

Ann. Rev. Psychol. 1981. 32:21-52
Copyright © 1981 by Annual Reviews Inc. All rights reserved

HUMAN LEARNING AND MEMORY

◆339

Wayne A. Wickelgren¹

Department of Psychology, University of Oregon, Eugene, Oregon 97403

CONTENTS

MEMORY CODING.....	22
<i>Associative Memory</i>	22
<i>Direct access</i>	22
<i>Specific node encoding</i>	24
<i>Type and token nodes</i>	25
<i>Nodes and codes</i>	26
<i>Types of links</i>	27
<i>Chunking and Vertical Associative Memory</i>	28
<i>Concepts</i>	28
<i>Concepts and words</i>	28
<i>Concepts and images</i>	29
<i>Constituent vs propositional and procedural meaning</i>	29
<i>Concepts vs semantic features</i>	30
<i>Fuzzy logic</i>	32
<i>Family resemblance</i>	33
<i>Prototypes</i>	33
<i>Abstraction</i>	34
<i>Basic concepts</i>	35
<i>Episodic and semantic memory</i>	36
<i>Propositions</i>	36
<i>Units of semantic memory</i>	36
<i>Constituent structure</i>	36
<i>Plans and the Coding of Order</i>	37
<i>Schemata</i>	37
MEMORY DYNAMICS.....	37
<i>Learning</i>	38
<i>Repetition</i>	38
<i>Spacing</i>	39
<i>Levels of coding and processing</i>	40
<i>Elaboration and distinctiveness</i>	41
<i>Chunking</i>	42
<i>Storage: Consolidation and Forgetting</i>	44

¹Preparation of this article was supported in part by NSF grant BNS78-08878.

<i>Retrieval</i>	44
<i>Speed-accuracy tradeoff functions</i>	44
<i>What do we know?</i>	45
<i>Short-Term Memory</i>	46

Do you admire the insights into the nature of the human mind that have been achieved by theoretical and empirical research on learning and memory? More times than I care to remember, I have heard my colleagues denigrate our understanding of human learning and memory. Many of the best known researchers in this field claim that we understand very little about learning and memory—that at best we have made some progress in improving our experimental methods for studying learning and memory, or at worst that we have merely been amusing ourselves at the taxpayer's expense. This pessimistic view is totally and tragically wrong. Its persistence in the minds of so many eminent researchers in cognitive psychology serves, for these individuals and those who listen to them, as a self-fulfilling prophecy. Because they believe it to be impossible, they do not strive for precise and general theories that accurately integrate the welter of specific facts we know regarding learning and memory and that bestow understanding as a consequence. The overall purpose of this article is to demonstrate that a very high level of integrative theoretical understanding of human learning and memory is possible today. Many aspects of this theoretical understanding will change in the future. Indeed, the pace of change will quicken as a consequence of taking the best present-day theories seriously. However, the future changes in our theoretical understanding of learning and memory will be evolutionary, incorporating the wisdom of the past with the wisdom of the future, not revolutionary in the sense of ignoring the wisdom of the past for an ever changing succession of fads.

MEMORY CODING

Associative Memory

Koffka (1935) proposed a nonassociative theory of human long-term memory according to which the mind lays down a continuous record of experience (trace column), much like a videotape. Popular accounts often talk about memory in this nonassociative way, and even sophisticated cognitive psychologists sometimes give credence to such nonassociative theories. This is a shame, because one of the important successes of cognitive psychology is that we can confidently assert that this nonassociative theory of LTM is false (Wickelgren 1972a; 1977a, pp. 233–47).

DIRECT ACCESS From a functional standpoint, the most critical defining feature of an associative memory is the capacity for direct access retrieval of traces without search. In artificial intelligence, associative memories are

often called content-addressable when they have this direct access property that the most common location-addressable computer memories lack. In a location-addressable computer memory, one must serially search through all memory locations to find the one containing a stored (trace) pattern matching the input pattern. However, if the properties of an input pattern determine exactly where in the memory that pattern will be stored, then the memory is content-addressable, that is, associative. One is directly addressing the contents of the memory location, not just addressing the location number. Present-day computers can be programmed to achieve a very limited degree of associativity (direct access to information), but in the future, computers with large parallel processing capacity may achieve associative memories closer to human capacity (Fahlman 1979).

The main argument against human memory retrieval being a random serial search of a location-addressable store is the incredible speed that would be required—on the order of ten million locations searched per second. Since synaptic transmission appears to be the most basic functional time unit in the brain, and that is on the order of 1 msec per synapse, it is plausible to assume that searching one memory location would require at least 1 msec. This yields an upper limit on search speed of 1000 locations per second, which is at least 10,000 times too slow.

Another retrieval theory that can be rejected is the hybrid search/direct access theory that postulates direct access to a category of memory locations, e.g. those storing species of birds, followed by serial search at modest speeds (tens of milliseconds per location) through this much smaller set of locations. Corbett & Wickelgren (1978) obtained evidence contrary to the versions of this theory that have been proposed by Rosch (1973) and Rips et al (1973). Furthermore, it is very unparsimonious to postulate a direct access process that does most of the retrieval work and then have some slow search process finish the job. If you can have fast, accurate direct access to all of the bird storage locations, you should be able to have at least equally fast and accurate direct access to the hummingbird location(s).

The most plausible alternative to direct access is the hierarchical search theory in which search occurs first for the correct high-level category of locations, then within this set for a correct subordinate category of locations, and so on until the correct individual location is searched for and found (e.g. Greeno et al 1978, pp. 24–27). If searching among categories of memory locations was on the order of 10 msec per alternative category, then a seven-stage hierarchical serial search process with an average of ten alternatives at each stage would result in the capacity to search ten billion locations in 700 msec. However, the information necessary to accomplish an hierarchical search process could be used more simply and efficiently to implement a direct access retrieval process.

With direct access, the features of the input signal are processed in parallel to intersect at the locations (nodes in memory that represent each feature, segment, concept, and proposition signaled by the input. With any serial search process, one must either possess an independent means of accurately segmenting temporally distributed input (which has proved to be an intractable problem for speech recognition by serial computers) or else try potentially thousands or millions of alternative parsings into phonemes and words of even short phrases, for example. An analogous problem exists for grouping (parsing) spatially distributed input (e.g. determining the form constituents of a complex picture). The more one considers the difficulties of serial search retrieval processes for recognition of any of the tens of millions of things humans beings can recognize in a matter of seconds, the more one appreciates the power of parallel processing, direct access retrieval systems (Fahlman 1979).

Retrieval speed considerations argue strongly that the basic human long-term memory retrieval process is direct access. Furthermore, the phenomenon of redundancy gain—e.g. faster discrimination of a large bright circle from a small dim circle than of either a large from a small circle or a bright from a dim circle (Biederman & Checkosky 1970)—argues compellingly for parallel processing of retrieval cues.

SPECIFIC NODE ENCODING I have emphasized direct access as the defining property of an associative memory from a processing standpoint. From a structural standpoint, direct access appears to require the property I call specific node encoding or specific element representation (Wickelgren 1977a, 1979a), after Johannes Müller's similar neurophysiological doctrine of specific nerve energies. For every idea (feature, segment, word, image, concept, proposition, etc) that we can represent in our minds, there is assumed to be a particular set of elements that represents (encodes, stands for) the idea. Call this set of elements the node representing the idea. More generally, a node is a fuzzy set of elements, that is, a vector of weights between zero and one ($\dots w_i \dots$) representing the degree to which each element in the memory participates in representing that particular idea (e.g. Anderson 1973). The difference between specific localization and more global theories of mental representation concerns what percentages of the w entries in the vector for any particular idea are zero. Global theories assume few zero entries. Specific localization theories assume that almost all (e.g. 99% or more) of the entries for a particular node are zero. One way to add inhibition to such a specific node encoding theory is to allow w_i entries to have negative values as well as positive values, e.g. from -1 to 1 .

Psychologically, a node means nothing more than "whatever represents an idea." There must be at least one node representing any idea we have

encoded in our minds, or else, by definition, we could not think about that idea. So the only alternative to specific node encoding is multiple node encoding of each idea. What I shall call mental multiple node encoding is the encoding of an identical idea in separated (not associated) locations in memory, typically because the idea was used to encode experiences or thoughts at two different points in time (usually in two different propositions or images). The identity of the codes for the multiple encodings can be determined by search and comparison of the codes found in the separated locations. Mental multiple node encoding can be rejected because it lacks the capacity for direct access retrieval, which, as discussed in the prior section, is an established property of human memory. There is also a great deal of evidence that the encoding of an idea that is identical or similar to a previously encoded idea has different consequences from the encoding of a new idea (Wickelgren 1977a, pp. 234–43). This evidence rules out the simple tape recorder theory of human long-term memory, which has mental multiple node encoding.

Physical vs mental specificity If every occurrence of the same idea is encoded in a separate memory location, but all of these locations are directly accessible in parallel, so that whenever one is retrieved, the others are retrieved at the same time, then from a cognitive standpoint, all of the physically separate nodes are functioning as a single mental node. There may be some good physical reasons for such redundant multiplicity; to overcome limited connection capacity of a single neuron, to provide protection against cell death, etc, but we can ignore this and assume specific node encoding at a mental level.

Multiple nodes vs multiple links It is possible that repetition of an association between two nodes does not strengthen the same physical links connecting these nodes, but instead strengthens some new series of links. Still a third alternative is that repetition strengthens the same series of links, but in some way that is physically distinguishable from a stronger version of one trace. The single vs multiple link or trace issue is theoretically important and not resolved. However, so long as the same idea is always represented by the same node, there is no conflict between multiple trace theories and specific node encoding.

TYPE AND TOKEN NODES If a new token node were created for every occurrence of precisely the same idea in a new context and all tokens of the same type were not directly accessible from any one of them, then there would be a conflict with the principle of specific node encoding. The preceding arguments against nonassociative memory theories apply to this ex-

treme overuse of token nodes. Many artificial intelligence models of semantic memory have token nodes without direct access to all other token nodes of the same type simply because they are implemented on existing serial digital computers. It is doubtful that many of the inventors of such semantic memory systems think this is a desirable feature for artificial intelligence or a true feature of human semantic memory. It is an undesirable necessity that the limitations of existing computers force them to live with.

Some theories of semantic memory (e.g. Norman & Rumelhart 1975) use a new token node for every occurrence of any relational concept (signaled by verbs and prepositions) in a new proposition. The purpose of this is to encode clearly the set of concepts that constitutes each proposition and avoid some associative interference problems. For example, encoding "John hit Frank" and "Peter hit Bill" by associations with the same type node for "hit" results in a memory that cannot tell whether John hit Frank or Bill. Using unique token nodes for the relational concepts in each proposition solves this associative interference problem. However, a more elegant way to solve this problem is to use a hierarchical associative memory that introduces new nodes to stand for new propositions and can also encode nonpropositional compound concepts (e.g. a green square) by the same higher-order node mechanism (Anderson & Bower 1973). The higher-order nodes encode compound concepts and propositions by associations to and from their constituent nodes. The philosophy of this latter approach to the introduction of new nodes into associative memory is that new nodes are introduced only to encode some new idea. Old ideas are always encoded by the same nodes, in conformity with the principle of specific node encoding and permitting direct access to each idea because it has a unique location in memory.

Finally, there is no conflict between the principle of specific node encoding and findings regarding frequency (Hintzman & Block 1971) and recency (Flexser & Bower 1974) judgments that support the hypothesis that somewhat different traces are sometimes established by the separate occurrences of the same item in different contexts (different lists or positions). The same idea node can be incorporated as a constituent of many different propositions (e.g. word *i* occurred near the beginning of the first list, word *i* occurred twice in the second list, word *i* occurred many times in the first list, etc).

NODES AND CODES A brief but important message about terminology is the subject of this section. In an associative memory, there is no such thing as a code *in* a node. According to the principle of specific node encoding, the code for an idea *is* a node (or set of nodes). It is not *in* a node.

There is no point in having a code in a node unless you could have any of several different codes representing different ideas. This makes the memory nonassociative to a greater or lesser degree depending upon the variety of possible codes in each node.

TYPES OF LINKS Until quite recently, nodes in associative memory were usually assumed to be connected by only one type of excitatory link (association), and there has been very little theoretical use of inhibitory links among memory nodes outside of the perceptual and neural areas. Prior to Quillian's (1966) thesis, people occasionally discussed the possibility of labeled associations (different types of links), but, with the notable exception of inhibitory links in neural net theories, nobody did anything with it. However, beginning with Quillian, an ever increasing number of associative network models of human and artificial semantic memory have been invented, virtually all of which use a large variety of link types to express a large variety of relations between concepts (e.g. Rumelhart, Lindsay & Norman 1972, Anderson & Bower 1973, Anderson 1976, Findler 1979). Virtually all semantic memory nets, whether designed to be artificial intelligence models or models of human memory, are considered by their inventors to be associative memory networks to a greater or lesser degree, contingent upon the degree of direct access capability, but irrespective of the variety of link types. Anyone who asserts that a critical defining property of an associative memory is the assumption of a single type of link between idea nodes (e.g. Greeno et al 1978, p. 21) is behind the times. Direct access (sometimes marching under the name of "content-addressable memory") has been recognized by computer scientists as the defining property of an associative memory, at least since the early 1960s. In cognitive psychology, Wickelgren (1965; 1972a; 1977a, pp. 11-22, 220-25, 233-51; 1979a, pp. 6-11) has repeatedly pointed out that the critical defining features of an associative memory are specific node encoding and the direct access retrieval such coding makes possible.

The number of different types of links is an important and as yet unsettled theoretical issue, but the issue concerns the specific type of associative memory we have, not whether or not human memory is characterized by specific node encoding and direct access retrieval, which are widely accepted as the critical defining properties of an associative memory in the fields of computer science and semantic memory and should be so considered in all of cognitive psychology. Cognitive psychology should recognize that a major theoretical problem has been largely solved, namely, the definition of the concept of associative memory, and that a great truth has been established regarding how the mind works, namely, that it is associative.

Links bond idea nodes together to encode more complex sets and se-

quences of ideas. One day we will know how many distinct types of bonds there are, like the electrovalent and covalent bonds that hold atoms together to form molecules or the four types of force that hold the nucleus together. Psychology has as much potential for intellectual beauty as any other science. Some of that beauty can be appreciated today if you have the eye for it. A great deal more can be created in the near future if psychologists have the will to do it.

Chunking and Vertical Associative Memory

In a horizontal (nonhierarchical) associative memory, the set of idea nodes is fixed after maturation is complete, not growing with experience. Learning changes the strengths of the links connecting these nodes to each other, but does not add any new nodes to the memory (e.g. Hebb 1949). In a vertical associative memory, this horizontal associative learning process is supplemented by a vertical associative learning process, chunking, that adds new nodes to associative memory, specifying new chunk nodes to stand for combinations of old nodes. George Miller (1956) originated the concept of chunking, and its meaning was extended by many others (e.g. Estes 1972, Johnson 1972, Anderson & Bower 1973, Wickelgren 1969a, 1976a,b, 1977a,b, 1979a,b). Of course, the new chunk nodes do not appear out of thin air. From a physical standpoint, new nodes are probably added by strengthening synaptic connections to neurons that have not previously been functionally connected to associative memory, though the anatomical connections might already exist prior to the chunking learning process. From a psychological standpoint, new nodes have been added to memory by the chunking process. Intuitively, it is clear that as we acquire concepts for ever more complex combinations of simpler ideas, we are not finding it ever more difficult to think with these concepts. The new higher-level concept nodes allow us to think about complex subject matter just about as easily and efficiently as we could previously think about simpler subject matter using lower-level concepts. This is a remarkable accomplishment. It is surely an important reason for the seemingly boundless potential of the human mind to understand ever more about anything.

Concepts

CONCEPTS AND WORDS A typical word probably has thousands of different meanings (e.g. "house" can refer to any particular house you ever experienced). Dictionaries list several meanings for the typical word, but each of these refers to a large family of different specific meanings. The modern word for any particular meaning of a word or phrase is "concept." Because words do not have unique meanings, words cannot be the atomic

elements of semantic memory. The atoms of semantic memory are either concepts or semantic features. Words are high-level structural units representing ordered sets of phonetic and graphic segments.

Understanding the relation between concepts and words and that concepts, not words, are the atoms of semantic memory clarifies thinking about many problems. For example, consider the phenomenon of synonymity (Herrmann 1978). It is of no interest whether there are any "true" synonyms for any given speaker of a language, that is, pairs of words (or phrases) which are associated to exactly the same set of concepts. If there are any, they are rare. The interesting point is that any given concept can be expressed by so many different words or phrases. What could be the purpose of such duplication? One plausible answer derives from the multiplicity of concepts associated with each word and consideration of a likely role of short-term memory in understanding speech and text. What should you do in writing a paragraph when you have used the same word to refer to two different concepts? Use a synonym in place of one of the occurrences of the word. Otherwise, the reader may mistakenly retrieve the former concept to the second occurrence of the word, because its strength of association to the word was temporarily increased by the prior pairing of word and concept. Synonyms probably exist to minimize the short-term memory interference problem that derives from our efficient use of tens of thousands of words to refer to millions of concepts.

CONCEPTS AND IMAGES Concept nodes not only receive input from words, but, at least for concrete concepts, also from nonverbal stimulus cues, e.g. the feel of fur, the shape of a cat, a meow, etc as cues to activate the cat concept. The combination of features that make up each of these cues constitutes an image. It seems clear and accurate to regard concept nodes as the first level of nodes that integrate verbal and nonverbal stimuli. The constituents of concept nodes are the chunk nodes for the words and images that cue them.

CONSTITUENT VS PROPOSITIONAL AND PROCEDURAL MEANING
The meaning of a concept is given partly by the constituent words and images that activate it from below and partly by the propositions and procedures of which it is a constituent, which can activate it from above (Woods 1975, Wickelgren 1979a). Propositional and procedural meaning are equivalent to what linguistic philosophers refer to as intensional meaning. Constituent meaning is a generalized and more psychological analog of the philosophical notion of extensional (referential) meaning. Both are essential components of the meaning of concepts.

CONCEPTS VS SEMANTIC FEATURES The two principal theoretical approaches to semantic memory are the associative network theory used throughout this article and the semantic feature theory developed by Katz & Fodor (1963), Schaeffer & Wallace (1969, 1970), Meyer (1970), Clark (1973), Rips et al (1973), and Smith et al (1974). As Smith (1978, p. 23) describes these semantic feature models, "Each word is represented by a set of attributes, called semantic features . . . 'bird' would include as defining features . . . animate and feathered . . . and as characteristic features . . . a particular size . . ." "Robin" has many features in common with "bird" plus such features as "red-breasted."

The experimental testing of semantic feature theories has focused on verification and contradiction of category-example relations between concepts, e.g. a robin is a bird (high feature similarity true), a chicken is a bird (low similarity true), a bat is a bird (high similarity false), and a rat is a bird (low similarity false). Some network theories have had trouble explaining why high similarity trues are verified faster than low similarity trues, but high similarity falses are contradicted more slowly than low similarity falses. The assumption that is rejected by these data is that semantic judgment times depend entirely or primarily upon the number of links separating the concepts whose relation is being judged. If one assumes that there are two processes, verification and contradiction, initiated in parallel by these tasks, each with a characteristic asymptotic strength, and that judgments are made when the difference in retrieved strength of the two processes exceeds some critical value, then network theories account for category-example judgment findings better than existing feature theories (e.g. the results of Holyoak & Glass 1975 and Corbett & Wickelgren 1978). Such a two-process retrieval assumption, married to the associative network theory, can doubtless account for the recent antonym judgment results as well (Glass et al 1979, Herrmann et al 1979). However, a feature theory can be easily married to the same two-process retrieval theory and account just as well for these results. Indeed, McCloskey & Glucksberg (1979) suggest just such a model, although the mathematical formulation leaves a bit to be desired.

What can be concluded concerning semantic features vs concept networks? First, almost all of the semantic features anyone has discussed would be considered concepts in semantic memory. No reasonable person could deny that there are nodes encoding these features somewhere in memory. From a coding standpoint there are two issues: (a) Are semantic feature concepts the most basic concepts of semantic memory, that is, the lowest level concepts, which serve as the constituents of higher-level concepts and/or the first concepts learned by children? (b) Are concepts other than semantic feature concepts represented by unitary nodes at all, or are

these concepts represented only as sets of semantic feature nodes? I believe that we can draw firm conclusions on both of these issues.

Although the above two questions are best posed in the order I used, they are best answered in the opposite order. There is a great deal of evidence supporting the chunking capacity of the human mind, as discussed earlier. If other modalities of memory form new chunk nodes to represent combinations of more elementary constituents, it would be peculiar indeed if semantic memory did not. Semantic memory is just where we need chunking the most to avoid the enormous associative interference problem that would result from associating concepts only by associating their featural constituents, each of which participates in thousands of other concepts and so would have thousands of competing associations. If the mind represented concepts only as sets of semantic features, retrieval of information by experts would be less efficient than by amateurs and retrieval by adults would be less efficient than by children, contrary to fact.

With respect to whether semantic features are the basic concepts of semantic memory, the answer is also clear. They are not. Semantic features are necessarily superordinate concepts with wide referential generality (large sets of examples), such as physical object, animate, white, living, male, feathered, etc. The research of Rosch et al (1976) and Rosch (1977) makes it quite clear that these very general superordinate categories are not the basic level of concepts in adult semantic memory, nor are they typically the first concepts learned by children. Neither are the basic concepts highly specific (narrow referential generality), e.g. ball peen hammer, delicious apple, or collie. As Rosch demonstrates, the basic concepts are of intermediate referential generality, namely, hammer, apple, or dog (see also Piaget 1954; Brown 1958; Anglin 1977; de Villiers & de Villiers 1978, pp. 126–32). Clark & Clark (1977, p. 500) have already pointed out this fundamental contradiction to the semantic feature theory.

Rejection of the semantic feature theory in favor of the more general concept-node-and-link network theory does not mean that the general concepts referred to as “semantic features” do not exist in semantic memory. They surely do. However, they do not appear to be more simply coded nor developmentally prior to other concepts. Indeed, on the average, these general concepts are probably coded more complexly and are certainly acquired later than concepts of intermediate referential generality.

Besides the general failure of psychologists to try to decide theoretical issues, there is probably a more specific reason for the persistence of this inadequate semantic feature theory. Some of these alleged semantic features refer to sensory attributes of the stimuli that activate concepts from below. These attributes are part of the constituent meaning of a concept, and they are certainly coded at levels below the concepts they activate in semantic

memory. However, such sensory features and semantic features are very different beasts. In the first place, most of the alleged semantic features are much too abstract to be considered elementary sensory features. In the second place, the concept red is activated by the word "red" as well as by red light. The concept red is not coded in the lower levels of the visual system where the sensory feature red is first encoded. We represent red at several levels of coding in the mind. Failure to realize this causes confusion. Consistent terminology might help. I propose that we use the otherwise synonymous terms "feature" and "attribute" differentially to help remind us of the distinction between sensory features and conceptual attributes. What more appropriate way to bury the semantic feature theory than to take away part of its name?

FUZZY LOGIC Fuzzy logic (Zadeh 1965) is a generalization of standard logic that is more applicable to human thinking. Instead of assuming that the degree of membership of some instance in any concept (category) is limited to the extreme values zero or one, fuzzy logic assumes any degree between zero and one. Fuzzy logic can express various relations between different concepts, more generally, between any idea nodes in the mind. For example, Kay & McDaniel (1978) discuss a couple of different alternative rules for fuzzy intersection ($C = A \text{ AND } B$), using the example of $\text{ORANGE} = \text{RED AND YELLOW}$. They reject the standard minimum rule for fuzzy intersection, namely, that the degree of membership of an event in the intersection of two categories is the minimum of the degree of membership in either constituent category. They also discuss the standard maximum law for fuzzy union ($C = A \text{ OR } B$), using the example of $\text{GRUE} = \text{GREEN OR BLUE}$, namely, that the degree of membership in the union equals the maximum of the degree of membership in either constituent category.

Zadeh's invention of fuzzy logic is an important intellectual accomplishment both for artificial intelligence and for cognitive psychology. However, there are a great variety of possible fuzzy logics yet to be invented, especially if one relaxes the restriction that fuzzy logic be a simple extension of classical logic, including it as a special case. Bellman & Zadeh (1970) suggest a multiplicative rule for intersection ($C = AB$) and an additive rule for union ($C = A + B - AB$), that also include classical logic as a special case. However, it seems to me that in describing the mind we ought not to restrict ourselves to logics that include classical logic. Whatever the value of classical logic to mathematics, it is a multifaceted failure as a description of the "laws of thought" (Wason & Johnson-Laird 1972, Wickelgren 1979a, pp. 360-68). Fuzzy logic corrects one of the flaws of classical logic as a theory of human thought: the restriction to two-valued set membership

functions. The various fuzzy logical operators that have been suggested are interesting to consider, probably as a point of departure, rather than being satisfactory as they stand.

So far, most psychological applications of fuzzy set theory by cognitive psychologists have been banal or vague, but I am hopeful this will change. For a striking exception, see the superb application of fuzzy logic to speech recognition by Oden & Massaro (1978), which uses the multiplicative rule for combining the featural constituents of a speech segment after power function weighting of each featural constituent. The weights can be considered to be the associative strength of each feature to each segment. In general, fuzzy logic is completely compatible with the assumption of graded strengths of associations between nodes in the mind, which is a very strong point in its favor. Oden has done a number of other pioneering studies on the application of fuzzy logic to semantic memory, speech, and reading (e.g. Oden 1979).

FAMILY RESEMBLANCE Frequently, only a tiny subset of all the characteristic attributes of a concept will be sufficient to cause us to think of that concept. There is nothing common to all of the sufficient cue sets for a given concept (Wickelgren 1969a), but there are many attributes that appear frequently in these cue sets. Following Wittgenstein (1953), we say that the cue sets for a given concept have a family resemblance to each other.

PROTOTYPES The prototype hypothesis is that a concept can be characterized by an ideal set of attributes (Evans 1967, Posner & Keele 1968, Reed 1972). If we include both verbal and nonverbal constituent, propositional, and procedural meaning as constituting the entire set of ideal attributes characteristic of a concept, then clearly no single example of most concepts, and certainly no single cue set, will contain all of the ideal attributes. Thus, it is probably correct to assert that for most natural concepts, the prototype is an ideal that can never be experienced at one time. In any case, the prototype hypothesis in no way depends upon whether a real example of a concept or possible cue set for a concept actually matches the prototype.

What does the prototype hypothesis add to our characterization of concepts as family resemblances with fuzzy boundaries from other concepts? This is not entirely clear (Neumann 1977). If it were true that activating a concept from a cue set involved matching that cue set to the ideal set (prototype) for each concept (Reed 1972), then the prototype hypothesis would have importance. However, this is almost certainly not what we do. Instead, any cue set will have some strength of association to every concept node in memory, and the strongest association will win most of the time, subject to random error. The most compelling reason to prefer this associa-

tive strength hypothesis to the prototype matching hypothesis is that the strength hypothesis is exactly what an associative memory with direct access predicts. A second reason to prefer the strength hypothesis is that a theory with predictions that are probably very similar to it, namely, the context theory of Medin & Schaffer (1978), has been shown by them to account better for transfer from concept learning to classification of new instances than the independent cue type of prototype matching theory. A third reason to prefer the strength hypothesis derives from consideration of the abstraction process by which concepts are learned.

ABSTRACTION Abstraction is the process by which the input links to a chunk node from other nodes are selectively strengthened and weakened through repeated experience with different instances of the concept. These input links should include the specification of constituent, propositional, and procedural meaning of concepts, since the relations between concepts often appear to be learned very early (e.g. chairs are to sit on). This strength learning hypothesis is both reasonable and consistent with associative network theory.

It would be highly unparsimonious to opt for some prototype learning theory that cannot be incorporated easily within the more general associative network theory. For example, one way to store an average prototype without storing examples is to add the value of each attribute dimension to a counter, add its square to another counter, and add one to a number-of-instances counter. This permits determining the average value and the variance on each feature dimension. When people speak of storing the average prototype without storing the examples, this must be the sort of process they have in mind. An associative computer could be built to do this, but it is a less accurate way to record the regions in a multidimensional attribute space that are associated with any given concept than the strength mechanism previously described. Such an average prototype abstraction mechanism could learn concepts whose examples cluster about a single prototype, such as "apple" or "cat," but would be quite inadequate to learn superordinate concepts such as "fruit" and "mammal" that do not cluster about a single prototype. Furthermore, the more adequate instance strength abstraction mechanism is also more natural for an associative memory, while the prototype abstraction mechanism would be much more complex to implement.

Recently, evidence has accumulated to support the hypothesis that for adults the best example of a concept is sometimes composed of the modal (most frequently occurring) values on each attribute dimension rather than the average values (Goldman & Homa 1977, Neumann 1977, Strauss 1979), though so far infants have been demonstrated to abstract only averages

(Strauss 1979). Whether adults abstract modes or averages depends on the discriminability of the values on a dimension—highly discriminable values cause the mode to have the greatest strength of association to the concept, less discriminable values cause the average value to have the greatest strength of association to the concept. As Neumann (1977) and Wickelgren (1979a, p. 310) point out, this can be accounted for by old-fashioned stimulus generalization in an associative network theory. No other unified theory, to my knowledge, explains the abstraction of averages when discriminability is low and modes when discriminability is high. Starting most recently with Brooks (1978), there is increasing acceptance of the hypothesis that the learning of a concept derives from encoding each instance more or less faithfully, subject to selective attention in learning and to the interference and facilitation in learning, storage, and retrieval that inevitably results from encoding in an associative memory. The most plausible theory at present is that the abstracted constituent structure of a concept is the result of accumulated association of the attributes of events that cued that concept.

In addition, inhibition is also involved in some manner so that the attributes that have the strongest association to a concept are those that discriminate best between this concept and other concepts (Rosch & Mervis 1975). See also Wickelgren (1979a, pp. 311–12) for an associative explanation of this using the Spencian overlapping excitatory and inhibitory generalization gradients mechanism.

BASIC CONCEPTS If it hadn't been so sad it would have been funny to hear psychologists at one time asserting that the earliest concepts learned by children were the most general ones followed by increasing conceptual differentiation, and at other times asserting that the earliest learned concepts were the most specific ones followed by concepts of increasingly greater generality. Brown (1958) pointed out that neither extreme was true and that the earliest concepts learned by children are of intermediate referential generality, and, as discussed earlier, subsequent research has confirmed this conclusion. Brown suggested that the reason for this was greater frequency of use of these concepts by other humans in the child's environment, and this is surely an important component of the explanation for why concepts like "cat" and "apple" are basic instead of "mammal" or "fruit" on the one hand or "Siamese cat" or "Jonathan apple" on the other. However, it was not until the elegant insight of Rosch (Rosch et al 1976, Rosch 1977) that we understood this phenomenon adequately. According to Rosch, basic concepts are the most general concepts whose examples are sufficiently similar that they hover around a single prototype. This high degree of example similarity means extensive strength generalization in the

integrative learning of separate examples, which, in effect, increases the frequency of pairing each attribute node with the concept node. On the average, the basic level concepts are those where the two factors contributing to associative strength, example frequency and example similarity, typically hit the maximum in their combined effect on speed of concept learning. Hence these concepts are typically learned first. At the same time, this explanation makes it clear how exceptions could occur when a child has a great deal of experience with a more general or less general concept than the corresponding basic level concept. I have taken a little license with Rosch's idea to make it compatible with the associative-strength, instance-learning theory of abstraction advocated in the prior section.

EPISODIC AND SEMANTIC MEMORY If you accept the instance memory theory of abstraction presented earlier, you have the best extant explanation of the relation between episodic and semantic (generic) memory. There is no reason to believe that these are two different forms of memory with either different coding structure or different cognitive processes. Unique experiences will by definition not be merged with subsequent experiences via the abstraction process. As a consequence, unique experiences will generally be forgotten unless they are frequently recalled. When they are recalled, they become subject to integration and abstraction with the recall experience. Thus, except for the first recall of a unique experience, there is no episodic memory, only various degrees of generic memory.

Propositions

UNITS OF SEMANTIC MEMORY A proposition contains a relational concept (verb) and one or more argument concepts (agent, object, etc). The cognitively correct definition of a proposition is not known, but people have a very high degree of agreement concerning the analysis of sentences into component propositions, so it has been possible to establish that propositions are units of coding above concepts in semantic memory (Wickelgren 1979a, pp. 317-21).

CONSTITUENT STRUCTURE A proposition is a set of concepts, but what kind of set is unclear at the present time. Neither predicate nor case constituent structure has proved completely satisfactory (Doshier 1976). Given that concept nodes and phonetic and graphic segment nodes are fuzzy combinations of their constituents, it is parsimonious to assume that propositions are fuzzy combinations of their concept constituents as well (Oden 1978, Wickelgren 1979a, pp. 321-25). Although it probably does not apply to networks with redundant link paths between nodes, Cunningham (1978)

has devised a very interesting scaling method for determining network structure from dissimilarity data.

Plans and the Coding of Order

Propositions are units of declarative knowledge. Plans are units of procedural knowledge. It may not be useful to distinguish these two types of knowledge, but in agreement with Anderson (1976), I currently believe it is. My theory is that propositions are unordered sets of concepts, while plans are ordered (most generally, partially ordered) sets of concepts. While propositions can encode all knowledge, including knowledge about actions and the temporal sequence of events, this knowledge is not in the proper form for the temporal control of human action (by which I mean all mental actions, not just those mental actions that directly control obvious physical actions by the muscles). Plans achieve this control of sequences of mental actions contingent upon internal and external events. Plans have been described using augmented transition networks (Woods 1970), production systems (Newell 1973, Anderson 1976), and context-sensitive coding (Wickelgren 1979a, pp. 357–60; 1979c). As a theory of the coding of order in phonetic and graphic segments in words, context-sensitive coding is incomparably superior to any other theory (Wickelgren 1969b,c, 1972b, 1976c, 1979a), but it is far too early to conclude anything about its application to the encoding of syntax and other higher-order procedural knowledge.

Schemata

Some people believe that there is a level in the mind above propositions and plans that integrates sets of these into single units called schemata. A more parsimonious alternative is that a schema is a mental set of primed (partially activated) nodes (and links?) that results automatically from spreading activation from the activation of nodes representing prior thoughts (Wickelgren 1979a, pp. 325–27). Priming via spreading activation is a natural part of an associative memory, and there is considerable evidence supporting such a hypothesis (Collins & Loftus 1975, Wickelgren 1979a). There is no evidence supporting the hypothesis of unitary schema nodes, although they could exist.

MEMORY DYNAMICS

Memory has three temporal phases: learning, storage (consolidation and forgetting), and retrieval (recall or recognition). Widespread understanding and acceptance of the learning, storage, and retrieval distinction has greatly facilitated scientific research on memory. Another basic and obvious dis-

inction is sorely needed, that between the macro and micro levels in the cognitive analysis of learning, storage, and retrieval. This distinction is analogous to that between macroeconomics and microeconomics.

Microlearning is concerned with the learning of a single association or a small set of associations encoding a single chunk or a small set of chunks. Microstorage is concerned with the subsequent consolidation and forgetting of such small atoms or molecules of learned information. Microretrieval is concerned with a single elementary act of recalling or recognizing some unit of information (e.g. word, concept, proposition).

To explicate the distinction, let us consider the retrieval process in more detail. One elementary act of microretrieval takes about one second, varying from .2 to 2 sec depending upon the coding level (Wickelgren 1979a, pp. 270–78). Fortunately, because of chunking, one elementary act of microretrieval can be of a very complex chunk of information with numerous constituents. Furthermore, under many conditions, several separate (unchunked) traces can be retrieved in parallel, within the attention span. Basically, if the retrieval takes place in a couple of seconds or less, you are studying microretrieval. If the retrieval is extended over tens of seconds, minutes, or longer, as in free or ordered recall of lists or text, this is a complex retrieval process consisting of a sequence of elementary acts of recall and recognition, and perhaps some inference processes as well. Such tasks are studying macroretrieval. The study of single inference operations or small sets or sequences of inference, recall, and recognition operations is at the micro level, but problem solving, creativity, and comprehension of large units of text are at the macro level. This review is only concerned with the micro level of learning, storage, and retrieval. Accordingly, I tend to ignore all multiple recall (especially free recall) studies, where there is little control over the control processes, retrieval cues, and retrieval time for each elementary act of retrieval.

Learning

REPETITION “Practice makes perfect”—or, in any case, it makes memory traces stronger. It is still possible that memory traces are composed of varying numbers of component traces each of which is learned, consolidated, forgotten, and/or retrieved in an all-or-none manner, but no evidence compels this assumption. To study memory dynamics independently from coding, it is simpler to characterize traces by continuous strengths, given that forgetting and retrieval functions are always continuous incremental functions of storage and retrieval time, and learning curves are almost always continuous incremental functions of study time. There do appear to be sudden jumps in strength during learning, which probably result from insights regarding good ways to code material (e.g. integrating the to-be-

learned material with related already-learned material). However, even when a single jump in learning boosts retrieval from chance to 100% accuracy, subsequent learning will further increase strength, as shown by increases in strength ratings and in retrieval speed (Corbett 1977, Wickelgren 1977a, p. 320). Further support for the incremental nature of memory traces comes from Nelson (1977), who found that repetition at the same phonetic level of coding, even with the same question each time (does the word have an r sound), produced increased learning, contrary to the assertion that increased learning results only from adding traces at higher levels of processing.

SPACING Spaced repetitions almost always benefit memory more than massed repetitions. Progress has been made in understanding the reasons for this phenomenon. We can classify theories of the spacing effect into three categories: encoding variability, deficient consolidation or rehearsal of the first presentation, and deficient learning of the second presentation. Hintzman (1976) argued that the first two categories of theories could be ruled out on the basis of prior evidence, and since then an ingenious study by Ross & Landauer (1978) provides more evidence against all encoding variability explanations. A novel experiment by Jacoby (1978) further supports the deficient second-trial learning hypothesis. Jacoby suggests that subjects do not go through the same recognition and other coding processes for a massed second trial as for a first trial or a spaced second trial. With a massed second trial, the end product of these processes can be activated without all of the preliminary coding required by first or spaced-second trials. This deficient coding hypothesis is doubtless the explanation of Jacoby's results, which were obtained using an unusual learning task. However, in more standard learning tasks, the levels (types) of coding and processing may be identical, but the speed of achieving the encoding is much faster for massed second trials. If increases in degree of learning are closely tied to the encoding process and not to the period of maintained activation after encoding, then faster encoding may necessarily produce less learning.

One of the persistent problems with almost all research on spacing effects is the failure to separately assess the effects of spacing on learning and retention. This can only be done by obtaining retention functions for the single-trial learning condition and for each two-trial learning condition (Wickelgren 1972c, Reed 1977). The increment to the memory trace contributed by second-trial learning can be estimated by subtracting the memory strength due to the first learning trial at the appropriate retention interval from the total learning strength obtained with two learning trials. This can be done at a variety of retention intervals following the second learning trial. Thus, not only can the second trial learning increment be

estimated, but also the retention function for the second-trial learning increment can be determined for comparison across spacing conditions and to the first-trial retention function. This makes the assumption that the second trial does not destroy or alter in any way the memory trace resulting from the first trial. The evidence reviewed by Hintzman (1976) and the results of Wickelgren (1972c) are in agreement with this assumption.

Retention function analyses of the spacing effect produce the following conclusions: (a) Second-trial learning increases monotonically with spacing, perhaps indefinitely, though the maximum increases are achieved during the first 10 to 30 minutes after first-trial learning (Wickelgren 1972c and unpublished data). (b) The memory trace for the first trial is probably unaffected by second trial learning (Wickelgren 1972c, Hintzman 1976). (c) Consolidation and forgetting of the second-trial's trace is probably identical to that for the first trial; certainly, retention of the second-trial trace is at least as good as that of the first trial (Wickelgren 1972c and unpublished data).

All of this tends to produce a beneficial effect of greater spacing between learning trials under many circumstances. However, shorter spacings have the advantage of shorter retention intervals for the first learning trial at any given retention interval following the second trial. This can produce advantages for shorter over longer spacings. It is foolish to do research on such a complex quantitative problem as the spacing effect without a mathematical modeling approach and without obtaining complete retention functions.

LEVELS OF CODING AND PROCESSING The levels of processing fad is over in the field of learning and memory. Of course, the concept of levels of coding (and processing) existed long before the fad started and will continue to be an essential, though not yet completely specified, concept for understanding the structure of the mind. One of the criticisms of the levels fad was that it failed to add to the precise theoretical specification of the levels concept in either coding or processing domains. Another flaw with the faddish conception was that its specific claims about learning and memory turned out to be almost entirely false. First, it is clearly not the level of processing that matters for degree of learning and retention, but the trace code that results from this processing. Semantic processing by itself is no guarantee of a high level of learning. For example, deciding that an ostrich is not a geographical location results in very little memory for having processed "ostrich," compared to deciding that an ostrich is a living thing (Schulman 1971). Second, high levels of performance on memory tests are not guaranteed by a semantic coding level either. What counts most is the relation between the information stored and the information questioned in retrieval. Phonetic traces are what will help you perform best on a phonetic

retrieval test, semantic on a closely related semantic test, etc (Morris, Bransford & Franks 1977; McDaniel, Friedman & Bourne 1978; Stein 1978). Third, substantial learning takes place at lower structural levels of coding and, as mentioned earlier, repetition of processing at lower structural levels further increases degree of learning; it is not necessary to process at semantic levels to achieve further increases in degree of learning (Nelson 1977). Fourth, despite the plausibility of the hypothesis that coding is more distinctive at semantic than at structural levels and thus less subject to interference, forgetting of many structural traces is not especially rapid and often there has been no difference in forgetting rate between structural and semantic traces (Nelson & Vining 1978). For all these reasons, the levels of processing framework has been largely abandoned. However, I believe it left a useful legacy of facts and ideas generated in its disproof, and it generated interest in levels of coding and processing, which is surely a critically important concept in understanding the mind. At least, now no knowledgeable cognitive psychologist should have trouble explaining to students what it means to read a page of words and suddenly realize that you haven't been processing the meanings at all.

ELABORATION AND DISTINCTIVENESS Two other explanations of the variability in learning are the elaboration hypothesis (Craik & Tulving 1975, Anderson & Reder 1979) and the distinctiveness hypothesis (Klein & Saltz 1976, Wickelgren 1977a, Eysenck 1979, Jacoby & Craik 1979). The elaboration hypothesis is that traces with many constituents or many associations to other traces are well learned and well remembered. The distinctiveness hypothesis has two principal variants, which are complete opposites, though no one seems to have realized this before: (a) Elaborate traces are more distinctive and therefore have more effective possible retrieval cues and are also less subject to interference in storage and retrieval. (b) Traces whose constituents have few interfering associations to other traces are more distinctive and thus remembered better because they are less subject to interference. The latter conception of what makes a trace less subject to interference is clearly the correct one, so the former conception appears to be left with only the multiple retrieval cues explanation for the effectiveness of elaboration-distinctiveness.

Choosing between these alternatives is difficult. On the one side is all the evidence from decades of verbal learning research and more recently from semantic memory research (e.g. Anderson 1974, King & Anderson 1976, Bower 1978) that learning multiple associations to the same nodes causes interference. On the other side is the evidence that elaboration often aids learning and memory, despite the fact that embedding the material in a sentence or image, relating the new material to schemata from existing

knowledge, etc. appears to be increasing the potential for proactive and retroactive interference. Such interference has been demonstrated (Owens et al 1979). To add to the problems with the elaboration hypothesis, elaboration has not always been found to provide a net benefit to memory (Nelson et al 1978, Stein et al 1978, Morris et al 1979). Nelson et al found that interactive images were superior to noninteractive images for learning word pairs, but that multiple interactive images produced no better memory than a single interactive image. Of course, we know for certain that multiple associations are important for making a piece of knowledge related to all that it should be related to in memory, but these results demonstrate that there is no huge benefit to redundant multiple associations connecting the same nodes. The Stein et al and Morris et al studies indicate that certain kinds of elaboration increase memorability more than other kinds. Elaboration that merely adds to what must be learned does not increase the memorability of the other material and often decreases it. Elaboration that increases the amount of the other material that can be encoded using already learned schemata is beneficial to learning. This means that elaboration of the stimulus input may be benefitting memory because it is requiring a *less* elaborate addition to memory than nominally simpler stimulus input.

CHUNKING Of course, it would be wrong to go overboard on what could be called the reduction hypothesis, that we learn best when we have the least to add to what is already in memory. In the extreme, when everything in some input material is already encoded in memory, we often do further strengthen the existing associations, but this is not when the greatest amount of learning occurs. Learning curves always indicate diminishing returns after an earlier period of positive acceleration, though only studies measuring learning on an unbounded strength scale are relevant for establishing such diminishing returns without ceiling artifact (e.g. Wickelgren & Norman 1971, Wickelgren 1972c). Information that is in the part of a sentence that the syntax signals to be new is better learned than information in the part the syntax signals to be given (Hornby 1974, Singer 1976). Nickerson & Adams (1979) discovered that people have extraordinarily poor recall and recognition memory for the visual details of a very familiar object, a United States penny. Thousands of experiences recognizing pennies clearly led to little learning of its features once it became highly familiar on the basis of some subset of features. All of these observations support the hypothesis that learning is maximal at intermediate degrees of prior chunking (integration) of the material to be learned (Wickelgren 1979a, pp. 119–20).

Probably the primary reason that maximal learning occurs for material with intermediate degrees of prior integration into chunks is that this is

where maximal new chunking (specification of new nodes) occurs. When the current experience is already integrated under one top chunk node, as in recognizing a penny, little or no further chunking occurs. When the experience is an unfamiliar combination of too many separate (ungrouped or unrelated) parts, as with some of my lectures, the entire experience is too complex to be integrated into one chunk at that time. Chunking may occur, but it will be of only one or a few subsets of the component features of the complex unfamiliar experience. Such partial chunking is under the control of the subject's attentional and grouping strategies, and as a consequence may be difficult for an experimenter to measure.

According to this theory of chunking, one way to obtain further chunking of a highly familiar entity would be to focus attention on subsets of features and the relations between them. This may explain why learning of familiar words and pictures has often been enhanced by increasing the difficulty of processing the material: presenting incomplete pictures (Kunene et al 1979), requiring completion of incomplete sentences or recognition of inverted text (Kolers & Ostry 1974, Masson & Sala 1978), presenting misspelled words for error correction (Jacoby et al 1979). Along the same line, Tyler et al (1979) found that greater difficulty in finding the correct word to complete a sentence or solve an anagram led to better memory for the target words.

The chunking and consolidation theory proposed by Wickelgren (1979b) can be extended in an elegant way to account for a variety of learning phenomena. The theory assumes the following: (a) There are two (fuzzy) sets of nodes in cortical associative memory, bound and free. Bound nodes have strong input and output links to other nodes in cortical memory and weak links to the hippocampal chunking arousal system. Free nodes have weak input and output links to other cortical nodes and strong input links from, and possibly to, the hippocampal chunking arousal system. (b) In chunking, the chunking arousal system primes the free nodes so that they can compete successfully for activation against other bound nodes to which the set of bound nodes to be chunked may already be associated. The node or set of nodes maximally activated (by the weak links from the bound nodes to be chunked) will increase in activation so that by standard nodal contiguity conditioning (Hebb 1949), the links associating the new chunk node to its constituent nodes will be strengthened. This binds the formerly free chunk node to represent the set of its constituent nodes. (c) Lateral inhibition among cortical nodes limits the total number of nodes (free or bound) that can be fully activated at any one time, preventing epilepsy (Milner 1957). Activating a set of bound nodes that is already well integrated under a bound top chunk node produces a very high level of activation of these nodes. This inhibits the activation of all other nodes, including free nodes, thus preventing further chunking. Activating less strongly asso-

ciated sets produces less inhibition, permitting some further integration of these bound nodes by strengthening associations from various combinations of these bound nodes to free chunk nodes. (d) Binding a chunk node to its constituents initiates a consolidation process that cumulatively disconnects the now bound chunk node from the chunking arousal system, preventing it from being recruited to represent some other combination of constituents.

Storage: Consolidation and Forgetting

The theory of consolidation and forgetting described in Wickelgren (1972c; 1974; 1977a, pp. 362–94; 1979b) is, in my opinion, much more general, accurate, and elegant than any other, and, except for the modifications concerning chunking and spacing effects in learning discussed in the last section, I have nothing to add to the theory at this time.

Retrieval

SPEED-ACCURACY TRADEOFF FUNCTIONS The study of the dynamics of memory retrieval, and indeed of all cognitive processes, has been given a powerful new experimental tool comparable in significance to the invention of the microscope. A. V. Reed (1973, 1976) had the insight that the phenomenon of speed-accuracy tradeoff in reaction time could be used to study the time course of all mental information processing (all of which is memory retrieval in the broadest sense). Speed-accuracy tradeoff functions have three major advantages over reaction-time measures (even accompanied by an accuracy measure): (a) It is not possible to meaningfully compare the difficulty of two conditions when one condition has a shorter reaction time, but a lower accuracy level, than another condition. Obtaining the entire speed-accuracy tradeoff function for each condition avoids this problem and permits comparison of two conditions without the possibility of an invalidating speed-accuracy tradeoff (Wickelgren 1977c). (b) Obtaining the entire speed-accuracy tradeoff function provides much more extensive information concerning the time course of retrieval dynamics than does a reaction-time experiment, which provides the equivalent of a single point of such a function, typically at a point near the asymptote, where retrieval dynamics is over (Wickelgren 1977c). (c) Speed-accuracy tradeoff functions permit separate estimation of the strength of a memory trace in storage and its retrieval dynamics parameters and functional form. Reaction-time measures completely confound storage and retrieval. A variety of phenomena, most obviously subject's ratings of trace strength, establish that there are many levels of trace strength above the minimum level necessary to yield 100% correct performance at asymptote (unlimited retrieval time). Unless those asymptotic (stored) strength differences are estimated and factored out using speed-accuracy tradeoff methods and, if necessary, incremental scaling (Wickelgren 1978, Wickelgren et al 1980), one cannot study re-

trieval dynamics, because storage and retrieval are confounded. Thus, in particular, just because subjects always respond correctly that “a robin is a bird” and “a chicken is a bird” does not mean that these traces have equal strength in storage. Indeed, subject’s ratings of trace strength establish that they do not. Virtually all reaction-time studies of memory are uninterpretable with respect to retrieval dynamics because of this confounding of storage and retrieval.

WHAT DO WE KNOW? Since the thousands of reaction-time studies of memory retrieval dynamics tell us very little that is definitive, and there have been only a few tens of speed-accuracy tradeoff studies, we know much less about retrieval dynamics than we ought to know. Nevertheless, we can draw some conclusions about memory retrieval dynamics with varying degrees of certainty: (a) Microretrieval in both recall and recognition is a simple direct-access process, not a search process (Corbett & Wickelgren 1978, Doshier 1976, Remington 1977, Wickelgren & Corbett 1977). Since this conclusion is supported by the arguments for associative memory described earlier, we can be quite confident of it. (b) Tentatively, microretrieval in recall is identical in dynamics to recognition (Wickelgren & Corbett 1977). (c) Within the span of attention, we can process several retrieval cues in parallel with little or no loss in efficiency (Doshier 1976, Wickelgren & Corbett 1977). The major uncertainty is just how little the loss is, what the span of attention is, and how each is affected by various conditions. (d) Retrieval of each link in a chain of links occurs, not strictly serially, but in a partially overlapping manner, such that partial retrieval at a lower level initiates retrieval at the next higher level and continually updates its input to the next level as its own retrieval becomes more complete. Retrieval of many, and sometimes all, levels is occurring in parallel. Turvey (1973) calls this parallel-contingent processing, Wickelgren (1976a) calls it chain-parallel processing, and McClelland (1979) probably has the best name for it, cascade processing. There is quite a lot of evidence supporting cascade processing. (e) Despite cascade processing, coding level is perhaps the most important determinant of retrieval dynamics. Higher-level traces (longer associative chains from input to output) have slower retrieval dynamics (Wickelgren 1979a, pp. 273–76). (f) Repeated retrieval of an associative chain in some cases can result, not only in increased strength, but also in a short-circuiting of the chain, speeding retrieval dynamics (Corbett 1977). Thus, both theories of automatization are correct (Wickelgren 1979a, pp. 276–78).

Finally, one of the most important accomplishments in the study of retrieval is purely conceptual. Some of us now clearly understand that memory traces should be considered to have at least two extreme states: retrieved (on our minds, conscious, in active memory) and unretrieved (not

on our minds, unconscious, in passive memory). This must be true, because we are not thinking all thoughts simultaneously, at least not to the same degree. Speed-accuracy tradeoff functions are most simply interpreted as indicating that the transition from the unretrieved state to the retrieved state is an incremental rather than an all-or-none process, but we have yet to establish this definitively. Persistence in the retrieved (active memory) state is one type of short-term memory.

Short-Term Memory

I will use "short-term memory" to refer to any rapidly forgotten memory, regardless of whether the trace on which the memory is based is the same or different from associative long-term memory. It is clear that one form of short-term memory is different from long-term memory, namely, active (primary) memory. I believe that this is the only distinct form of short-term memory. There is long-term associative memory (passive memory) and there is the subset of this passive memory that is currently in various states of activation. Such active memory is the only true form of short-term memory, and active memory is limited to the trace(s) currently being thought of and, to a lesser extent, the traces associated to this attentional focus. Active memory is equivalent to the span of attention, modified to include both a focus and an associative halo of decreasing activation. The focus is what is being consciously attended to, and the halo includes the entire attentional set of traces primed (partially activated) by the traces in the attentional focus.

Because we are not thinking all thoughts simultaneously and because time is required to retrieve new thoughts, it is clear that there is one kind of short-term memory, namely, active memory. In bridging the gap from active memory to the sort of memory that every cognitive psychologist agrees is long-term memory, there are three principal theoretical alternatives: (a) There is only active short-term memory and associative long-term memory. Active memory is limited to the currently activated thought and its associative halo (the span of attention). Previously activated thoughts, even the immediately previous one, are not in active memory. So-called short-term memory for short lists is just long-term associative memory under conditions where forgetting is rapid, presumably due to high levels of interference. (b) There is only active short-term memory and associative long-term memory. Active memory extends to encompass short lists of items (the span of immediate memory). (c) There are three dynamically distinct types of traces: active memory, short-term memory, and long-term memory. Active memory is limited to the span of attention. Short-term memory mediates the span of immediate memory. Long-term memory accounts for everything else.

The third alternative simply cannot be supported at the present time because all of the known memory phenomena can be accounted for more parsimoniously with but two traces, active memory and long-term memory (Wickelgren 1973, 1974, 1975, 1979b). I supported the second alternative for a time (Wickelgren 1979a), but a recent experiment (Wickelgren et al 1980) definitively rules out the second alternative in favor of the first. Wickelgren et al reasoned that since retrieval converts traces from the passive to the active state, traces that are still in active memory to some extent (primed or partially activated) should have faster retrieval dynamics. Accordingly, if the basis of short-term memory for lists (probe memory span) is active memory, then the decline in asymptotic recognition accuracy with more intervening items should have the same dynamics as the decline in the priming effect on retrieval dynamics. What we found was that the priming effect on retrieval dynamics was strictly limited to the very last item in the list. The retrieval dynamics of all other items in the list were identical, despite massive changes in asymptotic recognition memory accuracy. It appears that active memory is confined to the very last thought (and whatever it is associated to in long-term memory, to a lesser extent). The rest of the span of immediate memory is most parsimoniously attributed to associative long-term memory.

A very remarkable property of the retention function for long-term memory is that it automatically includes a rapidly decaying short-term memory buffer without the need for a separate short-term memory system. The reason for this is that the rate of forgetting in long-term memory is initially very rapid and continually slowing down with increasing trace age. This initially rapid fall in trace strength means that it is impossible to push the strength of long-term memory traces up to their maxima for more than a fraction of a second, and this, in turn, means that there is always room at the top of the current levels of long-term memory to add a short-lasting increment. This short-lasting increment to the long-term memory trace functions as a short-term memory buffer and is the basis for the span of immediate memory. Besides allowing us to dial telephone numbers without continually looking at the phone book, this short-term aspect of associative memory doubtless plays an important role in speech recognition, articulation, and reading.

It is also what accounts for "warm-up" effects when we sit down to think and write. At first we have trouble getting our thoughts to flow until we have retrieved some associations and strengthened them. As we all know, this strengthening dissipates to a large extent with time and interference. Hence, if our thinking on a particular problem is too fragmented in time, we find ourselves spending a large fraction of our time warming up our minds reviewing old thoughts and not enough time producing new thoughts. However, it is clear that such short-term memory is not as rapidly lost as

list memory in a typical probe memory span experiment. The difference is presumably due to differences in the susceptibility of the traces to interference (Wickelgren 1974, 1975), but there may also be important differences in the degree of learning that account for some of the differences in how long the memory lasts as well.

Another remarkable property of long-term retention is that modest differences in degree of learning are amplified to produce enormous differences in trace longevity, even with no difference in forgetting rate (Wickelgren 1977a, pp. 371–72). For example, a factor of two in learning can increase trace longevity by a factor of 100. Hence, what appears to be a difference in “forgetting rate” is very often merely a difference in degree of learning. When initial degree of learning in two conditions is sufficient to produce nearly perfect performance on immediate retention tests, people often erroneously assume that initial degree of learning is equal, when the differences could be quite substantial. With the large amplification factor that prevails between degree of learning and trace longevity, failing to control degree of learning has often led to wrong conclusions. In any case, judging by trace longevity, there are not just two categories of memories, short-term and long-term. There are memories that last seconds, memories that last minutes, memories that last hours, memories that last days, and memories that last months and years. Although we may find evidence for some biological separation into several stages of memory storage, there is currently no psychological reason to peel off the memories that last seconds and call them a different kind of memory trace. Associative long-term memory has solved the whole problem of having memories last for widely varying times, presumably because such limited longevity is functional. Functionally, there is a continuous spectrum of trace lifetimes. Theoretically, except for active memory, they all derive from the same remarkable associative memory. Even active memory is but a change of state for traces in the same associative memory system, but this state change is a dynamically different trace from passive long-term memory.

Literature Cited

- Anderson, J. A. 1973. A theory for the recognition of items from short memorized lists. *Psychol. Rev.* 80:417–38
- Anderson, J. R. 1974. Retrieval of propositional information from long-term memory. *Cognit. Psychol.* 4:451–74
- Anderson, J. R. 1976. *Language, Memory, and Thought*. Hillsdale, NJ: Erlbaum. 546 pp.
- Anderson, J. R., Bower, G. H. 1973. *Human Associative Memory*. Washington DC: Winston. 524 pp.
- Anderson, J. R., Reder, L. M. 1979. An elaborative processing explanation of depth of processing. See Eysenck 1979, pp. 385–403
- Anglin, J. M. 1977. *Word, Object, and Conceptual Development*. New York: Norton. 302 pp.
- Bellman, R. E., Zadeh, L. A. 1970. Decision-making in a fuzzy environment. *Manag. Sci.* 17:141–64
- Biederman, I., Checkosky, S. F. 1970. Processing redundant information. *J. Exp. Psychol.* 83:486–90

- Bower, G. H. 1978. Interference paradigms for meaningful propositional memory. *Am. J. Psychol.* 91:575-85
- Brooks, L. 1978. Nonanalytic concept formation and memory for instances. In *Cognition and Categorization*, ed. E. Rosch, B. B. Lloyd, pp. 169-211. Hillsdale, NJ: Erlbaum. 328 pp.
- Brown, R. 1958. *Words and Things*. New York: Free Press, 315 pp.
- Clark, E. V. 1973. What's in a word? On the child's acquisition of semantics in his first language. In *Cognitive Development and the Acquisition of Language*, ed. T. E. Moore. New York: Academic. 308 pp.
- Clark, H. H., Clark, E. J. 1977. *Psychology and Language*. New York: Harcourt. 608 pp.
- Collins, A. M., Loftus, E. F. 1975. A spreading-activation theory of semantic processing. *Psychol. Rev.* 82:407-28
- Corbett, A. T. 1977. Retrieval dynamics for rote and visual image mnemonics. *J. Verb. Learn. Verb. Behav.* 16:233-46
- Corbett, A. T., Wickelgren, W. A. 1978. Semantic memory retrieval: Analysis by speed accuracy tradeoff functions. *Q. J. Exp. Psychol.* 30:1-15
- Craik, F. I. M., Tulving, E. 1975. Depth of processing and the retention of words in episodic memory. *J. Exp. Psychol.-Gen.* 104:268-94
- Cunningham, J. P. 1978. Free trees and bidirectional trees as representations of psychological distance. *J. Math. Psychol.* 17:165-88
- de Villiers, J. G., de Villiers, P. A. 1978. *Language Acquisition*. Cambridge, Mass: Harvard. 312 pp.
- Doshier, B. A. 1976. The retrieval of sentences from memory: a speed-accuracy study. *Cognit. Psychol.* 8:291-310
- Estes, W. K. 1972. An associative basis for coding and organization in memory. In *Coding Processes in Human Memory*, ed. A. W. Melton, E. Martin, pp. 161-90. New York: Wiley. 448 pp.
- Evans, S. H. 1967. A brief statement of schema theory. *Psychon. Sci.* 8:87-88
- Eysenck, M. W. 1979. Depth, elaboration, and distinctiveness. In *Levels of Processing in Human Memory*, ed. L. S. Cermak, F. I. M. Craik, pp. 89-118. Hillsdale, NJ: Erlbaum. 479 pp.
- Fahlman, S. E. 1979. *NETL: A System for Representing and Using Real-World Knowledge*. Cambridge, Mass: MIT. 278 pp.
- Findler, N. V., ed. 1979. *Associative Networks*. New York: Academic. 462 pp.
- Flexser, A. J., Bower, G. H. 1974. How frequency affects recency judgments: A model for recency discrimination. *J. Exp. Psychol.* 103:706-16
- Glass, A. L., Holyoak, K. J., Kiser, J. I. 1979. Role of antonymy relations in semantic judgments. *J. Exp. Psychol.-Hum. Learn. Mem.* 5:598-606
- Goldman, D., Homa, D. 1977. Integrative and metric properties of abstracted information as a function of category discriminability, instance variability, and experience. *J. Exp. Psychol.-Hum. Learn. Mem.* 3:375-85
- Greeno, J. G., James, C. T., DaPolito, F. J., Polson, P. G. 1978. *Associative Learning: A Cognitive Analysis*. Englewood Cliffs, NJ: Prentice-Hall. 241 pp.
- Hebb, D. O. 1949. *The Organization of Behavior*. New York: Wiley. 335 pp.
- Herrmann, D. J. 1978. An old problem for the new psychosemantics: synonymity. *Psychol. Bull.* 85:490-512
- Herrmann, D. J., Chaffin, R. J. S., Conti, G., Peters, D., Robbins, P. H. 1979. Comprehension of antonymy and the generality of categorization models. *J. Exp. Psychol.-Hum. Learn. Mem.* 5:585-97
- Hintzman, D. L. 1976. Repetition and memory. In *The Psychology of Learning and Motivation*, ed. G. H. Bower, 10:47-91. New York: Academic. 247 pp.
- Hintzman, D. L., Block, R. A. 1971. Repetition and memory: Evidence for a multiple-trace hypothesis. *J. Exp. Psychol.* 88:297-306
- Holyoak, K. J., Glass, A. L. 1975. The roles of contradictions and counterexamples in the rejection of false sentences. *J. Verb. Learn. Verb. Behav.* 14:215-39
- Hornby, P. A. 1974. Surface structure and presupposition. *J. Verb. Learn. Verb. Behav.* 13:530-38
- Jacoby, L. L. 1978. On interpreting the effects of repetition: solving a problem versus remembering a solution. *J. Verb. Learn. Verb. Behav.* 17:649-67
- Jacoby, L. L., Craik, F. I. M. 1979. Effects of elaboration of processing at encoding and retrieval: trace distinctiveness and recovery of initial context. See Eysenck 1979, pp. 1-21
- Jacoby, L. L., Craik, F. I. M., Begg, I. 1979. Effects of decision difficulty on recognition and recall. *J. Verb. Learn. Verb. Behav.* 18:585-600
- Johnson, N. F. 1972. The role of chunking and organization in the process of recall. In *The Psychology of Learning and Motivation*, ed. G. H. Bower, 4:171-247. New York: Academic. 315 pp.

- Katz, J. J., Fodor, J. A. 1963. The structure of a semantic theory. *Language* 39:170-210
- Kay, P., McDaniel, C. K. 1978. The linguistic significance of the meanings of basic color terms. *Language* 54:610-46
- King, D. R. W., Anderson, J. R. 1976. Long-term memory search: an intersecting activation process. *J. Verb. Learn. Verb. Behav.* 15:587-605
- Klein, K., Saltz, E. 1976. Specifying the mechanisms in a levels-of-processing approach to memory. *J. Exp. Psychol.-Hum. Learn. Mem.* 2:671-79
- Koffka, K. 1935. *Principles of Gestalt Psychology*. New York: Harcourt. 485 pp.
- Kolers, P. A., Ostry, D. J. 1974. Time course of loss of information regarding pattern analyzing operations. *J. Verb. Learn. Verb. Behav.* 13:599-612
- Kunen, S., Green, D., Waterman, D. 1979. Spread of encoding effects within the nonverbal domain. *J. Exp. Psychol.-Hum. Learn. Mem.* 5:574-84
- Masson, M. E. J., Sala, L. S. 1978. Interactive processes in sentence comprehension and recognition. *Cognit. Psychol.* 10:244-70
- McClelland, J. L. 1979. On the time relations of mental processes: an examination of systems of processes in cascade. *Psychol. Rev.* 86:287-330
- McCloskey, M., Glucksberg, S. 1979. Decision processes in verifying category membership statements: Implications for models of semantic memory. *Cognit. Psychol.* 11:1-37
- McDaniel, M. A., Friedman, A., Bourne, L. E. 1978. Remembering the levels of information in words. *Mem. Cognit.* 6:156-64
- Medin, D. L., Schaffer, M. M. 1978. Context theory of classification learning. *Psychol. Rev.* 85:207-38
- Meyer, D. E. 1970. On the representation and retrieval of stored semantic information. *Cognit. Psychol.* 1:242-99
- Miller, G. A. 1956. The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychol. Rev.* 63:81-97
- Milner, P. M. 1957. The cell assembly: Mark II. *Psychol. Rev.* 64:242-52
- Morris, C. D., Bransford, J. D., Franks, J. J. 1977. Levels of processing versus transfer appropriate processing. *J. Verb. Learn. Verb. Behav.* 16:519-33
- Morris, C. D., Stein, B. S., Bransford, J. D. 1979. Prerequisites for the utilization of knowledge in the recall of prose passages. *J. Exp. Psychol.-Hum. Learn. Mem.* 5:253-61
- Nelson, T. O. 1977. Repetition and depth of processing. *J. Verb. Learn. Verb. Behav.* 16:151-71
- Nelson, T. O., Greene, G., Ronk, B., Hatchett, G., Igl, V. 1978. Effect of multiple images on associative learning. *Mem. Cognit.* 6:337-41
- Nelson, T. O., Vining, S. K. 1978. Effect of semantic versus structural processing on long-term retention. *J. Exp. Psychol.-Hum. Learn. Mem.* 4:198-209
- Neumann, P. G. 1977. Visual prototype formation with discontinuous representation of variability. *Mem. Cognit.* 5:187-97
- Newell, A. 1973. Production systems: models of control structures. In *Visual Information Processing*, ed. W. G. Chase, pp. 463-526. New York: Academic. 555 pp.
- Nickerson, R. S., Adams, M. J. 1979. Long-term memory for a common object. *Cognit. Psychol.* 11:287-307
- Norman, D. A., Rumelhart, D. E., eds. 1975. *Explorations in Cognition*. San Francisco: Freeman. 430 pp.
- Oden, G. C. 1978. Semantic constraints and judged preference for interpretations of ambiguous sentences. *Mem. Cognit.* 6:26-37
- Oden, G. C. 1979. A fuzzy logical model of letter identification. *J. Exp. Psychol.: Hum. Percept. Perform.* 5:336-52
- Oden, G. C., Massaro, D. W. 1978. Integration of featural information in speech perception. *Psychol. Rev.* 85:172-91
- Owens, J., Bower, G. H., Black, J. B. 1979. The "soap opera" effect in story recall. *Mem. Cognit.* 7:185-91
- Piaget, J. 1954. *The Construction of Reality in the Child*. New York: Basic Books. 315 pp.
- Posner, M. I., Keele, S. W. 1968. On the genesis of abstract ideas. *J. Exp. Psychol.* 77:353-63
- Quillian, M. R. 1966. *Semantic memory*. PhD thesis. Carnegie-Mellon Univ., Pittsburgh, Pa. (Reprinted in part in *Semantic Information Processing*, ed. M. Minsky, pp. 227-70. Cambridge, Mass: MIT Press, 1968. 440 pp.)
- Reed, A. V. 1973. Speed-accuracy trade-off in recognition memory. *Science* 181:574-76
- Reed, A. V. 1976. List length and the time course of recognition in immediate memory. *Mem. Cognit.* 4:16-30
- Reed, A. V. 1977. Quantitative prediction of spacing effects in learning. *J. Verb. Learn. Verb. Behav.* 16:693-98
- Reed, S. K. 1972. Pattern recognition and categorization. *Cognit. Psychol.* 3:382-407

- Remington, R. 1977. Processing of phonemes in speech: A speed-accuracy study. *J. Acoust. Soc. Am.* 62:1279-90
- Rips, L. J., Shoben, E. J., Smith, E. E. 1973. Semantic distance and the verification of semantic relations. *J. Verb. Learn. Verb. Behav.* 12:1-20
- Rosch, E. H. 1973. On the internal structure of perceptual and semantic categories. In *Cognitive Development and the Acquisition of Language*, ed. T. E. Moore, pp. 111-44. New York: Academic. 308 pp.
- Rosch, E. H. 1977. Human categorization. In *Advances in Cross-Cultural Psychology*, ed. N. Warren, 1:1-49. New York: Academic. 212 pp.
- Rosch, E. H., Mervis, C. B. 1975. Family resemblances: Studies in the internal structure of categories. *Cognit. Psychol.* 7:573-605
- Rosch, E. H., Mervis, C. B., Gary, W. D., Johnson, D. M., Boyes-Braem, P. 1976. Basic objects in natural categories. *Cognit. Psychol.* 8:382-439
- Ross, B. H., Landauer, T. K. 1978. Memory for at least one of two items: test and failure of several theories of spacing effects. *J. Verb. Learn. Verb. Behav.* 17:669-80
- Rumelhart, D. E., Lindsay, P. H., Norman, D. A. 1972. In *Organization of Memory*, ed. E. Tulving, W. Donaldson, pp. 197-246. New York: Academic. 423 pp.
- Schaeffer, B., Wallace, R. 1969. Semantic similarity and the comparison of word meanings. *J. Exp. Psychol.* 82:343-46
- Schaeffer, B., Wallace, R. 1970. The comparison of word meanings. *J. Exp. Psychol.* 86:144-52
- Schulman, A. I. 1971. Recognition memory for targets from a scanned word list. *Br. J. Psychol.* 62:335-46
- Singer, M. 1976. Thematic structure and the integration of linguistic information. *J. Verb. Learn. Verb. Behav.* 15:549-58
- Smith, E. E. 1978. Theories of semantic memory. In *Handbook of Learning and Cognitive Processes*, ed. W. K. Estes, 6:1-56. Hillsdale, NJ: Erlbaum. 311 pp.
- Smith, E. E., Shoben, E. J., Rips, L. J. 1974. Structure and process in semantic memory: A featural model for semantic decisions. *Psychol. Rev.* 81:214-41
- Stein, B. S. 1978. Depth of processing reexamined: the effects of the precision of encoding and test appropriateness. *J. Verb. Learn. Verb. Behav.* 17:165-74
- Stein, B. S., Morris, C. D., Bransford, J. D. 1978. Constraints on effective elaboration. *J. Verb. Learn. Verb. Behav.* 17:707-14
- Strauss, M. S. 1979. Abstraction of prototypical information by adults and 10-month-old infants. *J. Exp. Psychol.-Hum. Learn. Mem.* 5:618-32
- Turvey, M. 1973. On peripheral and central processes in vision: inferences from an information-processing analysis of masking with patterned stimuli. *Psychol. Rev.* 80:1-52
- Tyler, S. W., Hertel, P. T., McCallum, M. C., Ellis, H. C. 1979. Cognitive effort and memory. *J. Exp. Psychol.-Hum. Learn. Mem.* 5:607-17
- Wason, P. C., Johnson-Laird, P. N. 1972. *Psychology of Reasoning: Structure and Content*, Cambridge, Mass: Harvard. 264 pp.
- Wickelgren, W. A. 1965. Short-term memory for repeated and non-repeated items. *Q. J. Exp. Psychol.* 17:14-25
- Wickelgren, W. A. 1969a. Learned specification of concept neurons. *Bull. Math. Biophys.* 31:123-42
- Wickelgren, W. A. 1969b. Context-sensitive coding, associative memory, and serial order in (speech) behavior. *Psychol. Rev.* 76:1-15
- Wickelgren, W. A. 1969c. Context-sensitive coding in speech recognition, articulation, and development. In *Information Processing in the Nervous System*, ed. K. N. Leibovic, pp. 85-96. New York: Springer. 315 pp.
- Wickelgren, W. A. 1972a. Coding, retrieval, and dynamics of multitrace associative memory. In *Cognition in Learning and Memory*, ed. L. W. Gregg, pp. 19-50. New York: Wiley. 315 pp.
- Wickelgren, W. A. 1972b. Context-sensitive coding and serial vs. parallel processing in speech. In *Speech and Cortical Functioning*, ed. J. H. Gilbert, pp. 237-62. New York: Academic. 272 pp.
- Wickelgren, W. A. 1972c. Trace resistance and the decay of long-term memory. *J. Math. Psychol.* 9:418-55
- Wickelgren, W. A. 1973. The long and the short of memory. *Psychol. Bull.* 80:425-38
- Wickelgren, W. A. 1974. Single-trace fragility theory of memory dynamics. *Mem. Cognit.* 2:775-80
- Wickelgren, W. A. 1975. More on the long and short of memory. In *Short-Term Memory*, ed. D. Deutsch, J. A. Deutsch, pp. 65-72. New York: Academic. 411 pp.
- Wickelgren, W. A. 1976a. Network strength theory of storage and retrieval dynamics. *Psychol. Rev.* 83:466-78
- Wickelgren, W. A. 1976b. Subproblems of semantic memory: A review of *Human*

- Associative Memory* by J. R. Anderson, G. H. Bower. *J. Math. Psychol.* 13:243-68
- Wickelgren, W. A. 1976c. Phonetic coding and serial order. In *Handbook of Perception*, ed. E. C. Carterette, M. P. Freidman, 7:227-64. New York: Academic. 501 pp.
- Wickelgren, W. A. 1977a. *Learning and Memory*. Englewood Cliffs, NJ: Prentice-Hall. 448 pp.
- Wickelgren, W. A. 1977b. Concept neurons: A proposed developmental study. *Bull. Psychon. Soc.* 10:232-34
- Wickelgren, W. A. 1977c. Speed-accuracy tradeoff and information processing dynamics. *Acta Psychol.* 41:67-85
- Wickelgren, W. A. 1978. Wickelgren's neglect. *Acta Psychol.* 42:81-82
- Wickelgren, W. A. 1979a. *Cognitive Psychology*. Englewood Cliffs, NJ: Prentice-Hall. 436 pp.
- Wickelgren, W. A. 1979b. Chunking and consolidation: A theoretical synthesis of semantic networks, configuring in conditioning, S-R vs Cognitive Learning, normal forgetting, the amnesic syndrome, and the hippocampal arousal system. *Psychol. Rev.* 86:44-60
- Wickelgren, W. A. 1979c. I liked the postcard you sent Abe and I: Context-sensitive coding of syntax and other procedural knowledge. *Bull. Psychon. Soc.* 13:61-63
- Wickelgren, W. A., Corbett, A. T. 1977. Associative interference and retrieval dynamics in yes-no recall and recognition. *J. Exp. Psychol.-Hum. Learn. Mem.* 3:189-202
- Wickelgren, W. A., Corbett, A. T., Doshier, B. A. 1980. Priming and retrieval from short-term memory: a speed accuracy tradeoff analysis. *J. Verb. Learn. Verb. Behav.* In press
- Wickelgren, W. A., Norman, D. A. 1971. Invariance of forgetting rate with number of repetitions in verbal short-term recognition memory. *Psychon. Sci.* 22:363-64
- Wittgenstein, L. 1953. *Philosophical Investigations*. New York: Macmillan
- Woods, W. A. 1970. Transition network grammars for natural language analysis. *Commun. ACM* 13:591-606
- Woods, W. A. 1975. What's in a link: Foundations for semantic networks. In *Representation and Understanding*, ed. D. G. Bobrow, A. Collins, pp. 35-82. New York: Academic. 427 pp.
- Zadeh, L. A. 1965. Fuzzy sets. *Inf. Control.* 8:338-53



CONTENTS

MORE STATELY MANSIONS—PSYCHOLOGY EXTENDS ITS BOUNDARIES, <i>Leona E. Tyler</i>	1
HUMAN LEARNING AND MEMORY, <i>Wayne A. Wickelgren</i>	21
BEHAVIORAL DECISION THEORY: PROCESSES OF JUDGMENT AND CHOICE, <i>Hillel J. Einhorn and Robin M. Hogarth</i>	53
CATEGORIZATION OF NATURAL OBJECTS, <i>Carolyn B. Mervis and Eleanor Rosch</i>	89
DEVELOPMENTAL PSYCHOLOGY, <i>John C. Masters</i>	117
COCHLEAR PHYSIOLOGY, <i>Peter Dallos</i>	153
GRAPHICAL DATA ANALYSIS, <i>Howard Wainer and David Thissen</i>	191
CHILD PSYCHOPATHOLOGY, <i>Alan O. Ross and William E. Pelham</i>	243
COUNSELING PSYCHOLOGY: CAREER INTERVENTIONS, RESEARCH, AND THEORY, <i>John L. Holland, Thomas M. Magoon, and Arnold R. Spokane</i>	279
MASS COMMUNICATION EFFECTS, <i>Donald F. Roberts and Christine M. Bachen</i>	307
ATTITUDE AND ATTITUDE CHANGE, <i>Robert B. Cialdini, Richard E. Petty, and John T. Cacioppo</i>	357
CLINICAL PSYCHOLOGY: INDIVIDUAL METHODS, <i>Jeanne S. Phillips and Karen Linn Bierman</i>	405
CREATIVITY, INTELLIGENCE, AND PERSONALITY, <i>Frank Barron and David M. Harrington</i>	439
VISUAL NEURAL DEVELOPMENT, <i>J. Anthony Movshon and Richard C. Van Sluyters</i>	477
THE CENTRAL BASIS OF MOTIVATION: INTRACRANIAL SELF-STIMULATION STUDIES, <i>M. E. Olds and J. L. Fobes</i>	523
EARLY VISUAL PERCEPTION, <i>Bela Julesz and Robert A. Schumer</i>	575
TEST THEORY AND METHODS, <i>David J. Weiss and Mark L. Davison</i>	629
INSTRUCTIONAL PSYCHOLOGY, <i>Lauren B. Resnick</i>	659
CHAPTERS PLANNED FOR VOLUME 33 (1982)	705
INDEXES	
Author Index	707
Subject Index	733
Cumulative Index of Contributing Authors, Volumes 28 to 32	751
Cumulative Index of Chapter Titles, Volumes 28 to 32	753