The Concept of Memory

Abstract

The primary thesis here developed is that virtually nothing in modern research on “memory” has actually dealt with memory at all, for the simple reason that (put oversimply) this research has concerned itself with the retention of associations whereas memory proper is the recall of beliefs. Related objectives are clarification of the concept of “learning,” and introduction of a methodological distinction between process variables and state variables which has profound importance not merely for analysis of a behavior system’s formal dynamics but also for the practical development of psychological theory.

Among the great primordial concepts of psychology, few are so badly abused and poorly understood today as is “memory.” It is the word “memory” which is in this deplorable condition, not so much our knowledge of the realities to which the term has been applied, though to be sure there are many subtleties in these phenomena which still elude us and will continue to do so until the language in which we think about them is rehabilitated. In our technical and quasi-technical usage of “memory” and such related expressions as “remembering,” “memories,” “memory trace,” “recall,” “retention,” “learning,” and “information storage,” a number of fundamental distinctions and not-so-fundamental metaphors have become jumbled together in a monstrous snarl of ambiguity and confusion. This paper is an effort to tease apart the more important strands of this tangle.

Since conceptual analysis of the sort here assayed is seldom practiced or taken seriously by research psychologists,¹ let me attempt at the outset to reduce misunderstanding by listing some things which are not here at issue. (1) Although I shall complain that most technical uses of the word “memory” and its cognates are grossly at odds with the core meaning of these words and that very little of the research and discussion in which they appear is about memory at all, I intend neither to impugn the value of this research nor to deny others the right to pack new meanings into old symbols. I do, however, insist that if a word is to be redefined for some special purpose, this should be done knowingly, without loss of ability to

¹Early returns suggest that the prevailing reaction to the present endeavor is likely to substantiate a conclusion which I unhappily drew some time ago: “While we find a great deal of enthusiastic, if unskilled, discussion about methodology, particularly about general issues having little immediate implication for psychological research, . . . any attempt to do methodology in the course of an actual research problem is likely to meet only indifference, incomprehension, and at times open hostility.” (Rozeboom, 1961a, p. 473)
discriminate between the word’s old and new senses. (2) While I shall lay heavy emphasis upon the common-sense meanings of memory-words, this is not (never-never-never!) to imply, as is so frequent in contemporary linguistic philosophy, that this ordinary-language usage is in clean-limbed robust health, a worthy conceptual companion to precise insights into memory phenomena. It is only that the common-sense usage, despite all its deficiencies, has got hold of a vital ingredient of complex psychological functioning which is not addressed by current research on “memory,” yet which, precisely because it is the dominant theme of the common-sense concept, subtly colors, distorts and overgeneralizes the interpretation of this research. (3) The imperfections of which the present professional use of memory-words here stands accused are not just instances of the borderline fuzziness that to one degree or another roughens the cutting edge of any working concept. Rather, the charge is that this usage smears together polar distinctions both within and between at least two major dimensions of memory phenomena. Quarreling over the exact placement of category boundaries is a preciousity which has little point once the alternative possibilities are clearly perceived; here, the problem is to make the categories—or better, the continua from which they can be hewn—discriminable in the first place. As it is, in order to develop the major points of this paper on a scale broad enough to make them visible against the background of contemporary psychological thought, I have here had to employ oversimplifications, schematic arguments, and insufficiently clarified common-sense notions to an extent that does anguish to my methodological conscience.

To forestall the reaction that my theses about the nature of memory are already widely accepted and could be said more deftly in a fraction of the space, I should also add two further clarifications of intent: (a) While the substantive points I wish to make about “memory” are so intuitively familiar to common-sense psychology that exceptional perversity is required to dissent from them, it would be an abject non sequitur to conclude from this that these facts have received significant appreciation in the technical literature. The everyday meaning of “memory” still makes an occasional appearance in discussions of cognitive processes, but it seems to have vanished almost entirely from the mainstream of psychological thought. (b) Were the present concern merely to urge revival of old-fashioned language customs, a monograph would scarcely be needed for the purpose, especially insomuch as the particular sounds or shapes we choose to convey ideas which are themselves firm and unambiguous is for the most part a trivial matter. But my aim is far more ambitious than this—what I am ultimately after, and for which the present essay is largely blueprint sketching and foundation digging, is nothing less than a

---

2E.g., Ryan’s (1948) definition of “recollection,” and Reiff and Scheerer's (1959) concept of “remembrance.”

3A similar judgment has been passed by Rock and Ceraso (1964) on grounds which closely resemble some of the arguments developed in the present article.
wholesale re-examination and tightening up of how we think about psychological phenomena. It so happens that a particularly large number of seminal issues merge in our uses of memory-words, which is why I have directed the main flow of my argument through this channel. But the particular conclusions about “memory” at which we shall arrive are less vital than the way in which these are reached and the ground which is broken in transit.

From the phrasing of my opening indictment, it might be inferred that contemporary psychologists are in monolithic agreement in their usage and/or abusage of memory-words. This is, of course, egregiously false. Actually, there are four or perhaps five distinct conceptual traditions on the subject of memory, though many psychologists manage to embrace several of these at once. First of all there is the cognitive tradition, embedded in ordinary language and made most explicit in classical mentalistic psychology and its modern descendents. Secondly, there is a broad biological-memory tradition in which concern for brain mechanisms has been most prominent but which includes theories about the accumulated effects of environment upon living systems in general. Expressions such as “memory trace” and “fixation of experience” occur conspicuously in this tradition. Thirdly, there is the verbal learning tradition which flourished for several decades following its origin with Ebbinghaus, sank into decline during the ’30s and ’40s, and has recently hurst forth again with greater vitality than ever. It is in the verbal-learning literature that memory-words have been most extensively deployed in technical roles. Fourth, a distinctive conceptual tone heavy with metaphor-hardware has emanated from cybernetics theory, though this is perhaps too recent in origin and has too quickly infiltrated the older traditions to qualify as a separate movement of its own. And finally, the remarkable consistency with which specialists in behavior theory and conditioning manage to avoid the term memory and its cognates altogether4 demarks this, too, as a major conceptual tradition on the matter, namely, one which in effect denies that memory involves anything of behavior-theoretical importance which is not better addressed in other terms. Since the primary concern of this paper is methodological, not historical, I shall not attempt any systematic comparisons among these perspectives; however, some of their similarities and contrasts will be noted as we proceed.

The two critical dimensions of confusion—or at best of erratic discrimination—on which memory theory is currently in trouble are (1) the distinction between retention and repetition, which usually receives token recognition albeit little thoughtful appreciation, and (2) the difference between states or processes in the organism which are in some significant sense cognitive and those which are not. The first of these confusions has also massively contaminated our grasp of the logical nature of learning, and one useful way to get at the conceptual issues here

---

4For example, Kimble’s (1961) monumental compendium of conditioning data and theory contains not a single occurrence of “memory” or “remembering.”
involved is by comparing “learning” with “memory.” Before we begin to probe for meanings, however, some grammatical preliminaries are in order.

The Grammar of Memory-Words

So far as my linguistic sensitivities are able to discern, the generic concept of “memory” takes five primary grammatical forms in the English language, two verbs and three nouns. The verbs, “remember” and “memorize,” are easily recognized, but differences among the noun forms, which are morphologically indistinguishable, are rather subtle. The first of these is exemplified by the occurrence of “memory” in “I’ve searched my memory in vain,” “30 days to a better memory,” and “All the members of John’s family have excellent memories” (plural variant), and may be called the *abstract-particular* usage. This is an important contrast to the *thing-kind* usage which occurs in “He has a vivid memory of his great-grandmother” and “All of my memories are happy tonight” (plural variant), wherein “memory(ies)” is synonymous with “recollection(s).” In both its abstract-particular and its thing-kind contexts, “memory” is grammatically a common noun, but in the first case its definitive role is as a descriptor-radical in “x’s memory” which combines with person-identifying qualifiers “my,” “your,” “John’s” etc. to designate a single abstract object possessed by individual x and for which there exists no more direct description, whereas in the second case “memory” denotes a category of psychological entities which can occur in great profusion within each person and whose primary descriptions specify various detailed mental contents. The quickest way to appreciate the linguistic distinction between these two forms is to note that the phrase “my memories” is grammatically ill-formed in the abstract-particular sense, whereas in the thing-kind sense it denotes with perfect propriety a certain subclass of memories, namely, those which happen to be mine. Finally, “memory” takes still another grammatical form, which I shall call the *phenomenon-demonstrative* usage, illustrated by “We still don’t understand memory very well,” “Miss Smith assigns a lot of memory work to her students” (adjectival variant), and “Memory plays an important role in problem solving.” That this usage is syntactically distinct from the preceding two may be appreciated by observing that none of the phenomenon-demonstrative examples have a meaningful plural variant.

The reason why the niceties of *de facto* memory-word usage need to be made explicit is that only in this way can we come to recognize the beliefs about memory to which we are committed by a particular form of speech, and to appraise the extent to which our technical research on memory substantiates, refines, repudiates or is otherwise relevant to these beliefs. Not all of these usages are equally important, however, and I shall deal with three of the five rather summarily.

Perhaps most peripheral of all is “memorize,” the primary occurrence of which
is in phrases of form “s memorizes x,” where s is a person and x—a speech, a musical composition, a role in a play, etc.—is some complex action. The standard dictionary definition of “to memorize x” is “to commit x to memory,” in which “memory” occurs in its abstract-particular sense; however, in view of what we shall have to say about the latter, a better translation might be “to develop the ability to perform x,” which reduces memorizing to learning. In any event, there is probably nothing to be learned from analysis of “memorize” which cannot be pulled more directly out of other memory-words.

Neither need we loiter over the abstract-particular sense of “memory,” even though this is the form which occurs in ordinary language with perhaps the highest frequency of all. This usage is simply a hangover from faculty theory, which still dominates common-sense psychology but has largely disappeared from the more respectable professional literature. In this sense, a memory (yours, mine, John Smith’s) is a mental organ which its owner uses to remember with, just as he breathes with his lungs and walks with his legs. The primary image here is of memory as a tool, though there is also a secondary warehouse connotation which appears in such phrases as “I’ve searched my memory” or “his memory has tremendous capacity,” and which implies that one’s memory is a place wherein specific memories are stored. While technical memory theory is not altogether untainted by the warehouse metaphor, especially where the cybernetics influence is appreciable, and it is possible to imagine empirical findings which would justify a reconstructed abstract-particular concept of memory (e.g., verification of a strong common factor specific to performances on memory tasks, or evidence for a long-term memory phenomenon analogous to memory span), it is probably safe to say that few research psychologists today would regard “memory” in this sense as anything more than a figure of speech or have their thinking about memory phenomena contaminated by it to any appreciable degree.

Next we come to the phenomenon-demonstrative sense of “memory.” This might alternatively be called the evasive usage, for its function is to pertain to memory while remaining deliberately vague about a precise referent. Consider, for example, such statements as “Memory depends upon structural changes in the nervous system,” “Memory is an important factor in the behavior of all higher organisms,” or “Memory is a complex phenomenon.” These have the logical form “Φ(memory),” in which “Φ” is a predicate ascribed to a singular entity memory dis-associated from any particular organism or mental content. (This in contrast to the abstract-particular usage in which “memory” is only a radical in the description-schema, “x’s memory.”) Grammatically, the phenomenon-demonstrative usage is an ultimate platonic reification of memory-concepts; yet clearly this would be an overly harsh interpretation of the cited statements. All that is really intended by use of “memory” as a singular term here is to direct attention to a poorly delimited region of concern, rather like the use of “that” in “That’s what I call beauty”
uttered in the presence of some spectacular scenic or feminine display. To give this sense of “memory” an impressive title, it might be said to function as the name of a certain class of phenomena—which, however, still leaves thoroughly vague just what phenomena are memory-phenomena, or even what, formally, a “phenomenon” is in the first place. In any case, the phenomenon-demonstrative sense of “memory” is the most noncommittal of all memory-words, and rests entirely upon the others for its meaning. It might, in fact, be paraphrased as “whatever is involved in remembering and having memories.”

Thus we arrive at “remember” and whatever expressions are considered to describe specific thing-kind memories as the fountainhead of our linguistic commitments to the nature of memory. To close in on what these commitments are, we may note to begin with that the primary sentence structure in which the verb participates is “s remembers x,” while for the noun it is “s has a memory of x.” There is a delicate distinction between these two forms in that the verb-schema admits a psychology of mental acts wherein remembering is something the subject operantly does to entity x, while the noun-schema requires only a passive subject in whom memories can come and go. However, if we ignore or expressly repudiate such subtle overtones, “remembering x” seems to be essentially synonymous with “having a memory of x,” at least for most instantiations of “x.” (More on this later.) The ultimate logical question about the concept of memory, therefore, is: What is remembered?—i.e., what is purportedly designated by expressions which are grammatically proper instantiations of “x” in “s remembers (has a memory of) x”? This is strictly an issue about the use of language and has nothing to do with the facts of memory phenomena, though of course it could well occur that this usage rests upon presuppositions which are at odds with contemporary scientific knowledge. Parallel to the logical question, however, is a factual one: If assertions of the form “s has a memory of x” are ever true, then there exist psychological attributes describable by predicates of form “having a memory of x.” These cannot—or rather, must not—be confused with what is remembered for the simple reason that whatever x may be, it is logically distinct from any memory of x. The foremost empirical problem of memory, then, is: What are memories?—i.e., what condition of the organism constitutes having a memory? (Only after we have some idea of what memories are can we begin to study what is involved in their acquisition and loss.) Finally, completing the roster of definitive problems of memory is: How are memories about—i.e., what is the nature of the relationship

---

5If the reader thinks that he knows what the word “phenomenon” means in scientific discourse, I invite him to construct a definition which can actually be applied to discriminate phenomena from non-phenomena, or, failing that, to write down on paper a grammatically well-formed description of some “phenomenon.” Actually this is only one of many prestigious methodological words which everyone loves to throw around but no one cares to clarify. As a first approximation, I would offer phenomenon = def a recurrent pattern of events, but this only shifts the burden of analysis to spelling out what is meant by the definiens.
between a memory and what it is a memory of?

Memory as Retention

Those currently active branches of scientific endeavor which have most explicitly identified memory as a subject of their concern are the verbal learning and biological-memory traditions, and it is thus reasonable to expect the literature in these areas to yield a well-articulated grammar of “memory.” Actually, search of this literature reveals a remarkable fact: In technical writings on memory processes, memory-words scarcely ever occur except in the evasive phenomenon-demonstrative sense. This is especially conspicuous in the verbal learning literature, which abounds with technical expressions such as “memory span,” “immediate memory,” “memory trace,” etc. denoting various phenomena or theoretical entities whose more detailed descriptions are altogether innocent of memory-words and where nothing is said to convey how these phenomena or entities relate to memory—i.e., to what ordinary language is trying to get at with its memory-words. One partial exception to this rule is the important recent theoretical article by Melton (1963), in which the “proper domain of a theory of memory” is described as follows:

After some exclusions that need not concern us here, learning may be defined as the modification of behavior as a function of experience . . . It is useful to keep in mind the fact that learning is never observed directly; it is always an inference from an observed change in performance from Trial $n$ to Trial $n+1$. Furthermore—and this is the important point for theory—the observed change in performance is always a confounded reflection of three theoretically separable events: (i) the events on Trial $n$ that result in something being stored for use on Trial $n+1$; (ii) the storage of this product of Trial $n$ during the interval between Trials $n$ and $n+1$; and (iii) the events on Trial $n+1$ that result in retrieval and/or utilization of the stored trace of the events on Trial $n$. For convenience, these three theoretically separable events in an instance of learning will be called trace formation, trace storage, and trace utilization.

Obviously, a theory of learning must encompass these three processes. However, it must also encompass other processes such as those unique to the several varieties of selective learning and problem solving. Some advantages will accrue, therefore, if the domain of a general theory of memory is considered to be only a portion of the domain of a theory of learning; specifically, that portion concerned with the storage and retrieval of the residues of demonstrable instances of association formation . . .
The implication of this restriction on the domain of a theory of memory is that the theory will be concerned with post-perceptual traces, i.e., memory traces, and not with pre-perceptual traces, i.e., stimulus traces. It seems to me necessary to accept the notion that stimuli may affect the sensorium for a brief time and also the directly involved CNS segments, but that they may not get “hooked up,” associated, or encoded with central or peripheral response components, and may not, because of this failure of being responded to, become a part of a memory-trace system . . .

What, then, are the principal issues in a theory of memory? These are about either the storage or the retrieval of traces. (Melton, 1963, p. 1ff.)

What Melton says here about memory is more complex (and, as shown later, more questionable) than is immediately apparent, but it stresses a theme that is near-universal in the verbal learning and biological-memory traditions. Compare the following passages by the noted neurophysiologist, Ralph Gerard:

Exposure to light makes [linseed oil] turn gummy. A brief exposure may not cause any observable changes. But on later illumination the oil will change more rapidly than if it had not already been exposed. The oil “remembers” its past experience and behaves differently because of it. Its memory consists in the fact that light produces, among other things, substances which aid the light-induced oxidations that make it gummy.

However far removed this may be from remembering the Gettysburg address, it clearly points up one way in which memory can work—by means of material traces of the past—and the difficulty of defining what memory is . . . any concept which defines memory more narrowly than “the modification of behavior by experience” [italics added] will run into trouble . . .

Where then, shall we draw the line? A pebble, rubbed smooth in a stream, rolls differently from the original angular stone. Experience has here modified behavior; the past has been stored in a changed structure. Yet this does not greatly interest us as an instance of memory. Perhaps we should restrict the notion of memory to changes in systems which participate actively in causing the change. The linseed oil “remembers,” and so does the bulging calf muscle of a ballet dancer. Does a developing embryo “remember” the major steps, and missteps, in the long evolution of the species? Do trees “remember” good and bad seasons in the thickness of their rings? Is a film a memory of light in chemicals and a tape recording a memory of sound in magnetism?
Is a library a memory of thoughts in books and a brain a memory of thoughts in protoplasm? Even to identify memory, let alone explain it, is no simple matter... (Gerard, 1953, p. 118f.)

Memory involves the making of an impression by an experience, the retention of some record of this impression and the re-entry of this record into consciousness (or behavior) as recall and recognition.

Under certain conditions the interaction of the system and its environment leads to irreversible changes; the system has altered as a result of its experience. It fixes its experience and so becomes something different, and this I like to call “becoming.”

Becoming subsumes, of course, development of the individual, evolution of the species, history of the particular society or social group of any kind, and learning in the individual. And learning may include, if you accept a broad definition, changes as varied as: the hypertrophy of a muscle with exercise; the horny hands of a laborer; and the many other material changes that record the past—as in that lovely couplet on weatherbeaten trees:

Is it as plainly in our living shown,
By slant and twist, which way the wind hath blown?
That is memory in trees... Fixation of experience... is the basis of becoming and... includes in it memory. (Gerard, 1963, p. 22f.)

What do memory-words mean according to Gerard? His one suggested formal definition, namely of “memory” as a phenomenon (italicized above), is virtually identical with Melton’s definition of “learning,” while the rhetorical style through which the other usage-forms are sprinkled amounts to a clever repudiation of responsibility for any untoward implications that might accrue from taking them literally. Even so, there is an important similarity between what Melton and Gerard conceive to be the general pattern of events to which the term “memory” is to be attached somehow: Something happens which changes an organism’s (or system’s) condition in some respect, and this altered condition then persists in a form that is capable of being made manifest in some way. This formal schema has been made admirably explicit by Dingman and Sporn:

Ashby notes that “‘Learning’ and ‘memory’ have been given almost as many definitions as there are authors to write of them... but the theme is that a past experience has caused some change in the organism’s behavior...” Ashby also emphasizes the tenacity with which memory traces may persist; the persistence of such traces gives adaptive potential to the organism. Memory must therefore involve two processes: (1) the process of transition of a system to a new state...
and (2) the persistence of that new state. At this point, we shall dis- 
tinguish between the terms “memory,” “memory trace,” “recall” and 
“learning” as they will be used in this paper. While memory is the 
process of transition of a system to a new state and the persistence of 
the new state, a memory trace is the particular new state that has been 
formed. Recall is the process of using a memory trace. Learning is then 
defined as any process in which one can demonstrate the processes of 
memory and recall. (Dingman & Sporn, 1961, p. 1)6

In all these passages, what stands out most conspicuously is the view that 
whatever else the phenomenon of memory may be, it is at least the retention of 
an acquired characteristic. That this is, in fact, an interpretation which many 
psychologists profess to find congenial is attested by, e.g.,

Memory [is] the retention of acquired skills or information. (Miller, 
1962, p. 349)
“Retention” and “memory” are practically synonymous. (Lawson, 
1960, p. 400)
Memory [is] retention of what has been learned. (Munn, 1961, 
p. 718)
Basically, memory is the retention of contents, events or activities 
in some form by the individual. (Reiff & Scheerer, 1959, p. 24)

But this is agreement only in broad outline. Is the retention of any acquired 
characteristic an instance of memory? The biological viewpoint (e.g., Gerard and 
Dingman & Sporn) is disposed to answer yes, at least if the retention is by a 
living system, whereas the quotations from Melton, Miller and Munn suggest that 
memory-type retentions are restricted to those characteristics which can be learned.

Moreover, all the above quotations (with the partial exception of the one by 
Gerard) concern “memory” only in the phenomenon-demonstrative sense. What, 
more crucially, are memories and remembering? The biological and verbal learning 
literature has little to say about this, but since in these accounts the memory 
trace—i.e., the persisting acquired characteristic—is most conspicuously what 
the organism has, one might conjecture that for biological-memory and verbal 
learning theorists, the concept of “memory trace(s)” is intended to be a technical 
replacement for the everyday thing-kind sense of “memory(ies).” The verb-form 
counterpart of this would then be remembering = having a memory = retaining 
(storing) a memory trace (cf. Munn, 1961, p. 450: “Remembering is retaining”). 
On the other hand, each of the three longer quotations above also mentioned a

---

6I am indebted to R. C. DeBold for calling my attention to this reference, as well as for other 
less tangible contributions to this paper.
second thing, additional to retaining, which organisms do in memory phenomena, namely recalling, retrieving, or using the trace, and this might also be reasonably proposed as an interpretation of "remembering." (Cf. "To remember means to show in present responses [italics added] some signs of earlier learned responses" (Hilgard, 1962, p. 288); "In discussions of remembering, recall denotes the arousal of a memory trace" (Miller, 1962, p. 352).) But what does it mean to "use" or "to recall" a memory trace, anyway?

We have now pushed into what, for the verbal learning and biological-memory traditions, are muddy and shoal-laden waters. Before the bite of logical problems about "trace retrieval" and the significance of ordinary-language commitments to the nature of remembering can be fully appreciated, we need to polish up some methodological distinctions which are most accessible through analysis of the concept of "learning."

The Conceptual Anatomy of Learning

What is meant by "learning"? For many years psychologists have been volunteering definitions for this term, albeit usually in a spirit of impatience to get on with the facts, and their efforts have shaken down into what is by now a remarkably consistent formula. The first sentence in the quotation by Melton, above, is one recent example. Here are some others:

Learning [is] a more or less permanent modification of behavior which results from activity, special training, or observation. (Munn, 1961, p. 716)

Defined generally, learning is a change in behavior resulting from practice. (Kendler, 1963, p. 151)

Learning [is] a general term referring to a relatively permanent change in behavior that is the result of past experience. (Morgan, 1961, p. 677)

Learning [is] adaptive change in thought or behavior. (Miller, 1962, p. 349)

Learning [is] the more or less permanent modification of the response or responses to a stimulus or to a pattern of stimuli that results from experience with these or with similar stimuli. (Wickens & Meyer, 1961, p. 746)

Learning is a relatively permanent change in behavior potentiality which occurs as a result of reinforced practice. (Kimble, 1961, p. 6)

The first four of these, together with Melton’s contribution and still others collected by Kimble (1961, p. 2ff.), are all variants of the following basic pattern:
Learning is a relatively permanent change in behavior (i.e., in what the organism is doing) which has been brought about in a certain restricted way, where there is some disagreement about just what sources of behavioral change are to count as productive of learning. The last two offerings, by Wickens & Meyer and Kimble, do not quite fit this pattern, however, and therein lies a distinction which is of inestimable importance for behavior theory.

Let us see what happens if we take the standard learning formula literally. Consider the sequence of events lying behind some conventional learning curve, say a subject’s time-on-target plotted as a function of trials on a tracking task. The only suggestion of a relatively permanent change of behavior here occurs where performance is asymptotic, whereas we would normally consider the greatest amount of learning to have occurred when the curve was rising most rapidly—i.e., when the trial-to-trial behavior is most unstable and when behavioral changes have the least permanence. Again, consider Pavlov’s dog with the conditioned salivation reflex. Has any relatively permanent change in behavior been established here? The relevant behavior, presumably, is salivation, but—persistent salivation? Both before and after conditioning the dog is salivating at some moments, not salivating at others, nor is it even necessarily the case that the proportion of his day engaged in salivary activities is appreciably greater after conditioning than before. This example also serves to bring out that the vagueness of “relative permanence” is not the source of awkwardness here. How long the subject persists in salivating following a given learning trial has no logical connection with the extent to which learning has occurred on that trial. Still again, suppose that we take an animal which has been maintained at a normally warm external temperature and now place it for the rest of its life in a frigid environment. This experience causes our subject to undergo a permanent change in behavior, namely, sustained piloerection and probably a great many other behavioral adjustments as well; hence under the standard definition, the elicitation of piloerection by cold is an instance of learning, at least if the cold is sufficiently prolonged.

It is clear from these examples of what happens when the standard definition of learning is taken seriously that something is disastrously wrong with it. The wrongness lies in pretending that learning is ever a change in behavior, permanent or otherwise. With the possible exception of certain time-average measures, there is no such thing as relatively permanent behavior. What organisms do varies from moment to moment according to the environmental circumstances with which they are confronted. Stimulation, or change of stimulation, is the major immediate source of a change in behavior, but this is what, broadly, we mean by elicitation, which is altogether different from learning. Rather, learning is a suitably induced change in some condition of the organism which, together with the particular stimulus circumstances and perhaps other factors, is responsible for the organism’s behavior at a given moment, but which, unlike behavior, is relatively unaffected.
by moment-to-moment variations in the environment.

In order to make this last point with sufficient clarity and generality, it is helpful to introduce some formal concepts about the dynamic properties of a causal system. (A technically precise development of these is unfortunately impractical on this occasion, but the heuristic sketch which follows should suffice for present purposes.) Formally, an organism may be thought of as a reactive system in recognition that many of the organism’s properties at a given moment are importantly a function of the external (i.e., extra-organismic) conditions impinging upon the organism at that or a just-previous moment. That is, to use a well-worked analogy, a behaving organism may be compared to a computer or other responsive device whose output variables are a function of its input variables. (In this paper, the term “variable” occurs throughout in its scientific sense, namely, to denote sets of attributes which are mutually exclusive and exhaustive over the argument-domain in question, here time-slices of organisms. For details, see Rozeboom, 1961b.) Now, the variables which in totality characterize a reactive system—i.e., the variables whose constellation of values for the system at a given moment is an exhaustive description of the system’s condition at that moment—differ remarkably in the degree to which they are influenced by variation in the external-surround, or “input,” variables. For example, of the variables which characterize the total condition of a given computer, those which describe the readings of its print-out components depend strongly upon what particular numbers were fed into the computer a moment earlier, whereas the computer’s program variables—i.e., those whose values determine what the computer does with its input—are not in general affected by the input data at all. Similarly, while complete description of an animal’s momentary physical properties includes both the length of each limb-segment and the postural angles among them, the latter sustain very little moment-to-moment consistency within their ranges of possible variation, whereas the lengths remain essentially constant. To be sure, both a computer’s program variable and an animal’s limb-lengths are affected to an extent by the system’s external-surround. The computer’s program is changed from time to time, and the animal metabolizes nutrients which usually eventuate in some form of growth. Neither are these latter changes always gradual—quantum-jump transitions are normal for alterations in computer programs, and can occur even in limb-length through, e.g., traumatic encounters with surgeons or bear traps. Even so, the overall sensitivity to transient environmental details of computer output or postural angles on the one hand, and computer program or limb length on the other, are just not on the same order of magnitude. On the basis of such differences in reactivity, the variables constituting a reactive system may be classed as process variables on the one hand and state variables on the other, “process” variables being those whose stability is essentially no greater than that of the input variables (where the latter may also be included among the system’s process variables), while “state” variables
are those which remain essentially constant, though generally not altogether so, under major variation in the process variables. Correspondingly, specific values of state variable may be called state properties, while values of process variables are process properties or, in recognition of the flux in processes, process stages.

(Somewhat more precisely, though still heuristically, system variables $X$ and $Y$ are a process variable and a state variable, respectively, with respect to each other if there is a temporal duration $d$ such that the correlation between the values of $X$ for the system at the beginning and at the end of time intervals of duration $d$ is close to zero, while the correlation between the system’s values of $Y$ at the beginning and at the end of time intervals of duration $d$ is close to unity. In this more technical formulation, it is possible for a variable $Y$ to be a state variable with respect to one variable $X$ and a process variable with respect to another variable $Z$. In behavior theory, for example, deprivation drives are state variables when assayed against the flux or stimulation and behavior, but become process variables relative to the stability of habits. We shall, however, continue to treat the state/process distinction as a dichotomy corresponding to the powerful intuitive polarity between psychological attributes which are essentially stable and those which are not.)

While the distinction between process variables and state variables is, of course, ultimately a matter of degree rather than of kind, it is implicitly recognized in virtually all phases of psychology, especially those which are concerned with individual differences. For the psychological state properties of an organism are those of its psychological attributes such as habits, values, abilities, personality traits, etc. which transcend—i.e., reliably persist apart from—the passing circumstances of the moment and hence include all that is psychologically distinctive about this particular organism independent of its immediate environment. In contrast, the process stages in the organism at a given moment are those conditions, such as the excitations in its sense receptors, its percepts, thoughts and expectations, and the particular pattern of its motor arousal, which come and go with the transient details of its stimulus-surround, internal rhythms or other possible sources of moment-to-moment variation. Epigrammatically, an individual’s process properties are what he is doing and what is happening to him or in him, whereas his state properties are what he is. In particular, dispositional attributes characterized in terms of what process properties the individual would have were he to be exposed to a certain input condition—e.g., having a disposition to utter the word “black” when exposed to the soundsequence “Tell me the first word you think of when I say...

---

$^7$A psychological property is an attribute which in some sense is directly involved in psychological phenomena. While this seems like a hopelessly vague notion, it is one which in fact we repeatedly attempt to apply, as when, for example, a person’s motives and abilities are considered to be “psychological” while his hair-color and shoe-size are not. Analysis of what underlies this intuitive distinction is beyond the present scope.
White,” or a tendency to feel frightened at sight of a shrunken head—are (in general) state properties which persist even when these input conditions are not in fact present.

While the distinction between state variables and process variables is initially grasped most easily as a difference in temporal stability, this is only ancillary to its primary methodological significance, which is that each state variable causally involved in the doings of a reactive system is to a greater or lesser extent mirrored in mutable interdependencies among the system’s process variables. To show this as simply as possible, suppose that \( Y \) is an output variable whose value for the system at a given moment is jointly determined by the values of input variable \( X \) and state variable \( S \). For example, \( X \) might be a stimulus variable whose alternative values for organism \( o \) at time \( t \) are whether or not \( o \) is exposed to a flash of light just prior to \( t \), \( Y \) a response variable whose alternative values are whether or not \( o \) blinks his right eye at time \( t \), and \( S \) a reflex-strength variable whose value for organism \( o \) at time \( t \) is the probability that \( o \) blinks his right eye at \( t \) if exposed to the light flash just prior to \( t \). The general relationship between these three variables may be expressed by an equation of form \( \dot{Y} = f(S, X) \), in which the best estimate dot over \( Y \) avoids assumption that \( Y \) is perfectly predictable from \( S \) and \( X \). Suppose that state variable \( S \) remains essentially constant at value \( S \) during a certain period of the system’s history. The relationship \( \dot{Y} = f(S, X) \) then simplifies for this system during this period to \( \dot{Y} = f_S(X) = \text{def} \ f(S, X) \), in which \( Y \) is a function only of \( X \), but where the manner in which process variables \( X \) and \( Y \) are coupled depends upon state-property \( S \). Now suppose further that while \( S \) remains essentially constant at \( S \) for a given system through periods which are long enough to permit testing the system’s response to various values of input \( X \) and hence allow empirical determination of the relationship \( \dot{Y} = f_S(X) \), the system’s \( S \)-value alters appreciably over more extended intervals of time and/or we have observational access to an ensemble of similar systems, all of which are governed by the general law \( \dot{Y} = f(S, X) \) but which have different stable values of \( S \). Then the local dependency of \( Y \) upon \( X \) (i.e., the relation between \( X \) and \( Y \) at a particular brief period in the history of a particular system), as characterized by parameters in \( f_S \) which take different values from system to system and from time to time in the history of the same system, is an empirical structural variable (see Rozeboom, 1961b, p. 357) which acts as an observable counterpart to state variable \( S \) no matter how inaccessible to direct observation \( S \) may itself be.

(Our previous heuristic definition of the state-process distinction may now be amended as follows: With respect to each other, \( X \) is a process variable and \( S \) is a state variable if \( S \) has sufficient local constancy relative to the variability in \( X \) that the local relation of \( X \) to other process variables can be ascertained under an essentially fixed value of \( S \)).
Conversely, whenever local interdependencies among process variables are discovered to differ among comparable systems, or show long-term changes within the history of the same system, we usually find ourselves postulating the existence of theoretical state variables of which these local process-relations are the observable manifestations; while the empirical regularities which are found to determine the parameters in the local process-relations are the source of the laws hypothesized to govern these theoretical state variables (see Rozeboom, 1961b, p. 362ff). It is altogether possible that some of the writers who, during the heyday of the intervening-variable/hypothetical construct controversy, argued that “intervening variables” are the relationships between data variables were groping toward some version of this point. However, it is important to be clear that state variables, whether observed or inferred, do not intervene in process relationships in the classic fashion of a mediation variable \( M \) which partitions an observed local process-dependency \( Y = \Phi(X) \) into a product of relations \( Y = \Phi_1(X) \) and \( Y = \Phi_2(X) \). Such a mediator has the same trans-situational instability as \( X \) and \( Y \), and is thus likewise a process variable—which also serves to emphasize that the concept of “process variable” includes theoretical processes hypothesized to take place within the organism as well as observable stimulus input and response output. Inference of mediational processes which account for a particular local pattern of observed process interdependencies must not be confused with the inference of state variables from the broader-scale variability in these process interdependencies.

And now back to “learning” and, eventually, “memory”. This section began with the observation that the standard definition of “learning” as a subclass of changes in behavior is fundamentally misdirected as an approximation to what we intuitively mean by this term. More generally, none of an organism’s transitions from one process stage to another, internal or external, are instances of learning. Rather, the concept of “learning” attempts to recognize the fact that most trans-situationally stable psychological attributes are nonetheless modifiable by experience. The proper definition—or more precisely, explication—of “learning” is thus

\[
(D1) \quad \text{learning} =_{\text{def}} \text{change in a psychological state variable due to experience}
\]

where “experience” is a poorly defined term which attempts to focus upon sensory processes and to exclude state changes due to such nonexperiential causes as brain damage, effector-organ hypertrophy,\(^8\) and the like. While definition (D1) still contains considerable vagueness, I submit that this is precisely what is vague

\(^8\)Thus the version of the standard learning definition which holds that learning is a change in behavior resulting from practice is defective both fore and aft, first because it is not changes in behavior which count as learning, and secondly because changes due to practice are not necessarily
in the common-sense usage of “learning.” That is, whatever we mean by “psychological,” “experience” and “relatively permanent,” we would be willing to classify a change in an organism as an instance of learning if and only if we were also prepared to consider it an experience-induced modification of a relatively permanent psychological condition.

What is at stake in the definition of “learning” is not just the wording of a glossary entry, but a basic cognitive orientation toward the interpretation of psychological phenomena. Ever since the behavioristic revolution, tough-minded psychologists have been struggling to get clear about how theory should or should not be done in psychology, with many empiricists taking the position that learning theory, which has been home base for behavioristic perspectives, is properly concerned only with the observable relations of behavior to its environmental antecedents. Others, uncomfortable with so peripheralistically austere a psychology, have argued for a distinction between “learning” and “performance” and have popularized the slogan that “learning is an inference.” Unfortunately, the latter offers no enlightenment concerning what learning-as-inference is an inference about, while to assert that learning is not the same as performance (which would be true even if learning were nothing more than a change in performance) likewise leaves obscure the way in which they differ. However, once it is made explicit that the very meaning of “learning” envisions changes in relatively stable psychological attributes which in general are not directly observable but can only be inferred from local stimulus-response regularities, it also becomes evident that learning research is inherently committed to study of the organism’s internal psychological constitution. On the other hand, while the concept of “learning” is theory-dependent, (D1) is also theoretically neutral in that it contains no presuppositions about what kinds of psychological states underlie behavior. In particular, there is no reason at all to assume that the only experience-modifiable state variables are the traditional objects of study in learning research, namely, abilities, reflexes, S-R habits and/or expectancies, and verbal associations. The definition of “learning” thus reveals the fundamental substantive problem of learning theory to be What is learned?—not just in the familiar restricted sense of the S-R vs. expectancy controversy, but in dogma-shattering generality: What is there in the organism that makes a difference for how it reacts to the environment, and what observable features of an organism’s process-regularities justify the belief that it has internal states of these particular kinds?

accomplished through the mediation of sensory input. Suppose, for example, that as a weight-lifter diligently practices year after year, his muscular strength continues to grow long after his coordination techniques have reached asymptote. This last period of improvement in his weightlifting ability is then a growth of ability due to practice, but not to experience—i.e., his sensory processes make no difference so long as his muscles get suitably exercised—and would hence no longer be intuitively regarded as learning.
Learning vs. Memory

And what has the definition of “learning” to do with “memory”? Clearly these concepts must be distinguished, if only because “learning” addresses change while “memory” is concerned with something that happens after a change. It is, however, instructive to explore the extent to which the two concepts are *adjuncts* in the sense of applying to different phases of the same sequence of events. In particular, is memory always preceded by learning, and do all retentions of learned attributes qualify as instances of memory? Or less evasively, though not altogether equivalently, are memories learned and is everything which has been learned a memory?

To begin, let us recast our previous conclusions about the verbal learning and biological-memory usage of “memory” in terms of the state vs. process distinction. We concluded earlier that for both of these traditions, whatever else memory may involve as a phenomenon, it is above all retention. But retention of what, and by what? Let us confine our considerations to retention in living organisms, not because it may not prove literally appropriate to speak of memory in, say, computers, but merely because our paradigm uses of memory words apply only to organic cases. Very well, then, is any persisting attribute of an organism an instance of memory? Does, for example, the retention of my chest’s inflation as I hold my breath count as short-term memory, or is long-term memory illustrated by my sustained heart-beating behavior? No published passages of which I am aware reveal a biological or verbal learning stand on cases such as these, but they would undoubtedly be repudiated on grounds that they feel too remote from common-sense usage. More generally, there are probably very few persons, who if pressed, would be willing to class as “memory” a period of constancy in a process variable resulting from constancy in the stimulus conditions which control it. But what about hysteresis effects, wherein a process variable lags in its accommodation to a change in another process variable of which it is a function? Consider for example, the after-discharge momentarily persisting in a sense-receptor after cessation of the impinging stimulation, or the deformation of an organism’s bodily surface which lingers briefly after sudden removal of a hard object which has been pressing against it—are these technically instances of “memory”? While most persons would probably agree that these particular examples should be excluded, physical scientists do, on occasion, refer to hysteresis as “memory,” and much of the recent work on “immediate memory” would seem primarily to concern non-instantaneous decay of the internal process-effects of stimulation. On the other hand, Melton (above) is explicit that mere perseveration of a stimulus trace is not enough; for memory to occur there must be formation of “associations,” which is what memory traces are. For the sake of tidy formulation, let us follow Melton’s lead and conclude that in the verbal learning tradition, at least, mere hysteresis in stimulation-dependent process variables does not qualify as “memory.”
If memory is retention, therefore, it is presumably retention of state properties. But is the retention of any state property an instance of memory? The earlier quotations from Gerard and Dingman & Sporn indicate that biologists are disposed to answer affirmatively, at least if that state was initially brought about through impact of the environment. But this would include under memory the retention of scars, chips on teeth, surgically altered nose shapes, and the like, which are highly remote from what psychologists consider to be memory phenomena. Ordinary language feels even more strained (just why will be seen shortly) if it is claimed that scars and chips on teeth are memories. But we need not be jealously possessive. If there is desire in some quarters for a technical concept of “biological memory” which subsumes all retentions of externally initiated state properties, so be it—at least if it is very clear that this is a neologism which may have no significant relation to the meaning of “memory” in ordinary language or technical psychology.

Where the verbal learning tradition would draw the line between memory-type retentions and non-memory retentions is, to be sure, obscure. As a basis for discussion and not because it necessarily does full justice to the view of any particular verbal-learning theorist, let us adopt

(D2) \[ \text{memory trace} =_{\text{def}} \text{a psychological state property acquired through experience}, \]

(Even without the reservation that the retained state be a “psychological” one, this would exclude scars and chipped teeth in that while acquisition of such blemishes is usually accompanied by sensory input, the latter is not instrumental in causing the former. Whether or not (D2) agrees with Melton’s interpretation of “memory trace” depends upon whether or not he would consider there to exist any experience-acquirable state properties which are not associations.) From (D2) and (D1), it follows that

(T1) \[ \text{learning} = \text{the acquisition of a memory trace}, \]

and

(T2) \[ \text{memory traces} = \text{what is learned}. \]

Consequently, if memory as a phenomenon consists in the retention of memory traces, as apparently maintained by the verbal learning tradition, then

(?) \[ \text{memory} = \text{the retention of what has been learned}. \]

Equation (?) offers a particularly simple view of the relation between learning and memory; and the three premises which have here led to it, namely (D1), (D2) and the assumption that memory is equivalent to memory-trace retention, illuminate
the force of the prima facie plausible proposition in that anyone who concurs with it is logically committed either to accept all three of these premises or to reject at least two of them. (However, the question-mark serves notice that the legitimacy of (?) will later be challenged.)

Equation (?) is one possible explication of the phenomenon-demonstrative meaning of “memory,” but what are memories? When previously considering possible verbal-learning applications of “memory” in the thing-kind sense, we noted that while memory traces are the most obvious candidates for this, another possibility is that memories appear only in the use, retrieval or recall of memory traces—except that words like “use” and “retrieval,” which imply that memories are devices which get transported from storage bin to showcase or workshop, are presumably metaphorical descriptions of some phenomenon whose literal nature remains to be clarified. We are now in a position to see in abstract generality what this phenomenon must be. It is evident that what is envisioned here is something which is made possible by retention of a memory trace, but whose actualization depends upon the circumstances of the passing moment—i.e., which consists in the arousal of some particular process property. Thus in general, to “use” a memory trace—which is a state property—is for this to make a difference for the organism’s process stages at a given moment, while what is “recalled” is some process property activated jointly by the trace and the immediately antecedent stages of the processes acting within and upon the organism. For example, under (D2) a conditioned tone-eyeblink reflex (an S-R “association”) is a memory trace; the reflex is “used” when it assists a particular presentation of tone in eliciting an eyeblink; and the elicited eyeblink itself is what is “recalled.” Similarly, in someone who has learned a BIQ→CEP association of nonsense syllables, the memory trace consists of that state condition in virtue of which this person is disposed to respond with CEP when presented with BIQ, while the recall brought about through “retrieval” of this trace is (roughly speaking—see below) the response CEP. The suggestion before us, then, is that eyeblink and CEP-saying, when aroused in this way, are memories—unless it is the memory traces themselves, here tone-eyeblink reflex and BIQ→CEP association, which are the memories.

Memory as Cognition

Quite a few lines of thought have been started in the preceding sections, but each was set aside before it could work its way to any firm conclusion. Frustrating as this may have seemed, it was an amassing of forces and we are now mobilized for the final push.

A moment ago, we arrived at the suggestion that an eyeblink or verbal CEP response, if elicited with the aid of a previously learned association, may be a
memory. Correlatively, eyeblinking or uttering CEP under these circumstances would be an act of remembering. On the other hand, if memories are the same as memory traces, simply retaining a tone-eyeblink or BIQ → CEP association would be remembering. Now both of these suggestions emerged by a perfectly straightforward extrapolation of the verbal learning tradition’s use of memory-words, yet by ordinary language intuitions, neither feels at all comfortable. To begin with, realization that in common-sense usage, remembering \( x \) is something that we fluctuate in our doing of from moment to moment (e.g., “I always have to think for a long time before I can remember John’s middle name”), added to our previous willingness to equate “memories” with “recollections,” makes clear that remembering should be assimilated to recalling rather than to retaining, and that memories are stages of process variables. But are there ever circumstances under which it would seem right to class overt actions such as eyeblinking and CEP-saying as memories? Not if everyday mentalistic psychology has any truth in it, for common sense is adamant that whatever else memories may be, they occur in the mind and not in behavior. This simple observation explains why behavior theory and conditioning research have had so little to say about memory: Where mentalistic concepts have been banned, there memory-words are also excluded. And it also helps to expose an important ambiguity—or should I say an anachronism?—in verbal learning concepts. For despite all its technical advances, verbal learning theory has never shaken free from the primary learning construct of classic mentalistic psychology, namely, associations between ideas. This is most apparent in the near-universal failure to distinguish between afferent and efferent verbal processes. What does a subject learn if he is shown repeated pairings of the stimuli \( TOP \) and \( BAT \)? Why, an association between the words \( TOP \) and \( BAT \) of course. But what sort of psychological process is a “word”? Would it be meaningful to represent the “association between \( TOP \) and \( BAT \)” by a symbol like

\[
TOP \leftrightarrow BAT
\]

with a bi-directional arrow?—i.e., if the subject has acquired both a \( TOP \rightarrow BAT \) association and a \( BAT \rightarrow TOP \) association, do “\( TOP \)” and “\( BAT \)” have the same referents in both of these formulas? Clearly not, if we were to discriminate between words as sensory patterns and words as motor emissions as any conditioning theorist would do. And the fact that verbal learning theorists seldom distinguish sensory words from motor words nor to my knowledge feel qualms about the possibility of bi-directional associations makes evident that the association is conceived

---

\( ^9 \)This is not at all adequately rendered by the distinction between stimulus-words and response-words which is, to be sure, commonly drawn in the verbal learning literature. In their more expansive moments, psychologists have seldom hesitated to apply the term “response” to any arousable process stage, even sensations, within the organism, and the verbal association \( TOP \rightarrow BAT \) can be called a “stimulus-response” association while intending no more than that arousal of the sensory pattern \( TOP \) tends to elicit the sensory image \( BAT \) in an organism possessing this association.
to hold between elements of the same kind, presumably the “ideas,” “images” or afferent representations of the previously paired stimuli. Thus for better or worse, verbal learning theory is covertly grounded in an S-S interpretation of human learning, and thus remains comfortably able to assimilate common-sense psychological terms which classically denote mental contents of the sort broadly conceived as “ideas”.

Realization that in common-sense psychology, memories are transient mental contents, casts considerable doubt on the tenability of our previous verbal learning hypothesis that memory as a phenomenon comprises all retentions of learned psychological states and that all psychological process stages aroused through the support of a learned psychological state property are memories. But then, this hypothesis was a projection into the verbal learning tradition of much more than has ever been explicitly asserted therein. Perhaps all that we can in fairness expect from the verbal learning literature is some crisp paradigmatic examples of memory and memories. Very well, then, let us return to our hypothetical subject who has experienced repeated pairings of the trigrams BIQ and CEP as part of a paired associates list, and assume that he has learned this list so flawlessly that whenever prompted with the cue BIQ he is invariably able to respond with CEP. Where in all of this do memories lie? Certainly not in CEP-saying, for we have already decided that a CEP-utterance does not qualify as a memory, but is only the peripheral aftermath of processes which occur within our subject when he sees BIQ; and in fact, despite our subject’s perfect mastery of the BIQ→CEP association, he will not emit an overt CEP-response when stimulated by BIQ unless he has been suitably primed to do so by instructions from the experimenter. Being the possessor of a BIQ→CEP association requires only being disposed to think—i.e., to have an aroused idea, image or sensory representation—of CEP when prompted by BIQ. The way in which this thought passes over into action, if at all, depends on additional details of the momentary circumstances. That is, the verbal association BIQ→CEP is a state variable which establishes a local relationship between two internal process variables, which may be called the “degree-of-thinking-BIQ” and “degree-of-thinking-CEP” variables, respectively, such that the stronger the degree of the BIQ→CEP association, the higher is the local correlation of degree-of-thinking-CEP with degree-of-thinking-BIQ. (For simplicity, we shall consider the degree-of-thinking-x variable to be only two-valued: thinking x vs. not thinking x.) Now, is the aroused process stage, thinking CEP (or, if the verb occasions discomfort here, the aroused CEP-thought), a memory? Surely not in general, as when it is elicited by sight of a stimulus card bearing the letters CEP, or by instructions to think of all pronounceable English-letter trigrams beginning with

---

10 That “ideas” and associations between them do exist is introspectively indubitable, and I see no reason why psychological theory need refuse to recognize this. Whether an unexamined perpetuation of classical concepts is the best we can do in this regard, however, is another question.
and ending with P. But is CEP-thinking (or the aroused CEP-thought) then sometimes a memory, as when elicited by a BIQ-experience in someone who has a BIQ→CEP association, and sometimes not? If this were so, we would have arrived at the strange conclusion that whether or not an idea (image, ingredient of experience) is a memory depends not on its content but on the circumstances of its arousal—which flies in the face of common-sense judgments that certain of our mental contents are memories even though we have no idea of how those memories came to be activated. Moreover, there is an even more powerful and direct reason why thinking CEP (or the aroused CEP-thought) could never legitimately be considered a memory under any conditions of elicitation. For if CEP-thinking (or the aroused CEP-thought) is a memory, what is it a memory of?

Before riding the wave of this last argument to a final crashing conclusion, let us quickly dispose of the other major way in which the learning, retention and “use” of a BIQ→CEP association might be considered to involve memory, namely, that the association is itself a memory. (This is what would be implied if we were to back away from our more recent conclusion that remembering = recalling and to reinstate our earlier hypothesis that remembering = retaining.) The most obvious common-sense objection to this proposal is that associations never appear in conscious awareness. A person can infer that he has an x→y association from introspective observation of the regularity with which arousal of x is followed in his experience by the nonperceptual arousal of y, but the enduring state property which is responsible for this correlation is not itself present in experience. Appreciation of this point is important for straightening out a serious confusion which frequently occurs in the verbal learning literature: If a verbal association is considered to be a memory trace, it is dangerously misleading to speak of “activating,” “retrieving” or “recalling” the trace, for this implies that what is aroused is the association itself. Actually, what is evoked in a person with, say, a BIQ→CEP association by cueing with BIQ is not the association but only the CEP-thought, just as presentation of tone to a subject with a tone-eyeblink reflex elicits not the reflex itself but the eyeblink response. Conversely, while the CEP-thought once present in conscious awareness can certainly be reactivated, reinstated or “retrieved” (I hesitate to say “recalled” here for the same reason that I am reluctant to describe re-elicitation of eye-blink as “recall”), it is not something which persists as a memory trace, either by itself or as an ingredient in the BIQ→CEP association, any more than retention of a tone-eyeblink reflex involves sustained overt or covert eyeblinking.

In short, once it is appreciated that associations are state properties while mental contents (thoughts, sensations, ideas, images or whatever) are process stages which associations help govern, the suggestion that BIQ→CEP and other learned associations should be classed as memories appears conspicuously at odds with common-sense psychology. This is not a conclusive objection—there might, after
all, be reasons why the thing-kind sense of “memory” in ordinary language logically subsumes associations and perhaps other learned state properties even though common-sense usage has only recognized its application to certain process stages. But here again the touchstone question arises: If the $BIQ \rightarrow CEP$ association is a memory, of what is it a memory?

The of-ness of memories is so fundamental that its role in the logical grammar of memory needs to be made unmistakably plain. We saw earlier that the root meaning of memory-words derives from the verb “remembering” and the thing-kind sense of “memory,” and that the sentence structures which these terms carry have the primary forms “$s$ remembers $x$” and “$s$ has a memory of $x$,” respectively. Formally, remembering is thus a relation between a person (organism, system) and something else. Just what sorts of entities can be remembered is left somewhat ambiguous in ordinary language, one unhelpful possibility being that in some cases, $s$ simply remembers memories (see fn. 12, below). However, remembering is obviously related in some intimate fashion to having a memory, and the grammar of the latter is obdurately insistent upon a tertium quid which is neither the memory nor the rememberer. There are many different memories a person might have, and the only method available in ordinary language for distinguishing one from another is to say what they are respectively memories of. That is, “memory” in the thing-kind sense is like terms such as “father,” “wife,” “handle,” “piece,” etc. which are inherently relational concepts, the relation becoming grammatically explicit when “of” is suffixed to the noun-stem. Thus no matter what the circumstances of its arousal or acquisition, a psychological process stage or state property cannot legitimately be classed as a memory unless it has whatever special features are required of the first element in a memory-of relationship.

And to what sort of thing is a memory related in this concept-definitive way? In broad outline (though not in all details) common-sense language is highly explicit about this. I have (or would have, were they to be elicited by suitable cues) memories of the party last night; of the birth of my first child; of the radio bulletins on Sunday evening, Dec. 7, 1941, announcing the bombing of Pearl Harbor; of how unjustly I was ticketed for a harmless traffic violation a few years ago—memories, that is, of various past events with which I once had experiential acquaintance. What a memory is of is some fact, or state of affairs, while the memory itself is, or at any rate includes, a re-enactment of the awareness which the rememberer had of that fact on the occasion of its occurrence. This was well known in classical psychology, as demonstrated by William James’ forceful statement:

Memory proper, or secondary memory as it might be styled, is the knowledge of a former state of mind after it has already once dropped from consciousness; or rather it is the knowledge of an event, or fact, of which meantime we have not been thinking, with the additional con-
sciousness that we have thought or experienced it before . . .

A general feeling of the past direction in time, then, a particular date conceived as lying along that direction, and defined by its name or phenomenal contents, an event imagined as located therein, and owned as part of my experience,—such are the elements of every act of memory. (James, 1890, p. 648ff.)

I am not sure that the common-sense requirements for a mental content to qualify as a memory are quite so stringent as James asserts, but boundary details are here a small matter; what is important is that a memory is above all a cognitive condition whose core is a belief (proposition, judgment) that such-and-such a fact is the case or that such-and-such an event took place. The reader should have little trouble persuading himself that this is, in fact, an essential ingredient in his ordinary usage of “memory.” It is interesting to note that this theme occasionally surfaces even in the technical literature. It is half visible when Gerard (above) speaks of “storing the past” in changed structures and asks whether trees remember the goodness or badness of past seasons by the thickness of their rings, or when the cybernetically colored phrase “information storage” is used to denote trace retention. Even more explicit are the italicized phrases (all italics added) in the following recent examples:

Memory of any particular event is dependent on a specific reorganization of neuronal associations (the engram) in a vast system of neurones widely spread over the cerebral cortex...We may say that the remembered thought appears in the mind as its specific spatio-temporal pattern is being replayed in the cortex. (Eccles, 1953, p. 266. Note how explicitly the second half of the quotation equates remembering with recall of a thought process rather than with retention of a state property.)

Superficially, memories of past occurrences seem to change in different ways. Sometimes the memory of a previous event seems to fade . . . (Riley, 1962, p. 402)

If we assume that the ‘learning to criterion’ of a visual discrimination is largely based on recent memory, i.e., that the animal remembers from trial to trial and from day to day, which figure is positive and which is negative, then this would explain . . . (Konorski, 1961, p. 128)

The systematic significance of studies of incidental learning lies primarily in the opportunity which they afford to identify the principles governing selective discrimination and retention of environmental events. (Postman, 1964, p. 146)

First, why can’t we recall everything that we experience? And second . . . why do we recall some things better than others . . . a recent
and vivid event is likelier to come to mind than a dull one from the distant past. (Waugh, 1963, p. 107)

In the paradigm cases, then, whatever else a memory may be, it is at least a cognitive representation of a state of affairs once perceived directly by the rememberer, accompanied, if James is correct, by awareness that it was so previously experienced. Just what sort of psychological condition a “cognitive representation” may be is an exceedingly difficult question which lies beyond the present scope; as a first approximation we may consider it to be some sensory-like (ideational) process similar to the original percept, though fuller analysis would reveal this to be a seriously inadequate characterization. How similar the memory must be to the original perceptual awareness of the remembered event and how much accuracy it must preserve in order to merit the honorific title “memory” are wide open concept-boundary questions which, however, do not seem to me to have any great significance for technical psychology. It is of psychological importance, though, that memories have the moment-to-moment temporal instability characteristic of process stages. With one special exception, memories need to be recalled—i.e. the appropriate process variable must be returned to something like its earlier aroused value following a period during which it has subsided to an “off” level. The exception is the case of perceptual (or ideational) hysteresis in which awareness of an event has been elicited by sensory input from the event itself and lingers after this stimulation is discontinued. Common-sense usage is not altogether firm about whether such perseverations should count as memory or not; however, classical psychology recognized this phenomenon as “primary memory” (see James, 1890, p. 643 ff), and the intrinsic difference between a perseverating initial awareness on the one hand and a revived awareness on the other seems hardly great enough to warrant conferring memory-status upon the latter while withholding it from the former.

Insomuch as memories are (in general) recalled, it follows that memory phenomena must involve two importantly different kinds of psychological attributes. First of all we have the memory proper, which is a process stage which comes and goes according to the passing stimulus circumstances. But there must also be retention of a state condition, acquired through perception of the to-be-remembered event, in virtue of which it becomes subsequently possible for suitable cues to evoke the memory. This underlying state condition is sometimes called a “structural” or “latent” memory—which, however, is a dangerous locution which suggests that memory-potentiating states are basically like memories. This is no more appropriate than would be describing a reflex as a “structural response” or “latent response.” While memories proper, like responses, are not learned but elicited, acquisition of the state property which makes the memory available is a straightforward instance of learning—in fact, of one-trial learning insomuch as it is logically
impossible for a to-be-remembered event to be repeated.\textsuperscript{11}

And now, what about “remembering”? Earlier in this paper it was suggested that “s remembers $x$” is equivalent to “s has a memory of $x$,” but actually this is not quite correct: For any expression “$x$” such that “memory of $x$” is a grammatically acceptable phrase, “remembering $x$” and “having a memory of $x$” do indeed appear to be essentially synonymous (e.g., “John remembers the death of his grandmother” vs. “John has a memory of the death of his grandmother”). However, as shown by comparing, e.g., “John remembers \textit{that the dam broke last year},” “John remembers \textit{how to ice skate},” and “John remembered \textit{to mail the letter},” respectively, with “John has a memory of \textit{that the dam broke last year},” “John has a memory of \textit{how to ice skate},” and “John had a memory of \textit{to mail the letter},” a grammatically correct instantiation of “$x$” in “remembering $x$” is not always acceptable in “having a memory of $x$.” Careful analysis of these differences in usage could easily fill a separate article, but with my usual temerity I will try to cover the ground in a paragraph or two.

There appear to be three primary grammatical forms of expressions which can instantiate “$x$” in “remembering $x$”: (1) \textit{how-to clauses} in “s remembers how to such-and-such”; (2) \textit{that-clauses} in “s remembers that so-and-so”; and (3) \textit{event-descriptions} which are equally acceptable in “remembering $x$” and “having a memory of $x$.” Other variants, such as remembering to do something, may be analyzed as ellipses for or derivative from one or another of these three forms. Grammatically, forms (2) and (3) are extensively interchangeable in that most that-clauses can be paraphrased into event-descriptions, while conversely, all event-descriptions can either be paraphrased directly into a that-clause or regarded as elliptical for a more complete event-description which can be so paraphrased. Thus “John remembers \textit{that the dam broke last year}” is equivalent both to “John remembers \textit{the dam’s breaking last year}” and to “John has a memory of \textit{the dam’s breaking last year}”; while in, say, “John remembers \textit{his fifth birthday}” (which is equivalent to “John has a memory of \textit{his fifth birthday}”), the phrase “his fifth birthday” is incomplete in that it doesn’t tell what John remembers about his fifth birthday, and making this explicit brings out either a that-clause (e.g. “that one of his presents on his fifth birthday was a green sack of marbles”) or an equivalent event-description (“receiving a green sack of marbles for a present on his fifth birthday”). Rememberings of form (3) and many of form (2) thus consist in having a memory of some past event. To be sure, some that-clause rememberings are not comfortably reducible to having memories. For example, I can remember that 7 times 8 equals 56, that blue and yellow are complementary colors, and that George Washington was the first President of the U.S., but it does not seem quite

\textsuperscript{11}The prevalent fallacy that events can be repeated is due to confusion between events proper, which are designated by expressions with the logical structure of a sentence or ellipses thereof, and \textit{event-types} which are designated by predicates. (See Carnap, 1950, p. 35.)
right to say that I have a memory of 7 times 8 equaling 56, of blue and yellow being complementary colors, or of George Washington’s being the first President. The event-paraphrases of these that-clauses comes off satisfactorily, but whether timeless generalities and past occurrences which I have not personally witnessed are things of which I call have memories is more problematic. But again, this is only a concept-boundary wobble. Whatever else may also be required, remembering (recalling) that \( p \) is the case involves an aroused state of belief in \( p \), where this belief is a revival (or possibly, in the case of “primary” memory, a continuation) of knowledge achieved previously. Hence as psychological phenomena, there would appear to be no important difference between rememberings of forms (2) and (3) on the one hand, and having memories on the other.\(^{12}\)

On the other hand, how-to rememberings constitute a significant deviation from the pattern of memory-word meanings that we have isolated so far. It should require no argument to convince the reader that “\( s \) remembers how to \( d \),” in which “\( d \)” describes some skillful activity—e.g., row a boat, do long division, typewrite, disagree with a superior without making him angry, etc.—is equivalent to “\( s \) retains the ability to \( d \).” Thus ordinary language undeniably considers some cases of retention to be instances of remembering, namely, the retention of certain abilities. (Not all ability-retentions count as rememberings, though—a person would not be said, e.g., to remember how to read newsprint at a distance of 20 feet or how to bend a crowbar with his bare hands, even though he can do these things, unless there is some special trick by which the feat is accomplished. The difference between abilities which ordinary language allows to be “remembered” and those which are not so honored appears to depend on the manner in which their acquisitions involve learning, and perhaps the complexity of what it is that is learned.) And now that we have admitted that remembering can sometimes be retaining, it might as well also be confessed that we sometimes say that \( s \) remembers that \( p \) when we mean only that \( s \) can summon awareness of \( p \) when needed—i.e., that \( s \) retains a state property in virtue of which awareness of \( p \) is aroused by appropriate cues. (This parallels our custom of extending the notion of “knowing that \( p \)” to include the disposition to have active knowledge of \( p \) on relevant occasions.) Patently not all retentions of learned states qualify as rememberings, however. One does not remember a tone-eye blink reflex, for example, nor how to blink an eye. (Either how to eyeblink falls short of the complexity required for a potential action to characterize an ability, or it is not the sort of ability which ordinary language considers to be rememberable.) Neither is retaining a learned \( BIQ \rightarrow CEP \) association remembering, though the fact that \( BIQ \) and \( CEP \) previously occurred

\(^{12}\)When \( s \) remembers that \( p \), it is a moot question whether what \( s \) remembers is the fact \( p \) itself, or a memory (or some other psychological process stage) which is about \( p \). But while this raises some important issues in the philosophy of intentionality, it is ultimately a matter of definition and has negligible psychological import.
together can certainly be remembered.

Actually, the common-sense boundary between states which are rememberable and those which are not can be drawn with remarkable precision, relative to the concept of “knowledge.” Some years ago, Ryle (1949) pointed out the now-celebrated distinction between knowing how and knowing that. It appears that the verb “knowing” has two major usages, one occurring in sentences of form “s knows how to d,” in which d is some skillful activity, and the other in sentences of form “s knows that p,” where p is a state of affairs. These are readily seen to correspond exactly to the two primary senses of “remembering.” Necessary conditions for s to remember that p and how to d are for s to know that p and how to d, respectively; while conversely, s can know that p and how to d only if “p” and “d” are expressions such that it is logically possible for s to remember that p and how to d. To be sure, the non-cognitive “knowing how” of abilities is tantamount to a contradiction in terms, but that is the fault of everyday linguistic conventions; what is relevant for analysis of memory-words is that the expression “remembering how to d” is commonsensically acceptable for certain activities d only because “knowing how to d” passes muster. Hence even in its more liberal verb-form extensions, the ordinary-language concept of “memory” is strictly confined to ways of knowing.

Conclusions and Implications

Our analysis has shown that the common-sense concept of “memory” is considerably more intricate and specific than has been recognized by technical psychology of the modern era. Whereas contemporary research traditions have for the most part allowed themselves to speak seriously of “memory” only in the noncommittal phenomenon-demonstrative sense when they have not avoided the use of memory-words altogether, ordinary language has at least five major forms of memory-words, of which “remembering” and thing-kind “memories” are most basic. Memories, we have seen, are a subclass of aroused beliefs about past events, while remembering is first and foremost recalling a memory, though the verb also has three significant extensions beyond this, namely, (a) to arousal of awareness of any fact, not merely dated events in past experience, which the rememberer may be said to have “known” previously, (b) to the perseverating awareness of an event immediately following discontinuation of perceptual knowledge thereof, and (c) to retention of

It is instructive—psychologically as well as philosophically—to speculate on why ordinary language considers certain abilities to be a form of “knowledge.” Part of the answer may be that such abilities seem as though they should be describable as manifestations of well-mastered aggregates of cognitive knowledge (e.g., a typist knowing what letter is printed when the key at such-and-such a position is struck), but possibly more fundamental is that “knowable” abilities involve an integration of component processes in a structural complexity as great as that found in cognitive knowledge.
certain abilities and other dispositional state properties which ordinary language regards as “knowledge” in an extended (I would prefer to say “corrupted”) sense of the word.

It follows that except for some outgrowths of classical mentalistic psychology, notably the work of Bartlett (1932), virtually none of the modern research on “memory” and the retention of learning has been concerned with memory proper at all. The basis for this conclusion is very simple: Except for the special cases noted, remembering is not retaining but recalling, while moreover, those processes to whose arousal an organism is disposed in virtue of the state properties which have been studied or hypothesized by learning theorists are altogether devoid of the cognitive properties needed for them to count as memories. For example, even though John Smith is now saying or thinking the nonsense syllable CEP because he has previously acquired a BIQ→CEP association and was presented with the cue BIQ a moment ago, neither his CEP-utterance nor his CEP-idea is a memory, for there is no past event that the thought, much less the overt response, is a memory of. This holds true even for associations between cognitively meaningful words. If Smith is now thinking the word ZEBRA because he has had paired-associates training on FOX-ZEBRA and was just stimulated with FOX, his thought ZEBRA may be of zebras, but there is still no past event, involving zebras or not, of which this one-word thought is a memory. To make this important point as strongly as possible, consider an everyday-life type of association. Suppose that when song S was popular, Smith was having an affair with Mary Doe, so that now whenever Smith hears S he feels a complex bittersweet nostalgia and has a vivid image of Mary Doe’s face. No matter how similar this evoked emotion may be to what Smith was experiencing during the affair, it is still not a memory, for emotions have no referential properties. Neither is the image of Mary Doe’s face in itself a memory, for it is first of all highly problematic whether mere sensory resemblance to a percept suffices to endow an “image” with a referent (to say, as is customary, that the image is of Doe’s face actually begs some extremely important issues), while even if the image genuinely refers to Mary Doe, it lacks the propositional content required for it to be a belief. Only if the emotion or visual image is accompanied by some awareness that this is something like what I [Smith] felt then, or that this is the way Mary looked, or the like, can it properly be said that song S evokes any memories in Smith. Likewise, it is indeed possible for stimulus BIQ or FOX to elicit memories in which thought of CEP or ZEBRA is an ingredient—e.g., recollection that BIQ (FOX) occurred together with CEP (ZEBRA) several times

14That is, this is a possibility. Actually, there are some nice problems raised here about the semantic properties of isolated words in nonlinguistic contexts.

15Actually, the recalled image of Mary Doe’s face lies close to the (indistinct) borderline between memory and non-memory. While a mere pictorial reproduction of an object does not logically constitute a proposition about that or any other entity, it nonetheless has a structural organization which is very similar to that of a full-blooded assertion about that object’s appearance.
recently. However, the extent to which that-p processes accompany performance on retention tests has to my knowledge never been investigated, either empirically or conceptually, in verbal learning research.

This last point, incidentally, undermines one argument that verbal research has indeed been concerned with memory proper. The instructions given to subjects on test trials of a paired-associates or serial learning task are likely to stipulate that the subject remember what words were paired, or what word came next, on the previous learning trials. Likewise, in a recognition or reproduction procedure it is virtually impossible to give instructions which do not explicitly request the subject to recall what he experienced previously. Just because the subject is told to respond on the basis of cognitive knowledge of his past, however, it is not guaranteed or even necessarily made likely that such fancy processes actually mediate between test stimuli and test behavior, nor have verbal learning theories postulated that they do so.

If verbal learning research, which has at least tried to talk about memory, has failed to come to grips with the real thing, it is hardly to be expected that studies of animal learning would have come any closer. Obviously no conditioned responses are memories, but more importantly, none of the theoretical processes which behavior theorists have postulated to mediate overt responding include knowledge of the past, either. Occasionally one still encounters full-dress cognitive interpretations of conditioning, as in Konorski's suggestion, quoted above, that in visual discrimination the organism has knowledge about the environment, but these are vestiges of an earlier era when we had not learned how to discuss animal behavior without presupposing a traditional mentalistic explanation. It is true that one brand of behavioristics, Tolmanian expectancy theory, has experimented with quasi-cognitive mechanisms, but even if an S-S or S-R-S “expectancy” were genuinely a belief that this will be followed with that or that doing so-and-so in these circumstances will lead to such-and-such—and if it were sufficiently germane, I would argue that a belief-construal of behavior-theoretical expectancies goes to gratuitous excesses beyond what the formal structure of the theory requires—the belief would still be an anticipation of the future rather than a memory-belief about the past.

That memory has not, in fact, been studied by any branch of learning research in no way demeans the past accomplishments or present merit of this research. Quite the opposite—it would have been intellectual disaster had not the complexities of cognitive phenomena, which even today are too slippery to be grasped firmly, been set aside in order that we might first gain some technical understanding of the more elemental behavior mechanisms. In fact, perhaps the most significant achievement of conditioning research and behavior theory has been to show, however imperfectly, that a great deal of psychological functioning can be accounted
for by underlying states and processes much more primitive than those envisioned in classical psychology. (Contrary to the illusions of some anti-behaviorists, this in no way denies the generic existence of any “higher mental process”; rather, as Hebb (1960) has emphasized, it makes it possible for us now to tease out what is envisioned in these classical conceptions which is more than just, e.g., a constellation of conditioned reflexes described in humanistic terms, and challenges us to develop observable criteria for these extras.) Even so, heightened sensitivity to the ordinary-language commitments of memory-words leads to two important conclusions about the present status and future development of psychological theory, one concerning the conceptual sophistication of verbal learning research and the other having to do with the study of cognitive processes.

In my judgment, on the most unsightly blotches on the fair if adolescently acned face of contemporary psychology is the appalling abuse that cognition-words have received, perhaps the worst offenses having been committed against “concept, “information,” and “cognitive” itself, as well as, of course, “memory.” It is not just the frequent metaphorical intrusions of these terms into contexts where they do not apply, though this is reason enough for dismay. More inexcusable is that in studies expressly concerned with cognition, cognition-word labels have been indiscriminately applied to any psychological state or process lying in the general direction of the ordinary-language usage, while the genuinely cognitive aspects, if any, of the phenomena under study go unheeded. It particular, it seems to be felt that any mentalistic or quasi-mentalistic condition of the organism is “cognitive,” so that virtually any bump on the S-R hyphen becomes a “cognitive” process. (Another prevalent use of “cognitive,” namely, to distinguish holistic, phenomenalistic and existentialistic psychologies from behavioristic approaches, is so unspeakable that I shall not speak of it.) As strongly as I feel about this particular matter, it is impossible for me to give it any responsible attention in the waning paragraphs of the present essay, especially insomuch as recent trends are complex and by no means wholly bleak.16 Pending fuller discussion on another occasion, I shall merely indicate the general nature of my discontent.

By definition, the essence of cognition is knowing, and if we are not too squeamish about philosophical niceties, “knowledge” may be defined roughly as true belief, “true” here meaning truthful (veridical, accurate) rather than staunch, noble, faithful to doctrine, or other non-cognitive corruptions of this term. Whether a given belief is in fact true or false makes little difference for its nature as a psycho-

---

16I am thinking particularly here of works by Miller, Galanter and Pribram (1960) and George (1962), both of which in my judgment have struggled heroically to move in the right direction, the former in its concern for structural organization and latter in its attempt to extract cognition proper from a broader background of psychological phenomena. The past decade of work on linguistic structure and verbal “meaning” has also plowed a receptive field in which, with a little philosophical fertilization, a thriving psychology of semantics could germinate.
logical process, but unless an investigation is concerned one way or another with what differentiates beliefs from other psychological conditions (e.g., eyeblinking or CEP-thinking) to which ascriptions of truth or falsity make no sense it is not entitled to the status of “cognition research.” Now, for a psychological entity to be true, it must correctly signify a fact—i.e., a belief is about some state of affairs, and is true or false according to whether or not reality is as the belief represents. Moreover, the factual reference or truth-conditions of a belief are somehow determined by the referents (designata) of its descriptive elements together with the propositional structure by which the latter are concatenated in the belief. Thus the focal (though by no means the only) problem of cognition is the nature of reference, or “aboutness.” I propose that a piece of research or theory can properly be said to concern cognition only if it deals with some aspect of whatever it is that is distinctive about those special psychological states or processes which refer to, designate or are about some other entity. For example, concepts are inherently concepts of (about) something. Yet virtually all research on the formation and utilization of concepts has concerned the acquisition of categorizing habits (e.g., learning to discriminate the relevant cues associated with differential reinforcement in verbal labeling or object sorting) or the subsequent transfer-effects of such training. These “concept acquisition” experiments are straightforward instances of discriminating learning, revealing little if anything more than something about what stimuli their subjects can respond to, while the mediational responses postulated by some of the theory in this area have been ascribed no behavioral properties (other than observational inaccessibility) which are significantly different from those of any traditional response process. Contemporary attention to “concepts” has thus brought out nothing which cannot be comfortably handled in non-cognitive terms, and in fact the issue of conceptual reference scarcely ever appears in this literature beyond an occasional hint that what a concept refers to, or denotes, is the class of stimuli which elicit it. (And is, then, an eyeblink response also a “concept” when it has been conditioned to a variety of stimuli?) The only possible conclusion is that for the most part, concept formation, as something more than mere generalization and discrimination, is still virginally awaiting the first bold probe of serious research. Admittedly, the record is somewhat better than this in other branches of cognitive psychology—Zener and Gaffron (1962), for example, differentiate incisively between cognitive and noncognitive interpretations of perception, while Osgood’s work on psycholinguistics has made explicit provision for the the sign/significate relationship—but how typical it is for cognition theorists to discriminate knowledge from non-knowledge conditions, or to

17Scheerer (1954) makes a similar point when he emphasizes the “representational” aspects of cognition, but loses his grip on it when he assents to such global claims as “From the strictly cognitive point of view, memoria are also cognitive structures in that they have mediating and orienting characteristics. . . . This holds not only for knowledge but for habits and skills, as well as the maturing ego-apparatus” (Reiff & Scheerer, 1959, p. 36, italics added).
recognize the existence of an of-ness problem I shall leave to the reader’s own judgment.

It is far from my intent to imply that this sweeping neglect in contemporary psychology of the definitive problems of cognition is just a silly oversight which, now that it has been pointed out, can easily be set aright. As generations of philosophers have convincingly demonstrated, reference and truth are exceedingly treacherous matters to think about even confusedly, much less to pin down with tough-minded precision. About the only serious penetration into the mechanics of aboutness has occurred in philosophical semantics, and this is still very much in a flux of development with its classical foundations now tottering (see Rozeboom, 1962, p. 332ff). But this is all the more reason why psychologists must stop evading their rightful duties. It is very likely that epistemological theory has gone about as far as it can go until it receives incursions of factual knowledge about the fine structure of internal behavior mechanisms, and this knowledge will become available only when research psychologists begin to think seriously about the “philosophical” aspects of cognition. Undoubtedly the distinction between what is cognitive and what is not spreads over a multidimensional continuum, but the axes of this continuum will never be identified until we begin to discriminate among its various regions. What most urgently needs to be done now is to contrast psychological states or processes which by common-sense standards are paradigmatically cognitive with others which just as intuitively are not, in order that we may begin to decipher wherein, behavioristically, the difference lies.

Finally, let us take a look at current research on “memory” in the verbal learning tradition in light of present conclusions. We saw above that verbal associations, whether construed as S-R bonds or as ideational linkages, are not memories. But even more significantly, they do not involve cognition in any fashion—nowhere, nohow. Consider once again our learned $BIQ \rightarrow CEP$ association. Is there any aspect of the retention and use of this wherein knowledge, belief or aboutness of any sort enters? Certainly not in its “use,” for this is merely to provide a means by which occurrence of the idea (or percept, or sensation, or stimulus) $BIQ$ elicits the thought (or image, or idea, or response) $CEP$, and the $CEP$-thought does not prima facie even have a referent, much less contain a belief about anything. It would, of course, be possible for presentation of $BIQ$ to arouse in our subject a belief, say, that $CEP$ is about to occur, but this is quite different from the simple $CEP$-idea, nor is the state property in virtue of which stimulus $BIQ$ could elicit this belief be the same as a simple $BIQ \rightarrow CEP$ association. But could the $BIQ \rightarrow CEP$ association itself be or contain a belief? This is implausible on the face of it because introspectively, beliefs are processes which well up in and ebb

\footnote{For example, the earnest efforts of Morris (1946) and more recently Quine (1960) to work out the cognitive relations between language and reality both founder on the inadequacies of the primitive behavior-theoretical constructs at their disposal.}
from conscious experience with the passing circumstances, whereas associations
are state properties which are never themselves aroused, consciously or otherwise,
but only serve to make something else arousable. Even so, it would be ingenuous
in this post-Freudian era to assume that all a person’s beliefs are consciously ex-
perienced, and there remains the possibility that some enduring attributes of the
organism might be essentially propositional in nature. If cognitive state properties
exist, however, verbal associations are not among them. For were the \( BIQ \rightarrow CEP \)
(or any other) association to be or to contain a belief, what would this belief be about—i.e., what would verify or refute it? Merely to raise this apparently nonsensical question, whether or not a \( BIQ \rightarrow CEP \) association is true (veridical, correct, accurate), suffices to expose the association’s lack of propositional content, but let us press a little deeper. If the association were cognitively related to any state of affairs, this would surely have to be some fact concerning the objective co-occurrence of trigrams \( BIQ \) and \( CEP \). Following through on this suggestion while recalling also that the \( BIQ \rightarrow CEP \) state should more precisely be qualified with a measure of its strength, we might wish to consider whether an organism’s \( BIQ \rightarrow CEP \)-strength could possibly be construed as an opinion about the objective contingency of \( CEP \) upon \( BIQ \). But if so, what would constitute a match or mismatch between it and reality? For example, if the association were at maximal strength, would this be false (mistaken, erroneous, inaccurate) if the organism had ever been presented with \( BIQ \) without an accompanying presentation of \( CEP \), or if presentation of \( CEP \) has always accompanied presentation of \( BIQ \) in the organism’s history to date but will cease doing so in the future? Does a zero-strength \( BIQ \rightarrow CEP \) association have the force of a surmise that no objective relationship exists between occurrences of \( BIQ \) and \( CEP \)—and if so, is there then any way in which an organism can avoid having a cognitive hypothesis about the objective contingency of \( CEP \) upon \( BIQ \)? Moreover, have not these last speculations presupposed that the \( CEP \)-thought to which an organism is disposed in virtue of a \( BIQ \rightarrow CEP \) association refers to (designates, signifies, denotes, is about) the objective trigram \( CEP \)? If so, what grounds have we for this assumption? (Does any internal process elicited by a stimulus \( S \) or resembling one elicited by \( S \) then refer to \( S \)?) If not, why should not, say, a tone-eyeblink association also be suspected of being a cognitive representation of some state of affairs—or is it, and if so, of what? Hopefully, this barrage of questions will help not merely to expose the alleged cognitive status of verbal associations as the imposture it is, but also to suggest some of the problems that are brushed aside when cognition-words are applied indiscriminately, as when under the influence of cybernetics jargon retention of state properties whose operating characteristics are no more complex than those of a simple association is called “information storage.”

And why should we become exercised about the lack of attention given by learning research to memory or other forms of cognition when thorough understanding
of the phenomena which have in fact been studied is a prerequisite to any head-
way we may eventually make on cognition proper? Because when we discuss basic
mechanisms in terms of inappropriate cognitive metaphors, we not merely con-
taminate our ability to think straight about cognition, but also blind ourselves to
gaps and errors in our understanding of the subcognitive domain. Consider, for
example, the short-term and long-term “memory” phenomena currently promi-
nent in verbal learning research, in which subjects are first exposed to a simple
or complex stimulus and later given opportunity to make a response differentially
affected by this prior stimulation. To describe what happens in such experiments
as simply the storage and retrieval of memory traces (or of “information”) only
conceals under a glaze of common-sense pseudo-familiarity our inadequate tech-
nical knowledge of what, if anything, the subject learns from the first stimulus
exposure and how this later operates upon his retention-trial behavior, as well
as smearing together phenomena which are critically distinct in formal structure.
In particular, the storing-and-retrieving metaphor, which envisions some particu-
late entity being shuttled back and forth among compartments of the mind, fails
to distinguish between process stages and state properties, thereby allowing it to
seem perfectly plausible that short-term and long-term memory phenomena differ
only in their positions along a continuum (Melton, 1963). This point is sufficiently
important that I feel obliged to expand upon it, if only briefly and much more
crudely than is appropriate to the level of research sophistication that has been
achieved in this area.

Suppose that yesterday we gave a subject several serial anticipation trials on
a list of trigrams, in which the fourth and fifth items were MUH and ZIJ, respec-
tively, and today, when we ask him what follows MUH, he answers ZIJ. The tra-
ditional and simplest explanation for this behavior is that our subject has learned
and retained a MUH → ZIJ association, in virtue of which today’s presentation of
MUH evoked a ZIJ-idea (or image, or covert motor response, or possibly some
other sort of internal process stage) which, in some fashion for which the tradi-
tional theory has no explanation, was converted into an overt ZIJ-response by
our having instructed the subject to speak up. (If we had asked our subject to
tell us what syllable occurred on the list two steps before MUH, on the other
hand, we could no longer explain his ability to answer correctly—if he has it—
on the basis of simple associations alone, inasmuch as a once-removed backward
association should be overwhelmingly dominated by the immediate forward asso-
ciation. However, this latter phenomenon is not our immediate concern.) What
is “stored”—i.e., retained—according to this explanation is not a cognitive repre-
sentation of yesterday’s events nor even an idea (image, response) of ZIJ, but the
association MUH → ZIJ; whereas what is “retrieved”—i.e., aroused again—on to-
day’s retention trial is not the MUH → ZIJ association but only the ZIJ-idea. Thus
the storage-and-retrieval model here confounds ZIJ-thinking, which is a process
stage, with the state property of a $MUH \rightarrow ZIJ$ association.

On the other hand, suppose that in a short-term memory experiment, we show a naive subject the trigram $ZIJ$ for a moment, wait a bit, and then ask him to repeat what he saw. Here the subject’s correct response is most simply explained as perseveration of the $ZIJ$-idea which, as in the preceding case, is in some fashion made manifest in behavior by our instructing the subject to tell us what is on his mind. There is here no formation and storage of any association (or at least none is called for by the explanation), but only a process in hysteretic decay, nor is there any “retrieval” except in the sense of arranging for the internal ideational process to trigger an overt response.

Finally, suppose that in a free-recall experiment we showed our subject the $ZIJ$ stimulus by itself a day ago, and today ask him to reproduce the nonsense syllable he saw yesterday. If he is successful, our interpretation is more problematic than in the previous cases. Operationally, this procedure is most like a short-term memory design with, however, a retention interval on the order of long-term effects. A short-term memory explanation to the effect that the ideational process aroused in our subject 24 hours ago by stimulus $ZIJ$ has not yet faded from experimenter-interrogational access, might well be considered less plausible than the hypothesis that our present instructions to remember yesterday’s trigram interacted with some state property acquired yesterday to arouse the $ZIJ$-idea anew—except that here verbal learning theory is rather at a loss to say what that state (an association? If so, of what with what?) may be.\footnote{The free-recall experiment and its attendant interpretive difficulties for association theory have recently been emphasized by Asch (1964), though it may be questioned whether Asch’s presentation does full justice to the capabilities of association theory to handle such data.}

The distinctions addressed in the last three paragraphs are obscured not merely by the storage-and-retrieval metaphor but also by the uncritical use of “memory” in the phrases “short-term-memory” and “long-term memory.” These imply that what is being retained is the same sort of thing, namely, “memory traces,” and that what primarily differentiates the two phenomena is the length of retention. But in fact, the retention-time difference is only secondary—what distinguishes short-term memory (STM) from long-term memory (LTM) is the logical structure of their theoretical explanations, which in turn ideally (though not always so cleanly in practice) follow with inductive immediacy from the operational structures of the experimental procedures. In both cases, we put the subject into an instruction-aroused condition under which his overt behavior is presumably symptomatic of a certain internal process variable $X$ and try to discover how a certain internal condition alters as a function of time-since-acquisition, but in an STM design our inferred dependent variable is $X$ itself whereas in an LTM design it is the local dependence of $X$ upon some cue-variable $S$. That is, STM studies decay of $X$,
a process change, while LTM studies decay of X’s responsiveness to S, which is a state change. That different mechanisms are involved in STM and LTM has, of course, been widely accepted. However, the distinction is usually drawn in physiological terms, as in Hebb’s (1961, p. 41) contrast between “activity traces” and “structural traces.” What I want to emphasize here is that STM and LTM phenomena differ importantly in formal structure, not just in their physiological substrata, and that to classify both as “memory,” with the implication that they manifest the same empirical pattern of events differing merely in time-scale, is to obscure their most significant behavioral differences.

To repeat, our propensity for thinking in loose metaphors when we should be trying for carefully wrought sentences that say precisely what we want them to say too often fogs our recognition of the very technical details that are most seminal for further research and theory. I have complained enough about language abuses in this paper; let me conclude with three illustrations of the sorts of theoretical questions about which we should be brooding.

Shepard and Teghtsoonian (1961) have shown that if a subject is exposed to a long sequence of 3-digit numerals, some of which occur more than once, and after each exposure the subject guesses whether or not he saw that numeral previously, the probability of a correct response to a previously shown numeral, expressed as a function of the number of intervening presentations of other numerals and compared to suitable control rates, reveals a decaying effect of each specific stimulus which remains detectable even when accompanied by the residues of many dozens of more recent stimuli. This experiment is an efficient modernization of the old-fashioned “recognition” procedure, and it raises a nice little theoretical problem: What’s going on here, anyway? That is, what sort of psychological condition resulting from stimulation by a particular numeral is being retained (or more precisely, is gradually fading), and how does it interact with the stimulus input on the recognition trial to produce the identification response? It is hard to argue that the retention here is of an association, for what association could be learned on the first presentation of numeral N (which at that time most frequently elicits a denial-response) in virtue of which the later repetition of N gives rise to an affirmation-response? (I can think of possibilities, but they are labored.) The most likely bet is that what is retained is a decaying “trace,” in some sense, just of the stimulus N itself, akin to an afterimage but more centrally located. This is the sort of nonperceptual process that is envisioned by the classical concept of “idea,” except that in the present experiment the “trace” or “idea” of N is presumably long gone from conscious awareness by the time of the recognition trial. And since we can thus not appeal to common-sense intuitions about introspective reports for an explanation of the Shepard-Teghtsoonian data, we are left with the problem of how the second stimulation by N combines with the residual trace of the first exposure to produce an affirmation-response. Might this not be a summation
phenomenon in which stimuli receiving facilitation from a matching trace-residue have a better chance of exceeding the affirmation-response threshold than stimuli which receive no such assistance? Quite possibly—but more significant is that none of our traditional mentalistic or associational concepts seem quite appropriate here, and we had best be painstakingly explicit about the formal properties of the theoretical constructs we develop to account for data such as these, especially insomuch as the relative permanence of some recognition-dispositions (see Rock & Ceraso, 1964, p. 115) indicate that complex stimuli may deposit trace states in their perceiver over and above the transient trace processes which presumably account for STM phenomena.

Next consider a simple little experiment reported by Peterson (1963), in which subjects were given sometimes one, sometimes two exposures to paired-associates stimuli comprising word-digit couples, and were then tested after 0, 2 or 8 seconds (corresponding to different numbers of intervening exposures to other stimuli) for recall of the numeral in response to the word-cue. It was found (a) that probability of a correct response decreases sharply with increasing time between test trial and last presentation of the paired-associates stimulus, and (b) that the parametric details of this decay function are themselves a function of the number of previous learning trials on that pair. Taken at face value (and I have overstated the empirical solidity of the data as well as oversimplifying my analysis thereof in order to make a methodological point), these results implicate the existence of two different underlying variables, one which is aroused by presentation of the paired-associates stimulus and decays rapidly after this stimulation is discontinued—let us call this the “stimulus trace”—while the other is a function of the number of stimulus presentations and influences how rapidly the stimulus trace decays after arousal. How should these two theoretical variables be characterized in traditional terms? The stimulus-trace variable behaves something like an “association” in that it enables the subject to emit a response corresponding to the second component in a previously experienced compound stimulus when the first component is presented as a cue. The dynamics of this variable are quite different from those of a classical association, however, and it seems altogether reasonable to think that it is more like a classical “idea” of the compound paired-associates stimulus. (Note that even in traditional terms an association $x \rightarrow y$ is fundamentally different from a percept, image or idea in which mental elements $x$ and $y$ occur jointly.) Further, the second inferred variable acts like a state variable governing the compound-stimulus trace process, and grows in strength with repeated stimulus presentations just as a learned association is traditionally supposed to do. But is this second variable,

---

20 The “experience of recognition” which Rock and Ceraso (1964) find so inexplicable on association-theoretical grounds (a judgment with which I agree) may well be little more than an increased perceptual vividness or, more likely, something akin to this on a more central ideational level, due to matching-trace enhancement.
then, a word-numeral association? Perhaps—but if so, we have discovered a new way in which an association may affect processes for here its prima facie function is not to help the word cue elicit the numeral response but to retard decay of the compound stimulus trace.

Most important of all, this casual phrase, “compound stimulus trace,” brushes aside behavior-theoretical problems so profound that it is difficult even to make intelligible what they are, though the terms “structure” and “organization” point in their direction. A choice example of what I have in mind is contained in an immediate-memory phenomenon recently discovered by Hebb (1961) and explored in greater detail by Melton (1963). The procedure consists of reading (or, in Melton’s version, presenting visually) to the subject a succession of 9-digit numerals, after each of which the subject tries to reproduce the numeral just received. Since nine chunks is at the upper limit of memory span, the subject’s accuracy of reproduction is ordinarily substantially less than perfect. However, it is found that if a given numeral duplicates one which occurred earlier in the series, the accuracy of its immediate reproduction is greater than what it would be had not this numeral been experienced previously, the improvement being an increasing function of both the number of previous exposures to that numeral and their temporal approximation to the immediate-reproduction trial. Like the Shepard-Teghtsoonian recognition data, these results can most readily be explained as a summation effect in which a process aroused by present stimulation is enhanced by a trace of that same process still lingering from a previous arousal. But what sort of process? All the stimuli in the series are composed of elements drawn from the same highly familiar 10-member domain of digits, yet reception of a given 9-element sequence of these digits leaves a trace which later enhances immediate reproduction of that same sequence but not (or at least not to the same degree) of a different permutation of those same elements. Hence a critical aspect of the trace left by a compound stimulus must be the organization therein of component traces. Hebb (1961, p. 43) has contended on the basis of these data that the trace of a compound stimulus must be “structural,” i.e., to consist of associations or association-like state properties. However, this does not follow any more than initial perception of the order in a compound stimulus requires the subject to possess associations among the component processes elicited by the elements in the stimulus. Whatever it is in a complex of process stages that constitutes perceived organization should be able to perseverate to the same extent that a nonrelational component of the percept can do so. That is, inasmuch as we need in any case, to admit complexly structured (organized) internal processes to account for the immediate effects of compound stimuli upon the organism, there is no reason to doubt that these organized process

21I am treating these data as though all numerals in the series were composed of the same digits in different orders. Unfortunately, this was not in fact the case in either the Hebb or the Melton experiment, but had it been so the results would presumably have been about the same.
conditions can show hysteresis and/or be mnemonically recalled.

The fundamental problem which emerges from this is simply: What is psychological “organization” (“structure”), anyway? What does it do and what is the proper way for us to talk about it? The need for a theory of structure in response processes was stressed by Lashley (1951) some years ago and has grown increasingly acute in recent work on language-syntax behavior. Correlatively, on the stimulus side of the organism, I have elsewhere pointed out that the “structure” of an environmental event is an essential contributor to its elicitation-impact (Rozeboom, 1960), while this environmental organization corresponds to the propositional form taken by cognitions within the organism (Rozeboom, 1961a, p. 482–483) And with this thought we return full circle to “memory” and its relation to traditional learning research. For a memory of some past event, whether it be carried by an iconic sensory image, an internalized verbal representation, an “imageless thought” or whatever, must be a complex of elements with the same sort of structural integrity that differentiates a cognitively meaningful sentence from a mere aggregate of disconnected words. (Otherwise it could not embody a proposition.) Roughly speaking, moreover, while the constituent elements of a given memory complex can readily be aroused in any number of ways singly or in combination, the recallability of this particular memory in its organized totality is due specifically to the rememberer’s having previously experienced a stimulus configuration to which the memory is essentially isomorphic. Over and above the cognitive “aboutness” of memories, therefore, the acquisition of memory dispositions (i.e., state properties in virtue of which memory processes can be revived) may be differentiated from the learning of elemental associations in that the former critically involves a response-integration phase (in the broadest sense of “response”) in which a complex structured process previously arousable only by correspondingly structured stimuli now becomes responsive to redintegration by cues whose own structure, if any, is not inherently relevant to the elicitation. This aspect of learning, the integration of more elementary ideational or motor processes into a unitarily elicitable compound whose internal organization corresponds to that of the stimulus which initially evoked it, is present to one degree or another in most if not all instances of acquisition and retention, including those studied in traditional learning research, and is surely one of the more important dimensions along which noncognitive processes pass over into cognitive ones. Thus while neither an image of Mary Doe’s face nor a CEP-idea is literally a memory of anything, arousal of the Doe-image by a cue previously paired with the Doe-stimulus bears closer resemblance to a genuine act of remembrance than does elicitation of the CEP-idea by stimulus BIQ. Similarly, research on “memory for form” and other investigations of a subject’s reproduction or recognition of relatively novel stimulus patterns encountered previously is closer to concern for memory proper than is study of simple association learning.
It is not feasible to go more deeply into the nature of psychological “organization” on this occasion. But not until we have learned how to think effectively about the formal structure of psychological processes will we be able to pursue the more provocative leads which have been appearing in contemporary learning research or to attain any genuine understanding of the phenomena to which the concepts of common-sense mentalistic psychology are dedicated.

References


