlated with learning in both cases, but much more so on the stimulus than on the response side.

The results of one of our experiments may be used to illustrate this pattern. Subjects received one study trial and an immediate forward (left-to-right) recall trial on three lists, each consisting of a third of the three classes of pairs as shown in Table 3.3. At the end of the session, a final forward recall test and a backward recall test were given for the pairs from all three lists. Results of these final tests are shown in Table 3.3. Comparing forward recall, one finds a large effect of stimulus I-value (85 versus 56) but only a small effect of response I (85 versus 81). The other feature of interest in Table 3.3 is the relation between forward and backward recall of a pair. With high-low pairs, forward exceeds backward recall; with low-high pairs, backward exceeds forward.
recall. It thus appears that retrieval of the imaginal association is best achieved by using the high-imagery member of the pair. Lockhart (1969) has reported a similar finding.

TABLE 3.3 Forward and Backward Recall Percentages for Pairs Varying in Imagery.
High Denotes High Imagery Concrete Nouns; Low Denotes Low Imagery Words; a Third Each of Abstract Nouns, Verbs, and Adjectives.

<table>
<thead>
<tr>
<th>Percent Recall</th>
<th>Training Function</th>
<th>Forward</th>
<th>Backward</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Stimulus</td>
<td>Response</td>
<td>$R \text{ given } S$</td>
</tr>
<tr>
<td>High</td>
<td>Low</td>
<td>81</td>
<td>70</td>
</tr>
<tr>
<td>High</td>
<td>High</td>
<td>85</td>
<td>85</td>
</tr>
<tr>
<td>Low</td>
<td>High</td>
<td>56</td>
<td>71</td>
</tr>
</tbody>
</table>

That stimulus imagery-concreteness correlates with PA learning cannot be taken at face value; other word attributes concerning verbal meaningfulness may be related to I-value, and these may be the truly effective variables. It was to test such possibilities that Paivio and his associates carried out their extensive experiments, first scaling different word-attributes, and second, selecting subsets of words that varied one attribute while controlling others. The chief attribute contrasted with imagery-concreteness was “meaningfulness” ($m$) of the verbal unit, defined in terms of the number of verbal associates produced to the unit within a fixed time. This is a conventional index that has been found to correlate well with PAL and with serial learning, particularly when nonsense syllable units are used. There is a correlation of .72 in Paivio’s norms between the I-value and $m$ of his words, so it is not implausible that I versus learning correlations are really due to $m$ variations.

In a series of experiments (reviewed by Paivio, 1969), I and $m$ values of words were independently varied over equivalent scale ranges on both the stimulus and response sides of PA pairs. The conclusions were that $m$ has little or no effect on PA learning when I-value is controlled or is partialled out of the correlations. The most that has been found for $m$ when I is controlled is an occasional slight negative relation between response-$m$ and PA learning rate. On the other hand, for fixed $m$, PA learning correlates highly with the imagery-concreteness of the stimulus term, and correlates somewhat with the I-value of the response term. The significance of such results is clear: those PA studies reporting learning effects of $m$-variations with words could all be reinterpreted as displaying uncontrolled variation in I-values. The bulk of studies showing $m$ effects in learning have used nonsense trigrams, and there the $m$ variations are doubtless indexing the extent to which the trigram reminds subjects of a real word.

Mental Imagery and Ass

Paivio (1968) has related with I-value and $m$ and PA learning. Everyday differential frequency, familial imagery-concreteness, and with PA learning, especially in imagery-concreteness of PA learning yet unde

HYPOTHE

The evidence indicates that (1) subjects who have the following strategy learn high-imagery instructions to image related attributes the effect of I-value condition. So that such a powerful technique for planning will be consider

Motivation or Interest

A first possibility is that PA learning, making for learning, may be defined versus intentional-content view. But the basic problem of interest is to vary among the one or to employ more efficient explanation useless because

Selection from Incidental Cues

If the nominal stimulus word may select a cue, subject may select a cue word may arouse the imaginal scene contents alone, then there is a go
Paivio (1968) has searched for other word attributes that may be correlated with I-value and which might be causing the correlation between I-value and PA learning. Everyone has their favorite variable to suggest—Thorndike-Lorge frequency, familiarity, associative variety, emotionality, semantic differential ratings, and so on. But these all turn out to correlate only slightly with PA learning, especially when I-value is controlled. It thus appears that imagery-concreteness of the stimulus word is the most potent attribute for PA learning yet identified.

**HYPOTHESES ABOUT IMAGINAL FACILITATION**

The evidence indicating that imagery facilitates PA learning is of two sorts: (1) subjects who have not been specifically instructed to use a particular learning strategy learn high-imagery pairs faster than low-imagery pairs, and (2) instructions to image relationships improve recall. In fact, we can probably attribute the effect of I-variation in experiments of the first type to the second condition. So the crux of the problem is to explain why imaging is such a powerful technique for learning associations. Several possible explanations will be considered in turn.

**Motivation or Interest**

A first possibility is that imagery instructions enhance the subject's motivation for learning, making the task more interesting. The imagery-incidental versus intentional-control contrast presented earlier may be offered against this view. But the basic problem with this hypothesis is that it is difficult to test. One might try varying motivation by payoffs; but even if this had an effect, one could not be sure if high motivation caused the subject to rehearse more or to employ more efficient learning strategies like imagery. I consider this explanation useless because it explains nothing, or everything.

**Selection from Incidental Redundant Cues**

If the nominal stimulus contains a redundant array of possible cues, the subject may select a cue which is easiest to link to the response term. The stimulus word may arouse an image containing a large number of incidental cues. For example, my image for the word *boy* might be of my son in a particular setting, engaging in certain activities. Incidental cues might be background objects, things he is carrying or the clothes he is wearing. If the imaginal scene contains many more potential cues than the word *boy* alone, then there is a greater chance that a cue having a strong prior association
to the response will be contained in and selected from the former stimulus complex.

If this hypothesis is true, then one should be able to simulate this effect at the verbal learning level, without using imagery, simply by providing more redundant cues in the nominal stimulus. Accordingly, a PA experiment was run comparing the efficacy of a single-word nominal stimulus to a redundant three-word nominal stimulus. All were concrete nouns as were the single response terms. Half of the three-word compounds was unrelated and half was associatively related, the latter to simulate the interconnected components of an image. This amusing experiment was a total bust: contrary to the good-cue selection hypothesis, the pairs involving triad stimulus compounds were recalled worse than pairs involving single-word stimuli. The response recall percentages were: single-word cues, 48%; unrelated triad cues, 33%; associatively related triad cues, 35%. So, tripling is crippling. The effect is probably due to a reduction in effective study time for the pairing, and to increased variability of stimulus encoding for the trials.

A further nail in this coffin was provided by results of Jenkins (1967). He found that his subjects learned picture-word PAs faster if the perspective of the photograph (e.g., different snapshots of an elephant) was varied from trial to trial, rather than remaining constant. Since background incidental cues varied in the former case but not the latter, the incidental-cue-selection hypothesis would have predicted the opposite results.

Stimulus Distinctiveness

Another possibility is that imaginal coding facilitates learning because imagery causes the effective stimulus complex to become more distinctive. By distinctive I mean that the imaginal complex is better differentiated from the other stimuli in the list, as though a distinctive color had been attached to each cue word. Distinctiveness thus implies less intralist interference, less confusions among pairs, and more accurate recall.

There are problems with this account. First, in Gibson’s (1940) original sense of differentiation, practically all moderate-to-high frequency words are already well differentiated stimuli for adults. Second, experiments using predifferentiation of cue words or trigrams have reported small or inconsistent effects upon later PA learning. We might wonder if the highly personalized, idiosyncratic codings elaborated under imagery instructions produce the critical distinctiveness.

We have conducted one experiment to test this hypothesis. During the study trial the stimulus word appeared alone for 5 seconds, followed by the stimulus-response pair for five seconds. Four different groups of subjects were run. In the subject-elaboration condition, the subject was asked to provide a descriptive phrase for a familiar example of the word. The subject might produce "in the stimulus-alone this man uses the noun-stimulus with the root..."

In a second condition, subjects learned the critical cue word. These subjects were then paired, the stimulus word was presented on a cue card, and the subject had to repeat the pair silently when it appeared. The subject was told to repeat the pair silently when it appeared. The subject was told to repeat the pair silently when it appeared. The subject was told to repeat the pair silently when it appeared. The subject was told to repeat the pair silently when it appeared. The subject was told to repeat the pair silently when it appeared. The subject was told to repeat the pair silently when it appeared. The subject was told to repeat the pair silently when it appeared.

The question is whether verbal elaboration of the stimulus word to anything, recall is decreased for the stimulus word. And as before, the visual imagery group (experiments) is because half of the pairs were recalled one and a half times in this respect was practiced.

<table>
<thead>
<tr>
<th>TABLE 3.4</th>
<th>Mean Recall</th>
<th>Diffe...</th>
</tr>
</thead>
<tbody>
<tr>
<td>Repetition</td>
<td>Imagery</td>
<td>E-generated elab</td>
</tr>
</tbody>
</table>

Relational Organization

The former two hypothesis selection or cue distinctiveness.
provide a descriptive phrase for each stimulus word that identified a specific familiar example of the word. For example, for the stimulus word SLIPPER, the subject might produce “my mother's pink SLIPPER.” After elaborating the stimulus-alone in this manner, the subject was instructed merely to associate the noun-stimulus with the response word when it appeared later.

In a second condition, subjects simply read some context words along with the critical cue word. These stimulus elaborations were of the type generated by subjects in the former condition. The contextual elaboration of the critical stimulus word was provided during the study trial only and not during the cued recall trial. Along with these two conditions of stimulus elaboration, standard repetition and visual imagery control groups were run. The former subjects were told to repeat the stimulus word aloud, and then repeat the S-R pair silently when it appeared. The latter subjects were told first to image an object for the stimulus word and then, when the S-R pair appeared, to image the stimulus-response objects in interaction. Two study-recall cycles were done in this manner on each of three lists of 30 noun pairs, half of the pairs concrete and half abstract.

The question is whether verbal elaboration of a stimulus word makes it more distinctive, thus facilitating recall as much as imagery does. The answer to this question turns out to be distinctly negative. The relevant results are shown in Table 3.4 giving average recall probabilities over the two trials. Verbal elaboration of the stimulus word in context does not enhance recall. If anything, recall is decreased slightly below that of the repetition controls. And as before, the visual imagery subjects recall more than the rest of the pack. The low recall score for imagery subjects here (compared to the earlier experiments) is because half the pairs were abstract nouns. The concrete pairs were recalled one and a half times better than abstract pairs, and the split in this respect was practically identical in all four conditions.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Percent Recall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Repetition</td>
<td>53</td>
</tr>
<tr>
<td>Imagery</td>
<td>71</td>
</tr>
<tr>
<td>E-generated elaboration</td>
<td>47</td>
</tr>
<tr>
<td>S-generated elaboration</td>
<td>50</td>
</tr>
</tbody>
</table>

Relational Organization

The former two hypotheses—relating the imagery effect in PAL to cue selection or cue distinctiveness—relate the effects of imagery to embellishment
of the stimulus. The alternative view, which I consider to be closer to the truth, is that imagery enhances the associative connection between the two terms.

A simple experiment suffices to show how this relating or interactively organizing feature of imagery instructions is critical. An incidental learning experiment compared two imagery groups. One group received our standard interaction instructions, namely, to image a scene of the two objects interacting in some way. The second group received separation instructions: to image the two objects separated in their imaginal space, like two pictures on opposite walls of a room. One object-picture was not to be influenced in any way by the contents of the other object-picture. Each subject had 10 seconds to do his thing with each concrete-noun pair and to give a rating of how easy it was to visualize the objects. Cued recall was then tested with the result that the interactive imagery subjects recalled 71%, whereas the separated imagery subjects recalled only 46% of the response terms. This is a highly significant difference. But more importantly, the low recall of the separated imagery subjects is about the figure we would expect from repetition controls given 10 seconds per item. Thus, instructions to image the terms per se have relatively little effect on associative learning. The important component is the interactive relation between the imaged objects.

We may ask, what characterizes interactive scenes? If one examines the descriptive statements subjects use to characterize their interactive scenes, the phrases usually relate the two nouns as grammatical subject and object connected by a verb or a preposition. The usual scenes for the pair DOG-BICYCLE might be “a DOG riding on (or chasing, or barking at, or urinating on, or running into, or being hit by) a BICYCLE” or “a DOG on (behind, under) a BICYCLE”. The subject and object nouns may be embellished by adjectival modifiers (e.g., a mangy old DOG), but the base syntactic form contains either a verb or a prepositional connective. On the other hand, subjects in the separated imagery condition would describe their imaginal scene as “a DOG over here and a BICYCLE over there”. Again, embellishing adjectives may be added to each noun, but the connective remains the simple conjunction and.

There are parallel findings showing the same pattern in PA learning using pictures of familiar objects, and also using words related through different syntactic connectives. Since experiments by Epstein, Rock, and Zuckerman (1960), we have known that two pictured objects are more easily associated if they are shown in some kind of spatial interaction (e.g., lamp in a bottle) as opposed to just showing them side by side (lamp and bottle). Epstein et al. (1960) and Rohwer (1966) have shown similar variations in PA recall of noun pairs when, at input, the nouns are connected by a verb or preposition (leading to high recall) as opposed to a conjunction (leading to low recall). To test the facilitation to make sense for the two sentences “lamp sings bottle” or “lamp bottle”.

This recall pattern was produced by the same reasoning to which I do not subscribe primary, and that differing they do because of the difference generate. If one believed that sentence recall sentence context affects sentence. We have done some associates have done a lot to summarize it very brief.

Sentence Contexts in PA

With primary-grade children learning when the nouns are presented by conjunctions (and/or). The subjects who merely hear or read the nouns in the context of declarative sentences facilitation in learning. If the word pairs, then their recall of the nouns in the context of declarative sentences or nonsensical sentences pairs alone, and (2) the association of the noun is roughly symmetrical to the syntactic form of the sentence, affirmative, negative, or negative interrogative.

One interesting finding was generated much better than expected. In one condition, subjects were shown some sentences up and say some sentences. These subjects were shown the two similar declarative sentence pairs of seconds. Later recall (of 1.5 minutes) is about one and a half times higher for these subjects. This difference we attribute it to a difference in generating a linking sentence.
Mental Imagery and Associative Learning

recall). To get the facilitation, the prepositional and verb connectives have to make sense for the two nouns; semantically anomalous connectives like "lamp sings bottle" or "lamp how bottle" are of no help in promoting later recall.

This recall pattern with pictures, images, and words is probably being produced by the same relational generating system. An extreme view, to which I do not subscribe, would suppose that the linguistic medium is primary, and that differing pictorial or imaginal scenes have the effects they do because of the different descriptive statements that subjects generate. If one believed this, then one would question how variations in sentence context affect the association between critical words in the sentence. We have done some work along these lines; Rohwer and his associates have done a lot (cf. Rohwer, 1967, for a review), and I will try to summarize it very briefly before returning to the imagery topic.

Sentence Contexts in PA Learning

With primary grade children, Rohwer finds facilitation of noun PA learning when the nouns are connected by verbs or prepositions, but not conjunctions (and/or). This facilitation is assessed relative to control subjects who merely hear or read the noun pairs without a connective. Among adults, reading nouns in simple declarative sentences produces no net facilitation in learning. If adults are explicitly instructed to overtly rehearse the word pairs, then their recall is worse than that of subjects who study the nouns in the context of declarative sentences. In other experiments, (Bobrow and Bower, 1969), we have reported (1) that studying the nouns in anomalous or nonsensical sentences produces less recall than simply studying the noun pairs alone, and (2) the associative recall between subject-noun and object-noun is roughly symmetrical and about the same over several variations in syntactic form of the sentences studied (i.e., declarative, negative, interrogative, or negative interrogative).

One interesting finding is that subjects remembered sentences they generated much better than sentences we gave them. In the generate condition, subjects were shown noun pairs for 5 seconds, and they had to make up and say some sentence linking the two nouns. In the read condition, subjects were shown the two nouns capitalized as subject and object in a sensible declarative sentence, and had to read this aloud and study it for 5 seconds. Later recall (of the object-noun when cued with the subject-noun) is about one and a half to two times higher for the sentence-generating subjects. This difference has held up over four experiments. We now attribute it to a difference in comprehension of the sentence in the two cases. Generating a linking sentence forces the subject to search out, find, and
understand the relationship in a more reliable way than does mere rapid reading of a large number of similar sentences.

This led Bobrow and me to do incidental learning experiments in which subjects processed declarative sentences in ways designed to affect their comprehension of the propositions asserted. Comprehension of a sentence establishes a number of cognitive dispositions. So in one experiment we required our subjects to disambiguate sentences, in another to continue the action expressed in the sentences. Incidental recall of such subjects was compared to the recall of others who either searched for spelling errors in the sentences, or who read the sentences aloud three times as rapidly as they could. After processing 45 declarative sentences in this manner (at 7 seconds each), subjects were tested for recall of the object-noun when cued with the subject-noun. The disambiguation and continuation conditions produced about 50% recall whereas the spelling and rapid recitation conditions produced around 20% recall. This 2.5-to-1 difference presumably reflects the effect on memory of comprehension.

Unfortunately, deeper levels of analysis are stymied by the absence of a good theory of linguistic comprehension. There are various sketchy and incomplete proposals in the air. An appealing one is Quillian's (1966) which equates comprehension (of a simple declarative sentence) with finding and tracing a permissible pathway, corresponding to the predicate, between the semantic nodes — corresponding to the subject and object concepts. To comprehend a sentence is to find concepts preexisting in the semantic network that are related in ways compatible with the sentential relation. From the standpoint of memory, the process of understanding a sentence involves the construction or tagging of a retrieval scheme, whereby one semantic node points to a path which links it to the node of the other sentential concept. Thus, associative memory (between key concepts) is a by-product of comprehending a sentence.

Problems in Identifying Imagery Effect as Verbal Mediation

The excellent PA recall produced when subjects generate linking sentences suggests that the recall effects of imagery instructions are reducible to sentence generation. That is, subjects may be generating a linking sentence for the word pair. There may be no excess advantage to imagery once the effect of implicit sentence generation is partialed out.

I think this viewpoint is too extreme. A first fact is that if subjects read linking sentences, and also image scenes compatible with the sentences, their later PA recall is higher than it would be if they did not image. In one experiment, this produced a 62% versus 41% difference in recall; in another, a 69% versus 52% difference. Imagery subjects even recall somewhat better than subjects instructed to generate their own linking sentences for noun pairs. In one

Mental Imagery and Association experiment, the difference was 90% versus 82% differences are not very large, attributable to implicit sentence generation. But sentence-generating it spontaneously report having As reported by Yuille and Pa to generate verbal mediators the imagery value of the stimuli imagery stimulus terms are by generate verbal mediators. Cerebral mediators for abstract presence in verbal mediators later fact, abstract mediators were But there was a large difference.

Two Interconnected Systems

I think these attempts to likely to succeed. The alter imagery and verbal, with ver and Averbach, 1955). The i describe our imagery, to have assumption is that modality When the learning materials in the learning and the imagi upon both learning strategy verbal and imaginal repertoires (if given time), even pictorial in the verbal system. With use abstract paintings, visual inferverbal information.

According to this view, if conditions would cause the two imagery system and a corres vides a heuristic picture. Most interconnected verbal-sensory organized hierarchies con markers. For example, thing, level semantic tags correspond
experiment, the difference was 85% versus 73% in favor of the imagery subjects, in another, it was 90% versus 80%. Although statistically significant, these differences are not very large, so a portion of the imagery effect may still be attributable to implicit sentence generation.

But sentence-generating itself may not be a purely verbal process. Subjects spontaneously report having imagery of the scene described by the sentence. As reported by Yuille and Paivio (1968) and repeated by us, subjects instructed to generate verbal mediators for PAs will recall pairs at a level correlated with the imagery value of the stimulus member of the pair. That is, pairs with high-imagery stimulus terms are better recalled even when the subject is instructed to generate verbal mediators. One might attribute this to lower availability of verbal mediators for abstract pairs, but Yuille and Paivio (1967) found no difference in verbal mediator latencies for concrete versus abstract noun pairs. In fact, abstract mediators were reported slightly faster than concrete ones were. But there was a large difference in later PA learning in favor of the concrete pairs.

Two Interconnected Systems

I think these attempts to reduce imagery effects to verbal processes are unlikely to succeed. The alternative is to postulate two memory modalities, imagery and verbal, with very rich interconnections (cf. Paivio, 1969; Wallach and Averbach, 1955). The interconnections permit us to describe pictures, to describe our imagery, to have words evoking imagery, and so forth. The simplest assumption is that modality can be involved in most of our associative learning. When the learning materials are words, the verbal system is clearly involved in the learning and the imaginal system may be engaged to an extent depending upon both learning strategy and the strength of the connections between verbal and imaginal repertoires. Because adults usually label familiar pictures (if given time), even pictorial information will be partially coded and duplicated in the verbal system. With unfamiliar pictures, like cloud patterns or some abstract paintings, visual information may be stored with only very general verbal information.

According to this view, familiar picture pairs or word pairs under imagery conditions would cause the creation of two connected memory codes, one in the imagery system and a corresponding one in the verbal system. Figure 3.5 provides a heuristic picture. Meaning concepts correspond to nodes in an intricately interconnected verbal-semantic system. These nodes are arranged topologically in organized hierarchies corresponding to different categorizations or semantic markers. For example, thing, process, manner, and quality would be the highest level semantic tags corresponding to syntactic classes of nouns, verbs, adverbs,
and adjectives. A concept like DOG would be the name of a node pointing to interconnected sublists containing such information as:

(a) **Graphemic** — what the word looks like and how to spell it.
(b) **Phonemic** — what the word sounds like and how to articulate it.
(c) **Semantic** — definition of its meaning (several if ambiguous), subordinate and superordinate categories, verbal associates.
(d) **Sensory** — what the object looks like or sounds like. Criteria for class recognition, and lists for subspecies and individuals.

For ambiguous words, sublists of types c and d would be in coordinate pairs, with each meaning component tied to a particular sensory sublist. Such sublists would be used in stimulus recognition (of object or grapheme or acoustic stimulus) and in response generation. For the present it is immaterial whether we consider the stimulus recognition to be passive (categorizing feature lists) or active (matching generated patterns to those sensed). Illustrated in Figure 3.5 are the nodes and paths that are aroused and tagged by the phrase "a DOG on a BICYCLE." The sensory-information sublists corresponding to these semantic nodes may also be tagged and delivered to a production machine, which will generate an imaginal scene with the specified parameters.

The data suggest that the imagery code (or tag) has a slower decay rate than the verbal code. Perhaps subsequent materials are less likely to disrupt the imaginal memory code. We know from Shepard's (1966) experiments that recognition memory for easily identifiable pictures (e.g., magazine advertisements) is incredibly high, much better than for words. Gorman (1961) found a similar difference in recognition memory for concrete versus abstract words. Concrete words may be coded in both imagery and verbal forms, whereas abstract words may be coded only in verbal form. Free projects, next best with concrete results imply that the imagery sense is an important factor in memory storage and retrieval.

This dual-code theory of memory and learning task performance is the key to understanding the role of imagery and language in memory processes. The experiment shows that the production of sentences along associative pathways is influenced by the concreteness of the stimulus, with verbal production being faster and more accurate for concrete stimuli.

Mental Imagery and Association

This route information is linked in the experimental little to alter subjects' stores. It should be noted that the subject's store is not the same as in tables of association or the relation into which certain and BICYCLE fit because BICYCLE is not something. So by associating between terms.

There are many varieties of rhyme (rhymes) to subject-object ruled out by semantic rest in organization may be, they systems. There are similar to pictorial, images and sentences.

We may have a comprehension of sentences and drawn pictures. In this data that pictures have a "gram" which can be decoded for all the heuristics of Chomsky. 3-D may even be analogical to analysis (Shaw, 1968; Milner, and experimental, but we must focus on artificial intelligence. My speculation is that analysis and generation, or conceptual organization will...
Mental Imagery and Associative Learning

only in verbal form. Free recall is best with lists of familiar pictures (or real objects), next best with concrete nouns, and worst with abstract nouns. Such results imply that the imaginal code decays slower than a verbal code.

This dual-code theory still needs some assumptions concerning associative connections. The experiments on separate-imaging and type of connective (verb or preposition versus conjunctions) point to factors of relational organization. A bare PA learning task presumably activates a relation finding program which searches along associative pathways, fanning out from the two head nodes, looking for intersecting paths. Such associative pathways may be short or long depending on the semantic or phonological similarity (e.g., rhymes) of the key terms or their high associates. For example, a subject who learns the pair K-QUEEN via the path K-king-QUEEN is using the latter sort of route (cf. Schwartz, 1969). When it is successful, the relation finding program tags the associative pathway found, and attaches route information to the key nodes. This route information is usually temporary and is controlled by the set established in the experimental context. A few minutes' verbal learning task does little to alter subjects' store of semantic knowledge (cf. Slamecka, 1966). It should be noted that I am using *associative relation* in a very general sense, not as in tables of association norms. For example, "x rides on y" is an associative relation into which certain terms can be sensibly substituted for x and y. DOG and BICYCLE fit because our semantic network contains the information that bicycles can be ridden by something, and that a dog could sensibly ride on something. So by *associative pathway*, I mean a sensibly possible connection between terms.

There are many varieties of relational associations, from phonological overlap (rhymes) to subject-object predicates. A variety of predicative relations are ruled out by semantic restrictions. Whatever the general principles of relational organization may be, they appear to be similar within the imaginal and the verbal systems. There are similar patterns of associative recall whether the input material is pictorial, images aroused by words, or verbal connections established by sentences.

We may have a common generative grammar that underlies our verbal production of sentences and our imaginal production of visualized scenes or of hand-drawn pictures. In this day of computer graphics, it is hardly novel to suggest that pictures have a "grammar", that scenes can be parsed into subpictures, which can be decoded further into objects and contours. Such decoding contains all the heuristics of Chomskian grammar. Transformations of objects in imaginal 3-D may even be analogous to syntactic transformations. Grammatical picture-grammar (Shaw, 1968; Miller and Shaw, 1967, 1968) is still largely programmatic and experimental, but we can surely expect significant new ideas from this line of research in artificial intelligence.

My speculation is that a base grammar underlies our linguistic and pictorial analysis and generation. In particular, this suggests that Gestalt laws of perceptual organization will be the phrase-parsing rules of the picture grammar.
Pictorial constituents would correspond to perceptual groups, segregated according to the Gestalt principles of spatial proximity and similarity. Kohler (1947) and Asch, Ceraso, and Heimer (1960) have stressed the various perceptual organizations that lead to strong associations. If one figural component is a constituent of the other, or is a continuation of the other, or is an integral surface of the other, or is as part-to-whole with the other, then the figures are easily associated (cf. Asch, Ceraso, and Heimer, 1960; Prentice and Asch, 1958; and Kaswan, 1957). To illustrate, Figure 3.6a shows a constitutive relation (a square constituted of plus symbols) the terms of which are associated more easily than are those of control (Figure 3.6b). Similarly, Figure 3.6c shows a part-to-whole relation, the terms of which are associated more easily than are those of its control (Figure 3.6d).

Figure 3.6 Examples of figural-relations. In (a), the small plus signs are the mode constituting a square (unitary figure). In (b), a teardrop and a row of beads (nonunitary figure and mode). In (c), the two parts fit together as part-to-whole; in (d), the two parts are nonfitting.

It is probable that spatial relations like *is a constituent of, is composed of, is a surface of, is a part of*, and so on are going to be the basic relational predicates of a picture grammar, appearing as primitives in the most elementary statements about any scene. Thus, differences in associative memory for the unitary figure and mode of Figure 3.6a versus those in Figure 3.6b may be reflected in grammatical complexity of the two strings stored in the imagery system.

Whether or not these particular suggestions prove viable, I think we are in need of some heuristic but reasonably formal machinery for talking about relational organization. I am convinced that associating meaningful terms is largely a matter of relating them in some organization (see also Asch, 1969). And I think the mnemonic value of interactive imagery or sentence-generating devices arises from the creative production of relational organizations.
Mental Imagery and Associative Learning

ADDENDUM

In comments on my paper as delivered at the conference, Professor H. Simon asked what sorts of operations on information are possible with mechanisms in the imagery system, which are distinguishable from simple manipulation of verbal symbols. One class of tasks meeting these criteria comprise the major segments of batteries such as the Thurstone spatial abilities test, and the Minnesota paper form-board test (cf. Barratt, 1953). In such tasks, the subject is shown a spatial diagram, has to transform it geometrically in various ways, and must then select the transformed figure from a list of alternatives. Roger Shepard (personal communication) is measuring reaction times of subjects who are judging whether two 3-D rigid, block figures are identical except for rotation either in the vertical plane or in the picture plane. Essentially the subject asks himself, “Can I rotate figure A into correspondence with figure B?” Preliminary results show that errors are very infrequent and that reaction times for same judgments are a strict linear increasing function of the degree of rotation needed to convert figure A into figure B, with a slope of approximately 60 degrees per second of reaction time.

Pilot work has begun in my lab on another task, suggested by Lee Brooks. The subject is first taught a short stroke vocabulary using eight compass directions, for example, north, northeast, east, southeast, and practices drawing segments of unit length (about half an inch) in the appropriate direction when the vocabulary elements are spoken to him. The subject is then blindfolded and the testing begins. A sequence of four to ten strokes is read to the subject (approx. 1 per second) who draws them, connecting them into a continuous figure of line segments. He is asked to visualize the figure as he draws it. Figure 3.7 shows some example figures with the labelled strokes of the vocabulary as generated from left to right.

The subject is then asked to manipulate the information in one of several ways. In one task, he is asked to imagine a line connecting the first point with the last point in his visualized figure, and report how many line segments of the diagram are thus intersected. For the four figures in Figure 3.7, the number of crossings is 0, 1, 2, and 3, respectively. How well subjects perform on this test will surely vary with a variety of factors—the number of line segments, the redundancy or familiarity of the pattern drawn, the input rate of segments, whether the subject actually draws the diagram or merely visualizes it, and so forth. The important point of our pilot demonstration is that some subjects can perform these operations at a level far in excess of chance guessing. Significantly, the problem may not be soluble by a simple symbol-manipulating processor (operating on the list of strokes). Our heads must contain some geometry routines that allow us to treat sequential information as though it were being converted into a spatial diagram.
Figure 3.7  Examples of diagrams generated by stroke sequences (read left to right). Dashed line connects first and last points.

Another task that our subjects perform upon the completed diagram is to rotate it 90 degrees clockwise in their heads, and read out the names of the strokes that now appear in imagery. A final task we are now studying is immediate serial recall of the sequence of strokes. Half the subjects draw connected diagrams which they visualize as they draw them. The remaining subjects merely mark down the sequence of unconnected strokes in a line, and are asked to rehearse them verbally (rather than visualizing a continuous drawing). The hope is that those subjects who are drawing and visualizing a continuous, unitary figure will be better able to recall the stroke series than will those subjects who are simply recalling a series of compass directions. Preliminary evidence is encouraging, but many more subjects and conditions must be studied before we are sure of this conclusion.