THE BEHAVIORAL AND BRAIN SCIENCES (1984) 7, 223–268
Printed in the United States of America

# Précis of *Elements of episodic memory*

**Endel Tulving**
*Department of Psychology, University of Toronto, Toronto, Ontario, Canada M5S 1A1*

**Abstract:** *Elements of episodic memory* (Tulving 1983b) consists of three parts. Part I argues for the distinction between episodic and semantic memory as functionally separate albeit closely interacting systems. It begins with a review of the 1972 essay on the topic (Tulving 1972) and its shortcomings, presents a somewhat more complete characterization of the two forms of memory than the one that was possible in 1972, and proceeds to discuss empirical and theoretical reasons for a tentative acceptance of the functional distinction between the two systems and its possible extensions. Part II describes a framework for the study of episodic memory, dubbed General Abstract Processing System (GAPS). The basic unit in such study is an act of remembering. It begins with the witnessing of an event and ends with recollective experience of the event, with related memory performance, or both. The framework specifies a number of components (elements) of the act of remembering and their interrelations, classified under two broad categories of encoding and retrieval. Part III discusses experimental research under the label of "synergistic ecphory." Ecphory is one of the central elements of retrieval; "synergistic" refers to the joint influence that the stored episodic information and the cognitively present retrieval information exert on the construction of the product of ecphory, the so-called ecphoric information. The concept of encoding specificity and the phenomenon of recognition failure of recallable words figure prominently in Part III. The final chapter of the book describes a model, named the synergistic ecphory model of retrieval, that relates qualitative characteristics of recollective experience and quantitative measures of memory performance in recall and recognition to the conjunction of episodic-memory traces and semantic-memory retrieval cues.

**Keywords:** amnesia; encoding; episodic memory; knowledge; memory; recall; recognition; recollection; retrieval; semantic memory

## Part I: The episodic/semantic distinction[1]

### *Inchoate distinction*

I wrote my 1972 chapter in reaction to papers by Rumelhart, Lindsay, and Norman (1972), Kintsch (1972), and Collins and Quillian (1972), that had been given at a conference at the University of Pittsburgh in March 1971. These authors were concerned with what I thought were the processes involved in the understanding of language, whereas they suggested that they were studying memory in a broader sense than had been the case in the past. They all used the term "semantic memory" to describe their work, borrowing the term from Quillian (1966). I thought that the extension of the concept of memory to comprehension of language, question answering, making of inferences, and other such cognitive skills was inappropriate. Inspired by Bergson (1911), Reiff and Scheerer (1959), and Munsat (1966), as well as by others who had discussed similar issues, I wrote the essay on the distinction between episodic and semantic memory.

Episodic memory, I suggested, is a system that receives and stores information about temporally dated episodes or events, and temporal–spatial relations among them. Semantic memory, on the other hand, "is the memory necessary for the use of language. It is a mental thesaurus, organized knowledge a person possesses about words and other verbal symbols, their meaning and referents, about relations among them, and about rules, formulas, and algorithms for the manipulation of the symbols, concepts, and relations" (Tulving 1972, p. 386).

I contrasted the two forms of memory with respect to five issues: (a) the nature of stored information; (b) autobiographical versus cognitive referents; (c) conditions and consequences of retrieval; (d) vulnerability to interference; and (e) interdependence of the two kinds of memory. I assumed that the two forms of memory were interdependent, interacting closely most of the time, each influencing the other in many situations. But I also thought that such interdependence was optional rather than obligatory: it was possible for a person to acquire knowledge about a particular dated cooccurrence of novel and meaningless stimulus events; similarly, it seemed reasonable to assume that mere cooccurrence of two stimuli or language units would not change the structure of semantic memory.

These different conceptualizations of a person's knowledge of an A–B "association" corresponded to the distinction between recollection of events and recall of facts, discussed in textbooks of memory (e.g., Boring, Langfeld & Weld 1948). But it deviated from the commonly accepted assumption that the learning of an A–B association, as in a paired-associate list, essentially consists of strengthening or updating the existing association between the two items of the kind revealed by free-association tests. Donald Thomson and I questioned the validity of this assumption, on the basis of experiments showing effects of context changes on recall and recognition of studied words (e.g., Thomson & Tulving 1970; Tulving & Thom-

 0140-525X/84/020223-46/$06.00

son 1971). We thought that word-recall experiments were concerned with subjects' remembering of events rather than with the establishment of new associations, or the strengthening of old ones, between transsituationally invariant units of semantic memory. The distinction between episodic and semantic memory naturally fitted into this picture.

The 1972 distinction was inchoate: rudimentary, imperfect, incomplete, and somewhat disorganized. The summary statement of the essay was correspondingly cautious. What I had done in the essay was to "present a case for the possible heuristic usefulness of a taxonomic distinction between episodic and semantic memory as two parallel and partially overlapping information processing systems" (Tulving 1972, p. 401). With the wisdom of hindsight it is easy to see the weaknesses and shortcomings of this distinction. Lack of relevant empirical evidence was the major problem, but there were others. One of these had to do with the absence of emphasis on the similarities of the two systems; another concerned the implied exhaustiveness of the primitive taxonomy implicit in the distinction; the third difficulty had to do with the names of the two systems; and perhaps the most serious problem involved the lack of clear and definite ideas regarding the relation between autobiographical episodes and what we might call their "contents." I had somewhat hastily classified a large majority of laboratory experiments on memory that had been done up to that time as experiments on episodic memory: In these experiments subjects were tested for their knowledge of what they had seen or heard at a particular time in a particular situation. It seemed self-evident that a subject would have to remember the event of seeing a particular word in a list in order to be able to recall the word when given episodic instructions. In light of both old and new relevant data, identifying the recall of the contents of an event with the remembering of the event appears to have been unwarranted, however.

### Argument for differences

Episodic and semantic memory are two systems of *propositional* memory: Their function is to acquire, retain, and make available information that represents the reality external to the organism, information that can be expressed in the form of propositions. They can be contrasted with memory systems concerned with the acquisition and utilization of skills and procedures, or systems of procedural memory (cf. Winograd 1975b). Propositional and procedural memory systems differ in several ways: (a) Information handled by propositional memory systems has truth value, whereas that handled by procedural systems does not; (b) information retrieved from propositional memory can be contemplated introspectively or attended to internally, whereas procedural knowledge cannot; (c) propositional knowledge about something can be communicated to others in different ways through language or some other symbol system, whereas knowledge of a particular procedure can only be demonstrated through highly specific behavior; (d) propositional knowledge about something can frequently be acquired in a single act of perception or thought, whereas acquisition of skills and procedures usually requires intensive practice.

Thus, episodic and semantic memories are similar in that both are subdivisions of propositional memory. Despite this similarity, casual observations reveal a number of differences between episodic and semantic memory. These differences can be classified under three broad categories: differences in information, differences in operations, and differences in "applications," or the role that memory plays in a broad range of human affairs. A summary of these differences is given in Table 1.

Table 1. *Summary of differences between episodic and semantic memory*

| Diagnostic feature | Episodic | Semantic |
|---|---|---|
| *Information* | | |
| Source | Sensation | Comprehension |
| Units | Events; episodes | Facts; ideas; concepts |
| Organization | Temporal | Conceptual |
| Reference | Self | Universe |
| Veridicality | Personal belief | Social agreement |
| *Operations* | | |
| Registration | Experiential | Symbolic |
| Temporal coding | Present; direct | Absent; indirect |
| Affect | More important | Less important |
| Inferential capability | Limited | Rich |
| Context dependency | More pronounced | Less pronounced |
| Vulnerability | Great | Small |
| Access | Deliberate | Automatic |
| Retrieval queries | Time? Place? | What? |
| Retrieval consequences | Change system | System unchanged |
| Retrieval mechanisms | Synergy | Unfolding |
| Recollective experience | Remembered past | Actualized knowledge |
| Retrieval report | Remember | Know |
| Developmental sequence | Late | Early |
| Childhood amnesia | Affected | Unaffected |
| *Applications* | | |
| Education | Irrelevant | Relevant |
| General utility | Less useful | More useful |
| Artificial intelligence | Questionable | Excellent |
| Human intelligence | Unrelated | Related |
| Empirical evidence | Forgetting | Analysis of language |
| Laboratory tasks | Particular episodes | General knowledge |
| Legal testimony | Admissible; eyewitness | Inadmissible; expert |
| Amnesia | Involved | Not involved |
| Bicameral men | No | Yes |

*Source*: Tulving 1983b, Table 3.1, p. 35.

**Differences in information.** The two systems differ in the immediate source of the information they handle. The mere sensation of a stimulus can serve as a source of information in the episodic system, whereas comprehension is necessary for the semantic system. The prototypical unit of information in episodic memory is an event or an episode. In semantic memory, there is no single "basic" unit, but facts, ideas, concepts, rules, propositions, schemata, scripts, and other related terms have been used by philosophers, psychologists, and cognitive scientists in discussing the nature of people's knowledge of the world. Organization of knowledge in the episodic system is temporal: One event precedes, cooccurs, or succeeds another in time. Lockhart, Craik, and Jacoby (1976) have even argued that "episodic memory has *no* inherent structure" (p. 82). The organization of knowledge in the semantic system, on the other hand, is defined by many relations that could be classified as "conceptual." The temporal organization of the episodic system is relatively loose, whereas the conceptual organization of semantic memory is tight (e.g., Estes 1976).

The information in the episodic system refers to or represents events in the rememberer's personal past, and may thereby provide a basis for defining an individual's personal identity (e.g., Greenwald 1981; Grice 1941; Shoemaker 1959). The knowledge recorded in the semantic system is timeless: It has no necessary connection to the knower's personal identity and instead refers to the world. Finally, the rememberer's belief in the veridicality of the remembered event is an inherent feature of episodic remembering and independent of testimony of others, whereas the belief in the veridicality of semantic knowledge is supported by social consensus.

**Differences in operations.** The episodic system registers immediate experiences, the semantic system registers knowledge conveyed by referential events and language. Only the episodic system can keep track of temporal order of occurrence of personal events; the semantic system has no capability of *direct* recording and maintenance of such information, although it can solve problems of the temporal order of events by inferences. The episodic system is relatively limited in inferential capability, whereas the semantic system possesses a rich inferential capability. Affect probably plays a more important role in the recording and retrieval of information in episodic than in semantic memory.

It is generally thought that the operation of the episodic system is more context-dependent than the operation of the semantic system (e.g., Ehrlich 1979; Kintsch 1980). Yet the question of whether episodic and semantic systems can be differentiated in terms of context dependency is a complex one: It is quite possible that the acquisition and utilization of our knowledge of the world is as context-dependent as is our episodic knowledge.

Information in the episodic system is more vulnerable to interference than that in the semantic system. The actualization of episodic information tends to be deliberate, frequently requiring conscious effort; that of the semantic system tends to be automatic. The general form of the retrieval query directed at the episodic system is, "What did you do at time T in place P?" In the semantic system it is, "What is X?" where X refers to an object, a situation, a property or characteristic, a relation, and so on. Retrieval from the episodic system tends to change (recode) the stored information; retrieval of information from semantic memory usually leaves its contents unchanged. Retrieval from the episodic system takes the form of a synergistic combination of the information stored in the episodic system and the information provided by the cognitive environment of the rememberer, interpreted in terms of the person's semantic knowledge. In semantic memory, on the other hand, retrieval entails a process in which the dispositional knowledge is actualized, or in which it "unfolds," in a manner determined by the nature and organization of the stored knowledge and relatively independently of the nature of the instigating cue.

Recollective episodic experiences are interpreted by rememberers as being a part of their personal past, whereas actualized semantic knowledge represents the impersonal present. People use the word "remember" when referring to personal recollections, and the word "know" when talking about actualized semantic knowledge.

Although some writers have suggested that semantic memory develops "out of" episodic memory (e.g., Anglin 1977; Kintsch 1974), a more plausible argument is that, in the development of a child, semantic memory precedes episodic memory (e.g., Kinsbourne & Wood 1975; Schachtel 1947). A related speculation holds that childhood amnesia is a phenomenon of episodic (autobiographical) rather than semantic memory (Schachtel 1947).

**Differences in applications.** Formal education is aimed at the acquisition, retention, and utilization of skills and knowledge that have to do with the world; episodic memory is irrelevant to the accomplishment of these aims. The general utility of semantic knowledge for an individual is greater than is the remembering of personal events.

The prospects of endowing computers with episodic memories that faithfully mimic their human counterparts are decidedly less favourable than the prospects of making computers efficient language users, question answerers, inference makers, or problem solvers (cf. Schank & Kolodner 1979). In definitions and assessment of human intelligence, semantic memory occupies a central position, whereas episodic memory is unrelated to intelligence (e.g., Sternberg & Detterman 1979). The relevance of the distinction between episodic and semantic memory to legal testimony can be expressed by saying that for the testimony of eyewitnesses to be acceptable, it must be based on episodic memory, whereas for that of expert witnesses to be admissible, it must be based on semantic memory.

In the study of memory, the phenomenon of forgetting – discrepancy between input and output – defines the basic focus of interest in episodic memory; forgetting is of no interest to students of semantic memory. Most of the work on semantic memory has to do with people's knowledge of language (e.g., Anderson & Bower 1973; Lachman, Schaffer & Hennrikus 1974; Meyer 1973; Miller 1969; Rubenstein, Garfield & Millikan 1970; Smith 1978; Smith, Shoben & Rips 1974); episodic memory research need not involve language. In the laboratory, episodic-memory tasks require retention of information from a particular episode, whereas performance on semantic tasks is guided by general knowledge.

A number of writers (e.g., Kinsbourne & Wood 1975; Rozin 1976; Wood, Ebert & Kinsbourne 1982) have suggested that amnesia resulting from brain damage is a condition in which episodic memory is selectively impaired while semantic memory is less affected.

Finally, Jaynes's (1976) theory of the evolution of consciousness implies that although bicameral men had perfectly developed semantic-memory capacities they were deficient in episodic memory: They "could not reminisce because they were not fully conscious" (Jaynes 1976, p. 371).

### *Debate about memory*

**Agreements.** All students of memory seem to be willing to accept the distinction between episodic and semantic memory as a purely heuristic device that helps us to classify and describe experiments and observations. The heuristic use of the terms "episodic" and "semantic" aids communication and serves as a first step to deeper questions. We can describe different memory tasks as either episodic or semantic, and we can interpret, categorize, and organize outcomes of certain older experiments in terms of the difference between episodic and semantic memory without accepting the idea that the two represent different systems (e.g., Drachman & Leavitt 1972; Penfield & Perot 1963; Slamecka 1966). More recently, many researchers have related their own findings and observations to the episodic/semantic distinction in at least the heuristic sense (e.g., Caine, Ebert & Weingartner 1977; Gilhooly & Gilhooly 1979; Herrmann & McLaughlin 1973; Moeser 1976; 1977; Ojemann 1978; Petrey 1977; Russell & Beekhuis 1976; Underwood, Boruch & Malmi 1978).

There is also good agreement among theorists that "episodic" and "semantic" refer to different kinds of information. Anderson and Ross (1980), for instance, who reject the functional distinction between episodic and semantic memories, have no objection to the corresponding "content distinction" (p. 463). Similarly, there should be no disagreement regarding the separation between remembered episodes and their "semantic contents," and the possibility of answering questions directed at episodic memory on the basis of our general knowledge of the world. Finally, virtually everyone agrees that episodic and semantic memories are not only similar in many ways but also interact closely almost all the time.

**Open questions.** There are several identifiable matters on which disagreement does seem to exist at the present time. Probably the most basic issue of this kind concerns the problem of whether episodic and semantic memories represent different functional systems.

The position advocated in the book is that episodic and semantic memory are functionally distinct. This statement does not mean that the systems are completely separate, that they have nothing to do with one another, that there are no similarities between them, or that they serve completely nonoverlapping functions. It does mean that one system *can* operate independently of the other, although not necessarily as efficiently as it could with the support of the other intact system. It also means that the operations of one system could be globally enhanced without a similar effect on the operations of the other, and that the activity of one system could be suppressed without a comparable effect on the activity of the other. The functional difference also implies that in important ways the systems operate differently, that is, that their function is governed at least partially by different principles.

Some theorists (e.g., Craik 1979b; Jacoby & Craik 1979; Kintsch 1980; Naus & Halasz 1979) have argued that episodic and semantic memories constitute a continuum of some kind. Craik (1979b), for instance, has suggested that "the implied break between two memory systems is unsatisfactory," and that a better solution is the idea of "a continuum of representation, running from highly context-specific episodes at one extreme to abstract generalized knowledge at the other" (p. 451). In light of the currently available evidence, the idea of a continuum is not appealing.

Another open question concerns the status of lexical memory. Although many writers think of it as a part of semantic memory, or at least as a form of propositional memory (e.g., Collins & Loftus 1975; Kintsch 1980; Lachman 1973; Miller 1969; 1972; Schank 1975), and although the idea makes intuitive sense, it is possible to contemplate the hypothesis that lexical memory is a form of procedural memory that serves the function of transmitting and expressing episodic and semantic knowledge.

### *Empirical evidence*

On the basis of findings of transfer from an episodic to a semantic task with reaction time as the dependent variable, Anderson and Ross (1980) argued against a functional basis for the distinction between episodic and semantic memory. However, since other experiments (e.g., Jacoby & Witherspoon 1982; Tulving, Schacter & Stark 1982) have shown that what is transferred from episodic input to semantic retrieval can be uncorrelated with episodic information, Anderson and Ross's findings cannot be regarded as highly relevant to the episodic/semantic distinction.

Herrmann and Harwood (1980) obtained data that they thought supported the distinction between the two systems, and so did Shoben, Wescourt, and Smith (1978). The latter study was based on the logic of double experimental dissociation. In the semantic task, subjects verified the truth of sentences, whereas in the episodic task they made recognition judgments about the same sentences. Independent variables were semantic relatedness among the sentences and "fanning," defined in terms of the number of propositions learned about a concept. The results showed double dissociation: Semantic verification was influenced by semantic relatedness but not by fanning, whereas recognition was influenced by fanning but not by relatedness.

Single experimental dissociations between episodic and semantic tasks have been demonstrated by McKoon and Ratcliff (1979), Jacoby and Dallas (1981), and Kihlstrom (1980). In McKoon and Ratcliff's paper the relevant data were provided by Experiments 1 and 4 in which response latencies were measured in a (semantic) lexical-decision task and an (episodic) recognition task, as a

Table 2. *Response latencies (s) in McKoon and Ratcliff* (1979)

| | Relation of priming word to target | | |
|---|---|---|---|
| Task | Episodic and semantic | Episodic | Semantic |
| Semantic Lexical decision | 0.53 | 0.54 | 0.53 |
| Episodic Recognition | 0.57 | 0.62 | 0.74 |

*Source:* Tulving 1983b, Table 5.4, p. 88.

function of the relation between the target word and its preceding word. Their results are summarized in Table 2. These data show that the manipulation of the relation between the target and the preceding word in the series had no effect on lexical decision, but a sizable effect on the episodic-recognition task.

Jacoby and Dallas (1981, Exp. 1) compared subjects' performance in a semantic task (tachistoscopic identification of words) with that in an episodic task (recognition of previously studied words) as a function of the encoding operations performed on the target items in the first phase of the experiment. The results, expressed in terms of the probability of correct responses, are summarized in Table 3. These results show a clear dissociation between episodic and semantic tasks.

Kihlstrom (1980, Exp. 1) measured episodic free recall and semantic free association in different groups of subjects varying in hypnotic ability, in a situation in which subjects had learned the words and then were given posthypnotic suggestions to forget them. His data, summarized in Table 4, show that the effectiveness of posthypnotic amnesia suggestions varied greatly with the hypnotic ability of subjects in the episodic task but not at all in the semantic task. These data thus demonstrate a dissociation between episodic and semantic tasks in an experimental situation in which the independent variable was defined in terms of differences in brain states induced by hypnotic suggestions.

Wood, Taylor, Penny, and Stump (1980), in a well-controlled experiment, observed differences in regional cerebral blood flow between two groups of subjects, one engaging in an episodic, the other in a semantic memory task. They interpreted these results as suggesting "an anatomical basis for the distinction between episodic and semantic memory" (p. 113).

Table 3. *Response probabilities in Jacoby and Dallas* (1981, Exp. 1)

| | First-phase encoding condition | | |
|---|---|---|---|
| Task | Appearance | Sound | Meaning |
| Semantic Identification | 0.80 | 0.81 | 0.82 |
| Episodic Recognition | 0.50 | 0.63 | 0.86 |

*Source*: Tulving 1983b, Table 5.5, p. 89.

Table 4. *Response probabilities in Kihlstrom* (1980)

| | Hypnotizability of subjects | | |
|---|---|---|---|
| Task | Very high | High | Medium and low |
| Semantic Free association | 0.61 | 0.50 | 0.53 |
| Episodic Free recall | 0.01 | 0.47 | 0.86 |

*Source*: Tulving 1983b, Table 5.6; p. 90.

**Pathological dissociations.** Pathological dissociations supporting the episodic/semantic distinction have been discussed more fully elsewhere (Schacter & Tulving 1982). A few examples are mentioned here.

Warrington and Weiskrantz (1974) compared (episodic) Yes/No recognition performance with (semantic) word-fragment completion performance in four amnesic patients and four control subjects. They found that the control subjects' recognition-memory performance was much better than that of amnesics, whereas the two groups did not differ in the word-fragment completion performance. More recent evidence reported by Warrington and Weiskrantz (1982, Exp. 1) also points to a dissociation between episodic and semantic tasks when amnesic patients are compared with control patients.

Dissociations of episodic and semantic memory are found in many clinical descriptions of the amnesic syndrome (e.g., Claparède 1911; Williams & Smith 1954). A dissociation between episodic and semantic memory tasks has also been described in an experiment with a single patient who was suffering from a temporary functional amnesia (Schacter, Wang, Tulving & Freedman 1982): During the amnesic episode, the patient had great difficulty remembering events from his life, but no difficulty in identifying well-known people from their photographs (Albert, Butters & Levine 1979).

**Evaluation of the evidence.** The evidence reviewed shows that dissociations between episodic and semantic tasks have been observed in both laboratory experiments and clinical settings, with data provided by normal subjects, hypnotized subjects, and brain-damaged patients, as well as by functional amnesia patients. Semantic memory in these observations was tapped by a number of different tasks: sentence verification, lexical decision, tachistoscopic identification, word-fragment completion, free association, naming of category instances, production of opposites; both recall and recognition served as tasks of episodic memory.

The hypothesis of a functional distinction between episodic- and semantic-memory systems provides an economical explanation of the finding of the same pattern of performance – dissociation of tasks – in the face of a great deal of situational diversity: The manipulated variables, or different subject groups, produce differences in performance in episodic and semantic tasks, because the tasks tap different memory systems. In the absence of such an overall explanation, a large number of different, unique explanations would have to be provided for the results of different experiments.

***Extensions and contrasts***

If we accept the hypothesis that the two types of memory represent functionally distinct systems, we can proceed with the study of similarities and differences between the systems, possible extensions of the taxonomy of memory systems, and contrasts with other systems.

**Priming effects.** A persistent finding in experiments demonstrating experimental dissociations between episodic and semantic tasks, experiments we have just reviewed, was that of priming in semantic tasks. In experiments by Jacoby and Dallas (1981), McKoon and Ratcliff (1979), Kihlstrom (1980), and Tulving et al. (1982) – as well as in other related experiments (Morton 1979; Williamsen, Johnson & Eriksen 1965; Winnick & Daniel 1979) – performance in the semantic task, although not influenced by manipulated variables, was enhanced by virtue of subjects' prior experimental exposure to the target words. Thus, the complete results of these experiments can be schematically depicted as in Figure 1: A manipulated variable has an effect on the episodic task, no effect on the semantic task, and there is a priming effect, independent of the manipulated variable, in the semantic task.

No good explanations of priming are available as yet. Jacoby and Witherspoon (1982) and Tulving et al. (1982) have shown that priming in semantic tasks is uncorrelated with performance on episodic recognition tasks: Tachistoscopic identification and word-fragment completion were found to be indistinguishable for words that subjects thought they had seen before and words they thought they had not. These and other related findings (e.g., Kihlstrom 1980; Williamsen, Johnson & Eriksen 1965) seem to imply that priming effects are mediated by, and reflect the operations of, a system other than episodic memory.

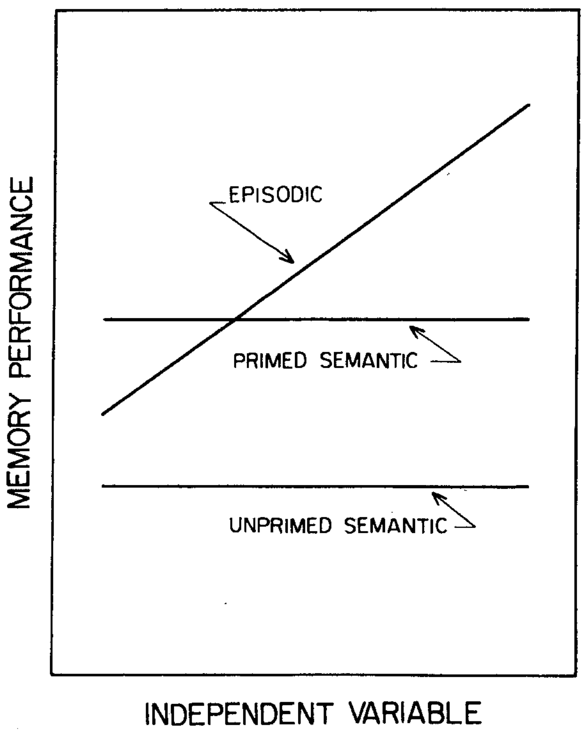


Figure 1 (Tulving 1983b, fig. 6.2, p. 106). A schematic pattern of data showing the effect of a manipulated variable in an episodic task, no effect of the same variable in a semantic task, and the priming effect, independent of the manipulated variable, in the semantic task.

Do priming effects reflect changes in the semantic-memory system? Since priming effects are defined in terms of changes in performance on semantic tasks, it would seem natural to answer the question in the affirmative. Certain facts, however, suggest that the answer may be more complicated. First, priming effects in semantic memory can be long-lived: Tulving et al. (1982), for instance, observed virtually no reduction in priming effects over an interval of seven days. The second fact has to do with the absence, or at least severe attenuation, of cross-modality priming effects (Jacoby & Dallas 1981; Morton 1979; Winnick & Daniel 1970): For priming to be optimal or to occur at all, the initial presentation of the target item has to be in the same sensory modality in which the item appears in a subsequent task. These two facts rule out *temporary* activation of modality-free semantic structures as responsible for the priming effects.

A third hypothesis is that priming reflects an improvement in the facility with which cognitive operations can be carried out, that is, that priming is a phenomenon of procedural memory. We know that many cognitive skills can be improved with practice (e.g., Cohen & Squire 1980; Kolers 1976b; Neisser, Novick & Lazar 1963; Peterson 1969), and priming effects may reflect nothing more than such improvement. The major difficulty with this hypothesis lies in the specificity of priming effects: Improvement in facilitation defined as priming occurs at the level of individual words or other small cognitive units. It has been customary to think of acquired skills in terms of their applicability to a wide variety of situations.

Priming effects that cannot be readily interpreted as reflecting changes in episodic, semantic, or procedural memory suggest the need for a modification of existing distinctions, or perhaps for an extension of memory taxonomy. Such a need is also hinted at by the existence of "free radicals" in memory.

**Free radicals.** Free fragments, discussed by Schacter and Tulving (1982), or free radicals, are bits of symbolic knowledge originally constructed as a part of the trace of an experienced episode that have become detached from episodic memory but have not, or not yet, been attached to any structure in semantic memory. Clinical descriptions of amnesia contain frequent references to patients' fragmentary knowledge of their recent experiences, in the absence of any awareness of the source of such knowledge (e.g., Claparède 1911; Luria 1976; Williams & Smith 1954). Evans and Thorn (1966) have described similar "source amnesia" for information acquired under hypnosis. The Warrington–Weiskrantz (1974; 1978) effect, discussed earlier as an example of pathological dissociation between episodic and semantic memory, could also be interpreted with the aid of the concept of free radicals.

**Working memory and reference memory.** The distinction made by Olton and his associates (e.g., Olton, Becker & Handelmann 1979; 1980; Olton & Papas 1979) between working memory and reference memory, based on work with animals, represents an interesting parallel to the episodic/semantic distinction. Working memory reflects an animal's knowledge of particular events in its recent past, whereas reference memory has to do with the animal's knowledge of relatively more permanent components of its world.

Olton and his associates have shown that bilateral destruction of the external connections of the hippocampus produces a permanent impairment in the working-memory component but not in the reference-memory component, reminiscent of similar dissociations between episodic and semantic memory performances in amnesic patients.

## Part II: General abstract processing system

The second part of the book describes a conceptual framework for the study and understanding of episodic memory. The framework is referred to as a General Abstract Processing System (GAPS) of episodic memory. It is *general* in that it is meant to apply to remembering of events of all sorts; it is *abstract* in that the specific nature of its components is not specified; it is a *processing* system since its major components have to do with the activity and the functioning of the system rather than its structure; and it is a *system* in the sense of an ordered and reasonably comprehensive collection of interacting components whose assemblage constitutes an integrated whole.

### *Elements of episodic memory*

The basic unit of the conceptual analysis of episodic memory is an act of remembering that begins with an event, perceived by the rememberer, and ends with recollective experience, the rememberer's private awareness of the event on a subsequent occasion, or with memory performance, the overt expression of the recollective experience.

GAPS can be described in terms of the componential structure of an act of remembering, summarized in Figure 2. It consists of 13 conceptual elements, organized in three groups: observables, hypothetical processes, and hypothetical states. Each element is tied to one or two other elements, indicated in Figure 2 by arrows, through relations such as "influences," "has an effect," or "brings about." (The broken arrows in the scheme represent relations that do not affect the ongoing act of remembering but may influence the outcome of a subsequent one.)

The processes of encoding, recoding, ecphory (actualization of a latent engram), and conversion in GAPS are to be thought of as "momentary" processes, or events, in the sense of Miller and Johnson-Laird (1976, pp. 443ff.). The states of the system can be thought of as processes held in abeyance, or as indicants that some processes have been completed and others have not yet begun.

The elements of GAPS can be classified into two categories, elements of encoding and elements of retrieval. The encoding part of an act of remembering begins with the perception of an event and ends with an original or a recoded engram; retrieval begins with the perception of a retrieval cue and ends with the recollective experience of the event, conversion of ecphoric information, or both.

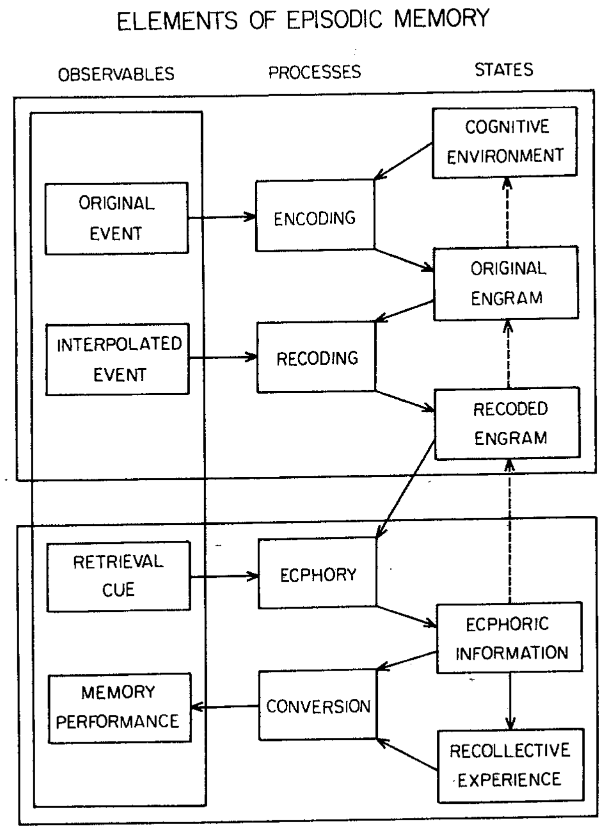


Figure 2 (Tulving 1983b, fig. 7.1, p. 135). Elements of episodic memory and their relations.

### *Elements of encoding*

**Original events.** The basic units of perceived time are events. An event is something that occurs in a particular place at a particular time. The closely related term "episode" refers to an event that is a part of an ongoing series of events.

It is useful to distinguish between the setting and the focal element (or elements) of an event (Hollingworth 1913, pp. 532–33). Setting refers to the time and place in which the event occurs, whereas focal element is a salient happening within the setting. Events are always unique, they are never repeated. But events may resemble one another, by virtue of the similarity of their settings, focal elements, or both. Similarity relations among events play an important role in recoding and in ecphory.

In experiments that can be thought of as having to do with episodic memory, the settings have usually been held constant and the subjects' recollection of them has been taken for granted. Only the focal elements, or the factual or semantic contents (Schacter & Tulving 1982) have been varied, in the form of discrete units of verbal or some other symbolic material. As the remembering of settings has not yet been studied, it can be argued that full-fledged research on episodic memory has not yet begun.

What is stored about an event in memory depends not only on the event as such and its own characteristics, but also on a large number of both temporary and permanent characteristics of the memory system. These characteristics have been known throughout history under a large variety of names. For example, McGeoch (1942, p. 501) referred to them as the "context of the individual's symbolic or ideational events," Bower (1972, p. 85) named them the "organism's cognitive state," and Donald Thomson and I (Tulving & Thomson 1971, p. 123) used the term "cognitive environment" as a label for the factors other than the event that influence the processing of the event.

**Encoding.** Encoding is the process that converts an event into an engram. Encoding processes are manipulated experimentally through encoding operations that subjects are instructed or induced to perform on perceived events.

The effects of encoding operations on remembering are revealed by experiments conducted according to the "encoding paradigm." A large variety of methods and techniques have been used to vary encoding processes (e.g., Craik 1973; Craik & Watkins 1973; de Schonen 1968; Geiselman & Glenny 1977; Hyde & Jenkins 1969; Johnson-Laird, Gibbs & de Mowbray 1978; McClelland, Rawles & Sinclair 1981; Tresselt & Mayzner 1960; Woodward, Bjork & Jongeward 1973). In these experiments, subjects have performed many different kinds of orienting tasks while inspecting the material to be remembered.

The effects of encoding operations on subsequent recall or recognition of the material can be considerable. For instance, in an experiment done by Mathews (1977, Exp. 3) subjects made different *semantic* judgments about word triplets, under either incidental or intentional learning instructions. Probability of recall was the same for incidental and intentional learning instructions, but varied greatly with the nature of the encoding operations performed, from .10 for a relatively ineffective operation to .68 for a relatively effective one.

A number of explanations of the differential effectiveness of encoding operations have been offered, beginning with the seminal paper by Craik and Lockhart (1972) in which differences in retention were attributed to differences in depth of encoding. None of the explanations and theories advanced (e.g., Anderson 1976; Anderson & Reder 1979; Craik & Tulving 1975; Eysenck 1979; Jenkins 1974; Lockhart et al. 1976; Nelson 1979; Postman 1975b; Postman, Thompkins & Gray 1978) has as yet gained general approval. In GAPS, explanations of encoding processes are tied to explanations of retrieval processes.

**Engram.** The product of encoding is an engram, or memory trace. Within GAPS, an engram (the word was coined by Semon 1904), like other hypothetical concepts, is defined in terms of its position in the overall scheme of things and its relations to other elements of the system. Engrams are specified in terms of both their antecedent conditions – particular events particularly encoded in particular cognitive environments – and their consequent conditions, including the circumstances surrounding their subsequent ecphory and retrieval. Different conceptualizations of engrams – whether as information stored about past events, records of operations, attunements, dispositions, images, copies, propositions, analog representations, or as particularly marked parts of associative networks – are compatible with GAPS.

A particularly useful idea is that the engram of an event is a bundle of features, or a collection of some other kind of more primitive elements. This idea, advocated by many contemporary theorists (e.g., Bower 1967; 1972; Estes 1959; Underwood 1969; Wickens 1970) helps us to talk about, and in some sense understand, phenomena of memory that could not have been equally gracefully handled by other languages. One of the advantages of the feature language lies in the fact that it allows us to think about engrams of different events as *qualitatively* different. Two engrams are similar to, or different from, each other to the extent that they possess shared and distinctive features, in keeping with the theoretical analysis of Tversky (1977).

**Recoding, interpolated events, and recoded traces.** One of the most distinctive characteristics of engrams of events is their mutability: Functional properties of engrams change over time. *Recoding* is the generic name of related operations and processes that take place after the encoding of the original event and thereby bring about changes in the engram. Research relevant to the concept of recoding has appeared in the literature under headings such as: repetition effects (e.g., Glanzer & Duarte 1971; Peterson, Saltzman, Hillner & Land 1962); rehearsal (Rundus 1971; Woodward et al. 1973); retroactive effects (e.g., Barnes & Underwood 1959; Postman, Keppel & Stark 1965); retrieval-induced recoding (e.g., Allen, Mahler & Estes 1969; Bartlett & Tulving 1974; Bjork 1975; Darley & Murdock 1971); "mental contiguity" (Glanzer 1969; Jacoby 1974; Wallace 1970); diffusion of trace elements (e.g., Shephard 1961; Shephard & Chang 1963); cue overload (e.g., Earhard 1977; Watkins & Watkins 1975); information integration (e.g., Bransford & Franks 1971; Loftus 1975; Loftus, Miller & Burns 1978; Pezdek 1977); "incrementing" (Raaijmakers & Shiffrin 1981); as well as others. The recoding of an original engram in all these cases is governed by the similarity of interpolated events to the original event, similarity of encoding operations performed on interpolated events, or both.

The most systematic study of recoding of events has been conducted by Loftus (e.g., Loftus 1975; 1977; Loftus et al. 1978; Loftus & Palmer 1974). In these experiments, it has been shown that references to an original event after it has occurred can change what the person reports about the original event. Loftus has interpreted these data to suggest that the interpolated reference to the event modifies the information stored about the event. In the language of GAPS, we would say that the original engram has been recoded. After recoding, utilization of some of the information contained in the original engram is no longer possible.

### *Elements of retrieval*

Engrams have no effect on ongoing mental activity unless they are retrieved. For retrieval to occur, two necessary conditions must be met: The system must be in the "retrieval mode," and an appropriate retrieval cue must be present.

We know little about the retrieval mode, since it has not been systematically studied. Although experiments have been done to compare intentional and incidental learning, "tests" of memory in experiments have always taken place under "intentional retrieval" conditions.

All retrieval is cued: Retrieval does not occur in situations in which appropriate retrieval cues are absent (Jones 1979; Tulving 1976; Tulving & Madigan 1970; Watkins 1979). An important research problem lies in the identification of the nature of "invisible" cues (Eich 1980; Tulving & Watkins 1975) in situations in which no cues appear to be present.

**Ecphory.** Ecphory is the process that combines the information in the retrieval cue and the engram into ecphoric information. The term "ecphory," too, was invented by Richard Semon (1904). Ecphory is one of the two central elements in the process of retrieval; the other is conversion. The distinction between ecphory and conversion is necessary not only because one precedes and the other follows the state of ecphoric information, but also because it helps us to conceptualize the relation between recall and recognition, as we will see presently.

The distinction between ecphory as a subordinate process of retrieval and the superordinate process of retrieval as a whole helps us to distinguish between the concept of retrieval as envisaged in GAPS and the same term as used by other writers (e.g., Anderson & Bower 1973; Indow 1980; Kintsch 1968; Mandler 1980; Mandler, Pearlstone & Koopmans 1969). Most contemporary theories of retrieval conceptualize retrieval as a kind of activation of latent information, or associations (e.g., Anderson & Bower 1972; 1974; Atkinson, Herrmann & Wescourt 1974; Jones 1976; 1980; LeCocq & Tiberghien 1981; Murdock & Anderson 1975; Norman & Bobrow 1979; Raaijmakers & Shiffrin 1981; Ratcliff 1978; Ratcliff & Murdock 1976; Shiffrin 1970). These theories assume, either explicitly or implicitly, that the contents of what a person recalls or recognizes reflect only the information that has been stored. Recall cues contribute to the experience of remembering only by virtue of determining what part of the stored information is activated. Thus, these theories attribute to episodic memory a retrieval process that characterizes semantic memory.

In episodic memory, according to GAPS, the process of ecphory is a constructive activity – a synergistic process – that *combines* the (episodic) information from the engram and the (semantic-memory) information from the cue. Similar ideas about the nature of remembering have been advocated by Bartlett (1932), Neisser (1967), Bransford and Franks (1971), Kintsch (1974), among others.

The complementary function of retrieval information in ecphory is somewhat conjectural, in that evidence for it is mostly indirect. The experimental work that comes closest to providing direct evidence for complementarity of engrams and cues is a series of experiments by Loftus and her colleagues mentioned earlier (Loftus 1975; Loftus & Loftus 1980; Loftus et al. 1978; Loftus & Palmer 1974).

**Ecphoric information and recollective experience.** Ecphoric information is the product of the process of ecphory. It determines the particulars of recollective experience and provides the input into the conversion process: What a person remembers of an event depends directly on the quantity and quality of relevant ecphoric information.

Recollective experience refers to the rememberer's subjective awareness of ecphoric information. The terms that have been most frequently used in descriptions of the mental experience of remembering are "memory image" and "consciousness": When a person remembers a past event, he has a mental image of it and is consciously aware of its being a mental replay of what happened once before.

The feeling that the present recollective experience refers to a past event and the feeling that the experience is veridical are determined by the intrinsic properties of ecphoric information. A reasonable assumption is that the intensity of the feeling of pastness is directly correlated with the relative contribution that the information from the (episodic) engram makes to the ecphoric information.

Ecphoric information can also serve as retrieval information. The product of ecphory involving ecphoric information as one source of input into the process is a new and different assembly of retrieval information. The recursive operation may be repeated until some stop rule is invoked (Kintsch 1974; Lockhart et al. 1976; Raaijmakers & Shiffrin 1981; Semon 1904).

GAPS describes a "snapshot view" of episodic memory: It focuses on conditions that bring about a slice of experience frozen in time which we identify as "remembering." The recursive operation of the process of ecphory, feeding upon the (changing) ecphoric information and combining it with the "fixed" stored episodic information, produces many snapshots whose orderly succession can create the mnemonic illusion of the flow of past time.

**Conversion and memory performance.** The act of remembering a particular episode may end with the recollective experience. The rememberer "just thinks about" the experience and does not express it in any overt fashion. At other times, recollective experience, or ecphoric information of which a rememberer is not directly aware, is converted into behavior. The form of conversion (e.g., recall, recognition, some "memory judgment") can be more precisely stipulated in laboratory experiments than in real life, but the general principles governing conversions are assumed to be very much the same.

The elements of the retrieval process in GAPS that have been labelled "ecphoric information" and "conversion" are related to the distinction between "memory" and "decision" in signal-detection analyses of memory (e.g., Lockhart & Murdock 1970; Murdock 1974). In recognition-memory tasks, "decision processes intervene between memory and response" (Murdock 1974, p. 8); in GAPS, conversion processes intervene between ecphoric information and overtly observable memory performance. The main difference between signal-detection analyses of recognition memory and GAPS lies in the process of ecphory.

## Part III: Synergistic ecphory

The third main part of the book discusses findings from experiments that have helped to shape the overall structure of the General Abstract Processing System as summarized in the second part. The section begins with a brief review of the history of the work that led to the idea of encoding specificity.

### From organization to encoding/retrieval interactions

With some imagination it is possible to see how the ideas discussed in the book grew out of my early work on subjective organization in multi-trial free recall (Tulving 1962) and intratrial and intertrial forgetting (Tulving 1964).

In 1966, Zena Pearlstone and I (Tulving & Pearlstone 1966) did a large experiment in which we compared cued and noncued recall. The finding that cued recall was better than noncued implied that recall depends on conditions of both storage and retrieval, that the failure of noncued recall of an item does not signify absence of stored information about the item, and that, with storage conditions held constant, successful recall varies as a function of the number and appropriateness of retrieval cues.

We used the term "availability" to refer to the hypothetical presence of information in the memory store, and the term "accessibility" to designate that part of the available information that could be retrieved. Given the idea that recall depends on both availability and accessibility, it was easy to imagine that sometimes some cues might be effective where others would fail. What determines the effectiveness of cues?

We did a number of small experiments, under relatively casual conditions, to try out a number of ideas relevant to this question. We quickly found out that the presence of preexperimental associations between cues and to-be-recalled list words did not always suffice for accessibility, and that effectiveness of cues seemed to depend on processes occurring at the time of study. For instance, a descriptive phrase such as *double letter in the middle* is an effective retrieval cue for a studied target word such as SUMMER, but *only if* the subject, while studying the list, had noted the fact that SUMMER was a word that had a double letter in the middle.

We formalized the operations of these "quick and dirty" experiments in a single but extensive experiment designed to examine the effectiveness of retrieval cues as a function of encoding conditions (Tulving & Osler 1968). The results of the experiment showed that "Specific retrieval cues facilitate recall if and only if the information about them and their relation to the [to-be-remembered] words is stored at the same time as the information about the membership of the [to-be-remembered] words in a given list" (Tulving & Osler 1968, p. 599). We thought that the same relation between effectiveness of cues and encoding conditions would hold generally, including situations where the learners were left free to encode the to-be-remembered words any way they wanted.

To test the generality of the conclusion, Donald Thomson and I did three experiments in which we varied both the encoding conditions and the preexperimental strength between the cue and target words (Thomson & Tulving 1970). We found that the cueing effectiveness of even a very strong preexperimental associate of the target word depends on what happens at the time of study. We contrasted these findings with the predictions made by the generation/recognition models of recall (e.g., Bahrick 1969; 1970; Kintsch 1970), according to which strongly associated words should have been effective cues regardless of how the target words had been encoded. We concluded that these models were wrong and that the two critical assumptions on which they were based – assumptions that we referred to as associative continuity and transsituational identity of words – were in need of revision.

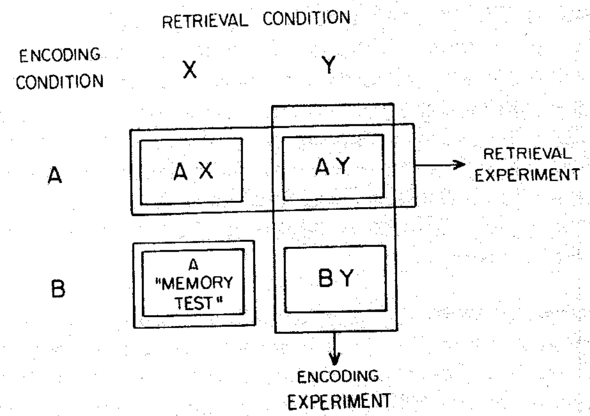


Figure 3 (Tulving 1983b, fig. 10.1, p. 220). A schematic diagram of the encoding/retrieval paradigm of episodic memory research.

Tulving and Osler's (1968) and Thomson and Tulving's (1970) experiments conformed to what we now refer to as the "encoding/retrieval paradigm." In this paradigm, both encoding and retrieval conditions are experimentally manipulated. Subject and material variables are usually held constant, although they could be varied as additional dimensions.

A schematic representation of a minimal encoding/retrieval experiment is shown in Figure 3. In it, two encoding conditions, A and B, are crossed with two retrieval conditions, X and Y. We can think of the total design as entailing two retrieval experiments and two encoding experiments, all conducted simultaneously. In a retrieval experiment, encoding conditions are held constant and retrieval conditions varied: Each of the two rows in Figure 3 represents a retrieval experiment. In an encoding experiment, retrieval conditions are held constant and encoding conditions are manipulated: Each of the two columns in Figure 3 represents an encoding experiment. An experimental situation in which both encoding and retrieval conditions are held constant represents a memory test of the kind used in psychometric measurement of abilities.

Outcomes of single retrieval experiments and single encoding experiments are theoretically uninteresting, because they seldom permit discrimination among alternative explanations. Thus, the finding that, say, recall is better after one way of studying the material than after another affords many possible interpretations; so does the finding that one set of cues leads to better recall than another. On the other hand, strong interactions between encoding and retrieval conditions that can be observed in an encoding/retrieval experiment do rule out explanations that fit individual encoding and retrieval experiments. The critical findings in the Tulving and Osler (1968) and Thomson and Tulving (1970) experiments had to do with such interactions between encoding and retrieval conditions. It was these interactions that led to the idea of encoding specificity.

### *Encoding specificity*

Although the concept of encoding specificity had its beginning in the pursuit of the problem of the effectiveness of retrieval cues, it has changed over the intervening years. The essence of the concept now lies in the emphasis on the *relation* between the specifically encoded (and perhaps recoded) memory trace and the particular retrieval information as the determinant of recollective experience: The engram and the retrieval cue must match and complement each other for remembering to occur.

The concept of encoding specificity denies the validity of a number of ideas that not too long ago were widely held: The idea that items (events) of a particular class are easier to remember than items (events) of another class, that a particular encoding operation is more effective than another encoding operation, that a particular type of retrieval cue is more effective than another type of cue, that copy cues provide automatic access to the stored information, that recognition-memory performance provides a measure of trace strength, and that memory traces have strength. These ideas are no longer tenable. According to encoding specificity, no absolute statements about the memorability of items and the effectiveness of particular kinds of encoding operations or particular kinds of retrieval cues are justified. Further, the effectiveness of all cues, including copy cues, depends on the conditions under which the target event was encoded. Finally, recognition-memory tests provide no better basis for estimating the strength of memory traces than do any other memory tests, or the application of any other type of cue. Indeed, traces have no strength independently of conditions in which they are actualized: Any given trace has many different "strengths," depending upon retrieval conditions.

Empirical evidence in support of encoding specificity is provided by findings from a large number of encoding/retrieval experiments. A representative list of such experiments includes those by Baker and Santa (1977, Exp. 2), Dong (1972, Exp. 2), Eich, Weingartner, Stillman, and Gillin (1975), Fisher and Craik (1977, Exp. 3), Geiselman and Glenny (1977), Godden and Baddeley (1975), Jacoby (1973), Masson (1979, Exp. 3), Morris (1978, Exp. 2), Ozier (1978), Roediger and Adelson (1980, Exp. 3), Stein (1978, Exp. 1), Thomson (1972, Exp. 4), and Till and Walsh (1980, Exp. 3). In all these experiments, crossover interactions (G. Loftus 1978) between encoding and retrieval were observed under a wide variety of conditions. The to-be-remembered materials included unrelated words, homographs, word pairs, categorized words, words embedded in sentences, and whole sentences. Encoding conditions were manipulated in terms of distributions of target words in study lists, test expectations, verbal context of target words, intentional learning, interactive imagery, pleasantness ratings, comprehension judgments, "thinking about" initial letters or categories of to-be-remembered words, judgments of meaningfulness of relations between comparison and to-be-remembered sentences, single or pairwise presentation, typicality of actions depicted by sentences, imagined voices speaking visually presented words, changed physical environments, and changed drug states of the rememberer. In most experiments, subjects engaged in intentional learning; in some, learning was incidental. Retrieval conditions were manipulated by asking the subjects to engage in free recall or in different kinds of recognition, or by presenting various types of retrieval cues: intralist and extralist cues, associatively related words, category names, initial letters, parts of studied sentences, words describing plausible inferences drawn from the studied material.

Thus, the relativity of "goodness" of encoding operations and the relativity of the power of retrieval cues, or the critical requirement of compatibility between specifically encoded engrams and retrieval cues, seem to hold, if not universally, then at least over a very wide set of conditions.

The initial conflict between encoding specificity ideas and generation/recognition theories of recall has been largely resolved. Generation/recognition models have been revised and brought in line with the concept of encoding specificity (e.g., Kintsch 1974). Moreover, the concept of encoding specificity is perfectly compatible with generation/recognition as an effective *strategy* of retrieval under certain conditions (Rabinowitz, Mandler & Barsalou 1979). And there is substantial agreement that the system can generate potentially effective retrieval information.

But some problems more fully discussed elsewhere (Rabinowitz et al. 1979; Tulving 1976; Watkins & Gardiner 1979) remain. These have to do with the question of whether the process of generation is guided by episodic or semantic information, and the question of whether the product of the generation process has to be some sort of a "copy" of the to-be-remembered item, rather than just (any kind of) useful retrieval information. Still another problem concerns the assumptions that are made about what it is that determines whether the generated information is accepted by the subject as "desired" or rejected as "not desired."

It is sometimes useful to distinguish between the encoding specificity hypothesis and the encoding specificity principle. The hypothesis is a tentative statement about the relation between the properties of the memory trace of an event and the effectiveness of retrieval cues; its tenability can be evaluated empirically. The principle is the assumption that the hypothesis is true; its usefulness depends on the truth of the hypothesis.

Three entities are involved in the testing of the encoding specificity hypothesis: the engram of the event, the retrieval cue, and the relation between them. When we test the hypothesis, we must know, or must be in a position to make reasonable assumptions about, the encoded features of engrams, as well as the features useful for retrieval. The results of the observations then tell us something about the third entity, the relation between the trace and the cue.

When we adopt encoding specificity as a principle, we can make inferences about the informational contents of memory traces on the basis of the observed effectiveness of retrieval cues. In so doing, we rely on the same logic that has been used in other situations to describe objects that are not directly observable.

### Criticisms of encoding specificity

Encoding specificity ideas have been criticized by a number of people on a variety of grounds. Some have declared the ideas to be untestable, others have said they are false; some have questioned the underlying logic, others have produced data that they have interpreted as contrary to encoding specificity; some have provided alternative explanations of findings supportive of encoding specificity, others have failed to replicate the findings. We will consider a sample of criticisms next, together with rebuttals of some of them.

The critical evidence for the concept of encoding specificity, as we have seen, has been provided by encoding/retrieval interactions. But what are we to make of experiments that yield no evidence of such interactions? One such experiment has been described by Postman (1975a, Exp. 1). It was designed to compare the effectiveness of different types of intralist and extralist cues following the study of target words in the presence or absence of associated list cues. The experiment was complex, comprising 16 experimental conditions, with a number of them nominally corresponding to the conditions in Thomson and Tulving's (1970) experiment.

Postman's data did not replicate the critical interaction found in the Thomson and Tulving experiment. Instead, they showed that a "strong" extralist cue (e.g., *bloom*) was as effective in eliciting the corresponding target word (FLOWER) following the target word's encoding in the presence of a "weak" cue as it was following the target's encoding as a single item in the study list, and that in both of these conditions strong-cue recall was considerably higher than noncued recall. Postman concluded that the kinds of results reported by Thomson and Tulving were obtained only under "narrowly circumscribed conditions" (Postman 1975a, p. 64).

This conclusion is defensible, as long as one compares only the two experiments in question. But when we consider encoding/retrieval interactions that, as we saw earlier, have now been observed under a wide variety of experimental conditions, Postman's criticism loses its force. Instead, a more reasonable interpretation is that, for reasons unknown, the intended manipulation of encoding conditions did not work in Postman's experiment.

Several writers have produced data and arguments claiming that, contrary to the encoding specificity hypothesis, unencoded cues are effective (e.g., Anderson & Pichert 1978; Baker & Santa 1977; Kochevar & Fox 1980; Light 1972; Marcel & Steel 1973). The relevant data are those showing that cues not present at the time of study and unlikely to have been encoded by the subjects at that time were nevertheless effective when given at retrieval.

There are two problems with the conclusions drawn by these critics. First, some of the relevant data have been drawn from simple retrieval experiments, rather than encoding/retrieval experiments, and such data are not compelling. Second, the criticisms would hold only if the experimenters' assumptions as to how their subjects encoded the material were true: Whether they are is not known. Baker and Santa (1977), for instance, chose to reject the encoding specificity hypothesis on the basis of comparisons of free and cued recall in retrieval components of their two experiments which together conformed to the pattern of the encoding/retrieval paradigm. The results of the two experiments considered together provide strong support for encoding specificity.

A different type of critical experiment is one that purports to show that encoded cues are ineffective. A representative experiment in this category is one reported by Humphreys and Galbraith (1975, Exp. 2). Using asymmetrical free-association pairs of words (e.g., *tobacco* and *smoke*) as a source of materials, they did an experiment whose results were very much what one would have expected on the basis of the generation/recognition theory, and contrary to those expected on the basis of the encoding specificity hypothesis. But this interpretation, as Humphreys and Galbraith acknowledged, depended on a critical assumption concerning the encoding of study-list words. When this assumption was put to experimental test by Ley (1977), it turned out to be wrong, thereby invalidating Humphreys and Galbraith's criticism.

Another type of objection to encoding specificity ideas concerned their circularity. Solso (1974, p. 28) put it as follows: "If a cue was effective in memory retrieval, then one could infer it was encoded; if a cue was not effective, then it was not encoded. The logic of this theorization is 'heads I win, tails you lose' and is of dubious worth in the history of psychology. We might ask how long scientists will puzzle over questions with no answers."

This criticism would be a serious one if it were indeed true that one always explained cue effectiveness by assuming certain properties of memory traces, and, at the same time, justified one's assumptions about the trace properties in terms of observed effectiveness of retrieval cues. But we can get out of the circle by doing what already has been done in many experiments: manipulate properties of memory traces by varying antecedent conditions, such as the events to be remembered or the way in which they are encoded. Although we never know exactly what information has been encoded into the trace, we can frequently influence it and make informed guesses about it. Experiments by Fisher and Craik (1977), Morris (1978), Geiselman and Glenny (1977), and Godden and Baddeley (1975), as well as many others, provide relevant evidence. Thus, in many cases, the observed effectiveness, or lack of effectiveness, of retrieval cues makes good sense in terms of what we know about the to-be-remembered events, their encoding, and their compatibility with cues at the time of retrieval. In other situations, however, we make inferences about how an event was encoded on the basis of observed effectiveness of retrieval cues (e.g., Ogilvie, Tulving, Paskowitz & Jones 1980; Tulving & Watkins 1975) and by so doing we extrapolate from what is known to what is unknown. The logic here is the same as the one used in estimating the strength of habits, associations, and memory traces on the basis of observed behavior, or characteristics of encoded features on the basis of observations of release from proactive inhibition (Wickens 1970), or false recognition (Anisfeld & Knapp 1968; Underwood 1965).

Until such time as the encoding specificity hypothesis is shown to be contrary to facts, it seems reasonable to claim that at some as yet undetermined level of abstraction, it probably holds for all phenomena of episodic memory, in the sense that there are no exceptions to it.

### *Recognition failure*

Recognition failure is the name of the finding that previously studied items cannot be identified as "old" although their names can be reproduced to other cues. Like the encoding/retrieval interactions we discussed, the phenomenon of recognition failure represents one of the many applications of GAPS to the understanding of remembering of word-events.

Scattered examples of experimental findings showing recall to be better than recognition – thereby implying recognition failure – had been reported in the literature before recognition failure became a full object of theoretical interest (e.g., Bahrick & Bahrick 1964; Bruce & Cofer 1967; Lachman & Field 1965; Postman, Jenkins & Postman 1948; Tulving 1968b). These initial observations failed to generate much interest, presumably because the magnitude of the effect was not striking, because they were made under rather special conditions, or because the effects could be explained in terms of uninteresting assumptions, such as "chance fluctuation" of attention.

In 1973 Donald Thomson and I reported findings that seemed to have more serious implications for theory (Tulving & Thomson 1973). The method we used in our initial experiments was rather cumbersome and at the present time it is of historical interest only. Subsequent research has shown that the essential ingredients for the production of recognition failure are three: (a) presentation of an A–B pair of words for study, (b) presentation of B alone as a test item in the recognition test, and (c) presentation of A as a cue for recall of B. In our very first experiment (Tulving & Thomson 1973, Exp. 1) we found that subjects could identify only 24% of the previously seen B items as "old," although they recalled 63% of these targets when A items were presented as cues. These data ruled out the generation/recognition theory of recall, since it was impossible, according to theory, for words to be recalled that could not be recognized. The finding was also contrary to common sense: How could perfectly normal and intelligent people claim that they had not seen a familiar word in the study list but produce it to its pair-mate? A general interpretation of the finding, however, seemed possible with the aid of the encoding specificity principle: It looked as if the encoded trace of the A–B event was such that cue A was more compatible with it than cue B.

Probably because of the counterintuitive nature of the phenomenon, the demonstration of recognition failure created a certain amount of interest, scepticism, and criticism (e.g., Light, Kimble & Pellegrino 1975; Martin 1975; Postman 1975a; Rabinowitz, Mandler & Patterson 1977; Reder, Anderson & Bjork 1974; Santa & Lamwers 1974; 1976). The criticisms, documented by John Gardiner in an in-house report, ranged from the denial of the existence of the phenomenon to the claim that it was trivial and theoretically irrelevant. Since a great deal of research has been done to demonstrate the occurrence of the phenomenon under a wide variety of conditions, and since the magnitude of the effect has been brought under experimental control, most of the criticisms have by now fallen by the wayside.

**Occurrence and magnitude.** The magnitude of recognition failure can be indexed by the conditional probability of failure to recognize a previously seen item given that it was recalled (Watkins & Tulving 1975). This measure is intuitively meaningful as it directly expresses the proportion of recallable words that cannot be recognized. For certain purposes, however, it is more convenient to work with the complement of the recognition failure measure. This "recognition success" measure is given by the recognition hit rate conditionalized on recall.

In individual experiments, recognition success is highly correlated with overall recognition hit rate across a large number of individual experiments and experimental conditions (Flexser & Tulving 1978; Tulving & Wiseman 1975). A representative set of data is shown in Figure 4. Each data point in the graph represents a large number of observations, pooled over many subjects and many words, in a particular condition in a particular experiment. The systematic relation between recognition failure and overall recognition hit rate shown in Figure 4 is invariant with large variations in levels of recall and recognition in different experiments, with little correlation between them, as shown in Figure 5.

The data in Figure 4 were fitted with a quadratic function of the form:

$$P(Rn|Rc) = P(Rn) + c[P(Rn) - P(Rn)^2] \quad \text{(Equation 1)}$$

This equation expresses recognition success as a function of recognition hit rate with a single constant. The equation is shown as the solid line in Figure 4, with the value of the constant $c$ set equal to 0.5 on the basis of the least-squares solution of Equation 1.

Recognition failure occurs, and its magnitude is governed by the overall level of recognition according to the function defined by Equation 1, in experiments in which

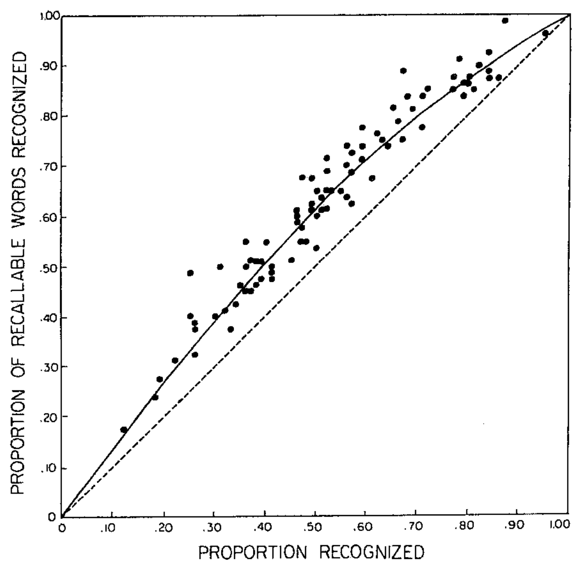


Figure 4 (Tulving 1983b, fig. 13.1, p. 282). Probability of recognition of recallable target words as a function of the probability of recognition of all target words. Each data point represents a separate experiment or a condition in an experiment.

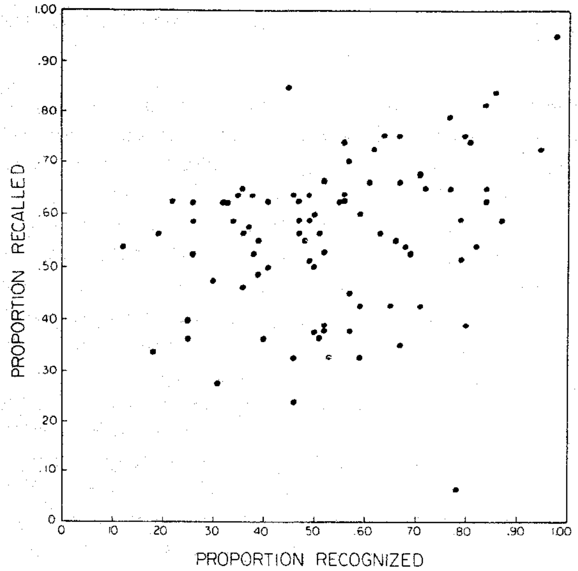


Figure 5 (Tulving 1983b, fig. 13.2, p. 283). Probability of recall of target words as a function of the probability of recognition of the same words. Each data point represents a separate experiment or a condition in an experiment.

recognition is better than recall as well as in experiments in which recall is better than recognition (Flexser & Tulving 1978; Wiseman & Tulving 1976); with naive as well as highly sophisticated subjects (e.g., Begg 1979; Bowyer & Humphreys 1979; Rabinowitz, Mandler & Barsalou 1977; Wiseman & Tulving 1975); in immediate tests as well as tests given a week after the learning (Tulving & O. C. Watkins 1977); with semantically related as well as unrelated study pairs (e.g., Begg 1979; Rabinowitz et al. 1977; Vining & Nelson 1979); with semantically related as well as unrelated recognition-test distractors (e.g., Begg 1979; Bowyer & Humphreys 1979; Postman 1975a; Rabinowitz et al. 1977); in situations in which no distractor items are presented at all in recognition tests (e.g., Begg 1979; Wallace 1978); with low-frequency words that have few semantic senses (Reder et al. 1974), as well as words that have only a single meaning in the dictionary (Tulving & O. C. Watkins 1977).

Recognition failure also occurs in experiments in which the study material consists of simple three-word sentences, such as "Shelter protected ESKIMO," and in which the cue presented in the recall test was not an explicit part of the input material, representing only an inference from the studied sentence (e.g., *igloo*). This experiment, using materials inspired by a study by R. C. Anderson, Pichert, Goetz, Schallert, Stevens, and Trollip (1976) was done at Toronto in collaboration with Norman Park. "Spectacularly large" recognition-failure effects have been demonstrated by Watkins (1974) in an experiment in which target items were meaningless letter bigrams. And Muter (1978) has demonstrated recognition failure, with the data adhering to the recognition-failure function, in a semantic-memory experiment.

Although sizable deviations of data from the recognition-failure function have been reported by Begg (1979), Nilsson and Shaps (1981), and Gardiner and Tulving (1980), the robustness of the phenomenon is no longer in doubt. A large number of explanations has been offered to account for it. One of the more popular ideas has been that recognition failure occurs because of associative asymmetry between the A and B members of a studied pair of items (e.g., Bartling & Thompson 1977; Rabinowitz et al. 1977; Salzberg 1976). There are two main problems with such theories. First, it is not clear whether associative asymmetry is a cause or a consequence of recognition failure (Tulving & Thomson 1973). Second, it is not clear that the relation between forward and backward retrieval, on the one hand, and recognition failure on the other, represents anything other than the general systematic relation between overall recognition and recognition failure, as depicted in Figure 4, a relation that requires explanation in itself.

Theories of recognition failure based on the idea of associative asymmetry, as well as some other theoretical accounts (e.g., Kintsch 1978; Reder et al. 1974) have been concerned with explanations of the occurrence of recognition failure. They have had little to say about the highly systematic nature of the relation between recognition and recognition failure. Such theories have been offered by Begg (1979), Humphreys and Bowyer (1980), and Flexser and Tulving (1978), with a refinement of the latter offered by Jones (1978). Of these, only the Flexser and Tulving (1978) model accounts for the quantitative relation between recognition failure and overall recognition, that is, for the observed value of $c$ in the equation of the recognition-failure function (Equation 1). It does so by working out the logical consequences of retrieval as matching of trace features, under three simple assumptions: (a) In the recognition-failure paradigm, the recognition cue and the recall cue are directed at the same episodic trace; (b) individual traces vary in terms of how well they have been encoded; and (c) information extracted from the recognition cue is uncorrelated with that extracted from the recall cue.

This model of retrieval independence has been tested by conducting a large number of simulated experiments, in which the model's six parameters are *randomly* varied and which, as a consequence, show large, uncorrelated variations in proportions of target items recognized and recalled (as shown in Figure 5), but which also produce recognition-failure data points closely adhering to the function (as shown in Figure 4). Thus, the model achieves a good qualitative and quantitative agreement between experimental and theoretical data; it does so *without* fixing the values of any model parameters. The value of parameter $c$ in Equation 1 (approximately 0.5) falls out of the assumptions and the structure of the model, independently of particular numerical inputs into the model.

The phenomenon of recognition failure represents a subclass of a class of phenomena showing that target items not retrievable in one retrieval situation may become retrievable in another (e.g., Arbuckle & Katz 1976; Buschke 1974; Erdelyi & Becker 1974; Estes 1960; Hoppe & Dahl 1978; Madigan 1976; Wallace 1978). But its counterintuitive nature and the fact that its magnitude can be very large have theoretical implications that other demonstrations of dependence of retrieval on retrieval conditions do not necessarily share. For instance, in what sense can one talk about "strength" of traces of target items in episodic-memory experiments, given that recallability and recognizability of these items are largely uncorrelated?

### Recognition and recall

Recall and recognition *tasks* differ in two basic ways. One of these has to do with the type of retrieval information: In recognition tasks, a copy of the to-be-remembered item is given, whereas in recall tasks it is not. The second has to do with conversion: In recall, ecphoric information has to be converted into a description (e.g., production of the name) of the target event, whereas in recognition the subject has to make a simple "familiarity" judgment on the basis of the comparison of ecphoric information with the test item. Thus, in the recognition task, copy cues are presented and familiarity judgments are to be made, whereas in recall no copy cues are given and the name of the target item is to be identified and produced.

With the help of Judith Sutcliffe, we compared subjects' memory performance in a "direct comparison" experiment in which study lists, encoding of target words, and retrieval cues were all held constant, and only conversion requirements varied. Thus, for instance, we compared subjects' performance on familiarity conversion (Yes/No recognition judgments) with the "name identification" conversion (e.g., recall of list words), using either copy cues or, in a parallel comparison, associatively related extralist cues.

The experiment yielded two interesting observations. First, with copy cues performance was better when subjects had to make familiarity judgments about the cues than when they had to produce the (identical) names of target items. Second, the correlation between false recognition and the valence (effectiveness) of associative cues was *negative:* Subjects were more likely to make false positive recognition responses to those associative cues that were least effective in eliciting target items under the "name identification" conversion (recall) conditions. These data are clearly contrary to some earlier speculations about the relation between recall and recognition (Tulving 1976) and necessitated rethinking of the issue of the relation between recall and recognition.

**Synergistic ecphory model.** The data from the "direct comparison" experiment can be accommodated by the "synergistic ecphory" model schematically depicted in Figure 6. The horizontal axis of the graph space represents trace information, the vertical axis represents retrieval information, and the two-dimensional space defined by the two axes corresponds to ecphoric information. The two curved lines in the diagram represent two conversion thresholds. One determines naming (recall) of target events while the other determines their episodic familiarity. Each bivariate point in the ecphoric space in Figure 6 corresponds to a particular bundle of ecphoric information whose qualitative characteristics define the nature of the rememberer's recollective experience.

The two conversion thresholds divide the total space of ecphoric information into three mutually exclusive subspaces. The region above the naming threshold contains ensembles of ecphoric information that are sufficient for both awareness of episodic familiarity and the naming of the target event; the region between the familiarity and naming thresholds represents ecphoric information that is sufficient for making positive familiarity judgments but insufficient for naming the target event; the region below the familiarity threshold consists of ensembles of ecphoric information that fail to give rise to any feeling of episodic familiarity. The placement of the naming threshold above the familiarity threshold means that it takes more, or higher quality, ecphoric information for the naming of the target event than for a familiarity judgment.

The model assumes that retrieval information is independent of trace information. A trade-off between trace information and retrieval information is possible within limits: Poor quality of trace information can be compensated for by high-quality retrieval information, and vice versa. Conversion thresholds are asymptotic with the coordinate axes: No recollective experience or conversion can occur in the absence of either the trace or the appropriate retrieval information. Quantitative measures of memory performance are reflected in the distance of bundles of ecphoric information from relevant conversion thresholds.

The model accommodates a number of basic observations about recall and recognition. For instance, the negative correlation between the valence of associative cues and their (false) recognition occurs because more effective retrieval cues produce ecphoric information that is sufficient for naming, and hence sufficient for the judgment that the target item is different from the cue (e.g., point $b,z$ in Figure 6), whereas less effective cues may produce ecphoric information (e.g., point $b,y$) that creates a feeling of familiarity but does not permit the retrieval of the target item, thus leading the subject to (falsely) accept it as "old." The phenomenon of recognition failure fits into the model, despite the fact that name (recall) threshold is higher than familiarity (recognition) threshold, because the thresholds are defined with respect to ecphoric rather than trace information, as used to be the case in classical strength-threshold models of recall and recognition, models that have now been thoroughly discredited (Anderson & Bower 1972; McCormack 1972; Tulving 1976). Other basic facts concerning the relation between recall and recognition can also be accommodated within the model.

The model can be used to illustrate graphically trace-dependent and cue-dependent forgetting, as well as "reversal" of forgetting when retrieval cues are changed (e.g., Tulving 1974). It shows how qualitative properties of memory traces and retrieval cues determine the qualitative properties of ecphoric information, and how the latter are related to quantitatively measured memory performance: "Proportion correct" in an experimental condition is determined by the proportion of bundles of ecphoric information that lie above the relevant conversion threshold. And it suggests that the feeling of pastness in recollection may be determined by the contribution of trace information to the ecphoric information: For instance, in Figure 6, ecphoric information represented by point $a,z$ contains little trace information, whereas $c,y$ contains more trace information but less retrieval information. We might expect, therefore, that the recollective experience corresponding to $a,z$ is tinged with a fainter flavor of pastness, and may seem subjectively less veridical, than the recollective experience based on $c,y$.

The synergistic ecphory model shares many ideas with other contemporary theories of recognition and recall. It is in good agreement with the theory proposed by Lockhart et al. (1976), and has a number of important features

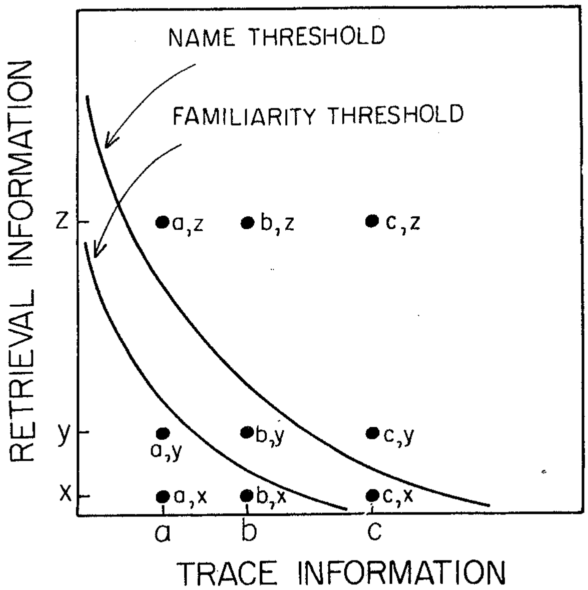


Figure 6 (Tulving 1983b, fig. 14.3, p. 312). A schematic diagram of the Synergistic Ecphory Model of Retrieval. The graph space represents ecphoric information, defined as a product of trace and retrieval information.

in common with theories proposed by Moeser (1977), Kintsch (1974), and Ratcliff (1978). Different recognition and recall thresholds postulated in the synergistic ecphory model are related to the idea, advocated by Mandler as well as others (e.g., Atkinson & Juola 1974; Humphreys 1978; Mandler 1980; Mandler et al. 1969; Tiberghien 1976), that recognition can occur either as a detection of familiarity or as a consequence of particular retrieval operations.

The unique feature of the synergistic ecphory model is the concept of ecphoric information as a conjunction of trace and retrieval information. Although there is as yet little direct evidence to support the concept, a number of facts about remembering and recollective experience, including the phenomenonal reality of variability of subjective feelings of pastness and veridicality, seem to necessitate the postulation of ecphoric information as a synergistic product of two memory systems, episodic and semantic.

ACKNOWLEDGMENTS
The writing of the book was made possible by an Izaak Walton Killam Memorial Scholarship awarded by the Canada Council and a Commonwealth Visiting Professorship at the University of Oxford; the research reported in the book has been supported by the Natural Sciences and Engineering Research Council of Canada and the National Science Foundation of the United States. I am deeply grateful for this support. I am equally grateful to many individuals who have helped me in the task in a variety of significant ways, especially Fergus Craik, Carol Macdonald, Henry L. Roediger III, Daniel Schacter, Michael Watkins and Lawrence Weiskrantz.

NOTE

1. The book is written at two levels, "serious" main text interspersed with in-text footnotes that provide personal, sometimes tongue-in-cheek, commentary on the main text. These more personal sections of the book are studiously ignored in the present précis.

# Open Peer Commentary

*Commentaries submitted by the qualified professional readership will be considered for publication in a later issue as Continuing Commentary on this article. Integrative overviews and syntheses are especially encouraged.*

## Neuropsychological evidence and the semantic/episodic distinction

Alan D. Baddeley
*MRC Applied Psychology Unit, Cambridge CB2 2EF, England*

While there is general agreement that the semantic/episodic distinction is heuristically useful, the claim that separate functional systems are involved is much less plausible. At first sight, the mass of evidence summarised in Table 1 of Tulving's accompanying Précis seems overwhelming. It does not, however, speak directly to the issue of whether separate systems are involved. An analogy might help explain why.

I first read this section of Tulving's book (1983b) while in a plane flying over wooded countryside. Out of the window, the forest beneath looked like a grey-green carpet, totally different from what its appearance would have been had I been standing in the forest. I could easily produce a long list of perceptual differences in terms of sight, sound, and even smell between the forest as experienced from the plane and as experienced from within. Would I therefore be entitled to conclude that they were quite separate forests? Clearly not. By analogy, one can reasonably argue that semantic and episodic memory emphasize different aspects of the same system.

With this in mind, the neuropsychological evidence for a distinction becomes particularly important. If it can be shown that one part of the brain is necessary and sufficient for episodic memory but unnecessary for semantic memory, while another part of the brain is necessary for semantic but not episodic memory, then the argument for two separate systems becomes vastly stronger. Tulving recognises this and appeals to neuropsychological evidence. How convincing is this appeal?

Consider first the blood flow study by Wood, Taylor, Penny & Stump (1980) cited in Tulving's Précis. This shows that the pattern of blood flow within the brain is somewhat different depending on how a subject is required to process a given word. More specifically, recognising whether a word has been presented previously gives rise to a somewhat different blood flow pattern from judging whether that word represents an object that could be contained in one's living room. This result is cited by Tulving as evidence for the separate location of semantic and episodic memory within the brain. Such an interpretation is possible but far from compelling. First, as Wood et al. (1980) point out, there are a number of differences between the two tasks other than the possibility that one relies on semantic and the other on episodic memory; they are, for example, different in difficulty. Second, even if such extraneous factors are ignored, the evidence merely suggests that the two processing tasks are different, and that this difference is reflected in cerebral blood flow. Given a fine enough measure of blood flow, it is conceivable that any two tasks that differ cognitively may be detectably different in blood flow pattern. Would one therefore wish to assume a physically separate system for each task? Clearly not.

Rather more compelling evidence for separate systems comes from the study of amnesia. A number of workers, including, alas, myself, have suggested that this implies a functional separation between semantic and episodic memory (Baddeley 1982a; Kinsbourne & Wood 1975). It is important in discussing this evidence not to confound the question of the distinction