2

CHAPTER

Endel Tulving

# Does Memory Encoding Exist?

Does memory encoding exist? What a silly question—of course it does! The whole world's literature on memory is full of papers and books on encoding, or phenomena involving encoding, and surely one cannot write about something that does not exist, at least not in science. Or can one?

At one time physicists wrote at great length about the aether, chemists about phlogiston, and biologists about élan vital, and everything that was said about these things turned out to be wrong, because they did not exist. While pursuing their interests, scientists did find out many things about their subject matter that became not only useful but true, but the fact remains that it is quite possible for scientists to think, talk, and write about nonexistent entities. Is encoding one of them?

There are two basic kinds of existences in the world: physical and mental. Many very clever people, from Descartes to the great thinkers of our own age, have written a great deal about the nature of the relation between the physical and the mental. Physical existences include things such as galaxies, stars, planets, oceans, pebbles on the seashore, tortoises, and worms, and the molecules and atoms that they are made of. Mental ones consist in emergent creations of an unbelievable physical thing, the brain: percepts, images, feelings, ideas, beliefs and the like, and their equivalents in nonhuman animals.

Whether or not encoding exists in the physical world is as yet unknown, but encoding certainly exists as a thought in the minds of many people. Encoding is a concept. In the science of memory, we have a very large number of concepts. We use them continually and are utterly dependent on their existence. We have linguistic, or sometimes mathematical, labels for them, *terms* as these labels are called, very much in the same way that we have symbolic labels for things that do exist as tangible physical entities. We create concepts and terms, describe and define and explain them, have faith in them, defend them against nonbelievers, and measure the progress in the understanding of our subject matter by evaluating the usefulness and validity of our terms and concepts.

Now, encoding may exist as a concept, but the bigger question is whether encoding also exists in the world, outside some individuals' imaginations. Is it a real part of something that happens even if there are no sentient beings who think about it? A falling tree in the woods makes no sound if there is no living soul around to hear it—sound is the product of an appropriately endowed nervous system—but it does produce compressions of the air regardless of the presence of witnesses.

It is appropriate to raise the question about the existence of encoding in the present context because our friend and colleague, Fergus Craik, whose brilliant scientific career this volume is meant to celebrate, has repeatedly expressed doubts about such existence. Others may use the term designating the concept, but he does not believe it corresponds to anything in reality. Encoding is, our friend implies, memory science's élan vital, a thought in the minds of some thinkers, but it does not exist in reality. What does exist is perception, different kinds of perception at that, and the term *encoding* is mistakenly applied to some of these kinds.

Is our friend right? If so, it follows that those others who believe in encoding must be wrong. The question of who is right and who wrong is obviously of some interest, because we cannot expect to get very far in our pursuits of the truth about memory unless we solve the encoding problem first. I am not suggesting that the encoding problem is the only hurdle to be crossed before progress in our understanding of memory can occur, but I am suggesting it is among the very basic ones.

The issue before us, then, is this: Is there something in nature that corresponds to the idea that (many) students of memory have and that they designate as encoding and that is different from *perception*? Tackling the issue also means asking how we can find out. How do we go about establishing the existence of something that is invisible, inaudible, and intangible, and how do we determine its "basic" difference from another such? And, most germane to the occasion, is there a way of persuading Gus Craik to accept encoding as a part of reality?

This essay consists of five parts:

1. A brief history of the science of memory, the concept of encoding, and Gus's place in it.
2. *Byproduct theory* of trace formation: What is it and why is it postulated?
3. GAPS (general abstract processing system): Summary of the opposing view. What is it, how does it differ from the byproduct theory, and why is it postulated?
4. Functional neuroimaging and the *novelty encoding hypothesis*: Fleshing out the encoding's skeleton.
5. These four main sections are followed by a concluding statement.

## ☐ A Brief History of the Science of Memory

As this is a historic occasion, let us begin by placing our friend's thoughts about encoding into some kind of a historical perspective. The following thumbnail sketch of the history of the psychological science of memory, borrowed from Tulving (1993), is extremely brief and undoubtedly biased, but it serves the current purpose. For a fuller and more objective view, see Bower (2000).

The short history of the study of human memory can be divided into four successive stages, each stage bringing in innovations not known in the preceding one. The stages are shown in Table 2.1. The first one began in antiquity and ended in 1885. In that year, the second stage was launched by the publication of Herman Ebbinghaus's *Über das Gedächtnis* (Ebbinghaus, 1885). It lasted some 75 years, to around 1960. During that time the field came to be dominated by English-speaking scientists, especially in America. Problems of memory were pursued by experimental psychologists under the general rubric of *verbal learning*. Most of the research activity had to do with precise measurement of basic phenomena of learning and forgetting of lists of verbal materials under tightly controlled experimental conditions. The central theoretical concepts were *association* and its single property, *strength*. Theories consisted of attempts to explain observed facts in terms of

**TABLE 2.1. Four successive stages of the science of memory**

| | |
|---|---|
| Metaphysics | From the early Greeks on |
| Verbal learning | After 1885 |
| Information processing | After circa 1960 |
| Cognitive neuroscience | After circa 1980 |

acquisition, retention, and transfer of, interference with, and mediation by associations. In America, the dissenting voice of Frederic Bartlett (1932) from across the Atlantic was not heard.

In the years around 1960, the associative verbal-learning framework saw a rival emerging in the form of the *information processing* paradigm. A much wider variety of problems, issues, approaches, methods, and theoretical interpretations was adopted. Paired-associate and serial learning were largely abandoned in favor of free and cued recall, as well as recognition and various kinds of *memory judgments*, such as recency and frequency. To-be-learned lists that had served as units of experimental and theoretical analysis were replaced by to-be-remembered (TBR) single items, and single-trial memory studies became popular. Associations among single items were studied under the rubric of *organization*. Later, the TBR items came to be thought of as TBR *events*. Experimental studies of short-term memory led to a theoretical distinction between primary (short-term) and secondary (long-term) memory. Experimental findings were interpreted in terms of processes such as encoding, decision, storage, and retrieval. The distinction between storage and retrieval became a significant experimental and theoretical concern. Influential concepts such as levels of processing, encoding specificity, and encoding/retrieval interactions emerged during this stage, as did *context* and *context effects*. Connections were established between the previously isolated disciplines of cognitive psychology and neuropsychology.

The concept of encoding found its way into the verbal learning and memory literature through the back door, so to speak. James Deese, as well as some other participants at a verbal learning conference organized by Charles Cofer in 1959 (Cofer, 1961), used the term encoding in the sense of *response integration*. The field was still under a strong Ebbinghausian influence, and the use of nonsense syllables as learning materials was widespread. Such materials, according to then-prevailing views, had to be "integrated" or "unitized" before they could enter into associative relations, and Deese used the term encoding to refer to this process. One of his original proposals was that encoding was an all or none feature of verbal units and did not covary with other features of words, such as their frequency of occurrence in natural language. A low-frequency word such as "giraffe" was as well encoded as a high-frequency word such as "dog." These were novel ideas in 1959, and vigorously debated at the conference.

Deese's encoding, in its intended meaning, did not make it into the permanent literature. However, it did influence the thinking of Arthur Melton (1963), who, at the next Cofer conference 2 years later (Cofer & Musgrave, 1963), used the concept in discussing the recently discovered phenomena of short-term memory (Brown, 1958; Peterson & Peterson, 1959). The distinction between perceptual traces, associated with "immediate" memory, and memory traces, associated with "delayed" memory, was under discussion. Melton accepted the new reality of short- and long-term memory, but rejected any notion of a sharp discontinuity between them: "I suspect that there is no fundamental dichotomy between perceptual traces and memory traces, but rather a continuum of *rates of encoding* sensory information into memory traces based on previous experience and the compatibility of the experience with such previous experience" (p. 359).

Melton's "encoding sensory information into memory traces" was a harbinger of things to come in at least two senses. First, *memory trace* was a new concept, meant to take the place that had been occupied by *strengthened associations*. Second, it introduced the concept of encoding in the sense in which it has become an inseparable part of the memory world today. The new ideas spread quickly. By 1972, the concept of encoding was so well ensconced on the verbal learning and memory scene that Arthur Melton and Edwin Martin saw a need to edit a book on the topic entitled *Coding Processes in Human Memory* (Melton & Martin, 1972). In his chapter in that book, Gordon Bower called encoding "the truly central concept in modern theories of memory" (Bower, 1972, p. 85).

The current era of research, beginning some time around 1980, could be labelled the *cognitive neuroscience of memory*, although not every student of memory would accept this designation. It is characterized by further expansion and liberalization of methods, techniques, and choices of questions and problems that had been introduced under the banner of information processing. Like the previous shifts, it did not replace the reigning practices. It was not a Kuhnian paradigm shift—the science of memory is still waiting for its first paradigm—but rather an expansion of the domain, initially methodologically and conceptually, and eventually in terms of fact and theory. Among the central concepts of the era so far have been *priming, implicit memory, multiple memory systems, and functional neuroimaging*. There has been a steadily growing convergence between cognitive psychology and neuropsychology, interest has deepened in the study of memory processes in memory-impaired patients, more attention is being paid to autobiographical memory in real life and to its development over the life-span, theoretically motivated and precisely controlled psychopharmacological studies of memory have begun to appear on the scene, mathematical and computational modelling of memory processes has become increasingly sophisticated, and functional neuroimaging approaches to the study of memory are gradually overcoming their initial teething problems and beginning to generate large amounts of data. What was once the private sandbox of psychology has become the playing field of a larger collection of brain sciences. The unprecedented flourish in research on human memory was paralleled by increasingly vigorous investigations of learning and memory in nonhuman animals at the levels of behavior, neuroanatomy, and neurobiology. Today the multidisciplinary study of memory has taken off and is proceeding with abandon.

Gus Craik discovered memory as a possible scientific love object near the middle of the information processing epoch, safely outside the era of the (not terribly exciting) research on paired-associate learning and transfer. Like all of us, his approach to memory was shaped by a myriad of unfathomable forces. But three basic background factors, shared by many at the time, seem to have played a dominant role in shaping his thinking.

First, psychology was an autonomous scientific discipline that did not have to learn from any other and was best kept from the possibly corrupting influences of other disciplines, such as neuropsychology or physiology. Every graduate student of experimental psychology at the time knew that there was nothing useful that one could learn from physiology and that smart experimentalists stayed away from brain-damaged amnesic patients, because we had problems enough with more-or-less normal sophomores and their minds and did not need any extra troubles.

Second, the greatest virtue of any scientific idea or theory was its parsimony. The ultimate psychological bliss would be a single-factor theory of the whole mind, or at most a two-factor theory. Any kind of admitted complication was an anathema, possibly a sign of mental weakness on the part of the thinker. Because of psychology's isolationist independence, few people had heard of Einstein's dictum that the scientist's duty is to simplify nature as much as possible but not more. As we had not heard of it, everybody was playing Procrustes with a vengeance.

Third, and more specifically with respect to his chosen field, our friend, at a safe distance from the American preoccupation with Ebbinghaus's legacy and closer to where Bartlett had reflected on perceiving and remembering, had learned to think of memory in terms of what we know about it from everyday observations in real life. One of the most obvious things about real-life memory is that remembering occurs naturally and effortlessly. School learning aside, people walk through life and remember what happened to them without making any special attempt to commit any life experiences to memory. They perceive, they comprehend, they think, and then they remember, if not everything then at least many things. It is easy to imagine, then, that memory is an automatic byproduct of such perceiving, thinkings, and understandings. No special learning, or acquisition, or impressing-on-the-mind process is needed. Indeed, memory as a special, separate faculty of the mind is not needed. Memory is delayed perception of internal representations of earlier perceived stimuli.

## Byproduct Theory of Trace Formation

The famous levels-of-processing (LOP) paper by Craik and Lockhart (1972) presented a framework for the study of memory. A framework is like a theory, except it is usually broader than a theory and may be more vague. A framework in psychology is not unlike Thomas Kuhn's *paradigm* in mature sciences. A major difference is that a framework lacks the universal or near-universal acceptance by the practitioners in the field that a paradigm enjoys. The purpose of a framework, such as LOP, is to provide guidance for thinking about phenomena within its purview and to suggest ideas for research of a kind that would not be undertaken in the absence of the framework. History shows that LOP triumphed on both of these counts.

A major message of the LOP paper was "that the memory trace can be understood as a byproduct of perceptual analysis and that trace persistence is a positive function of the depth to which the stimulus has been analyzed" (Craik & Lockhart, 1972, p. 671).

Further: "Retention is a function of depth, and various factors, such as the amount of attention devoted to a stimulus, its compatibility with the analyzing structures, and the processing time available, will determine the depth to which it is processed" (Craik & Lockhart, 1972, p. 676).

The most important consequences of the framework are spelled out in the concluding comments of the paper: "If the memory trace is viewed as a byproduct of perceptual analysis, an important goal for future research will be to specify the memorial consequences of various types of perceptual operations" (Craik & Lockhart, 1972, p. 676).

Note the presence here of several basic concepts: memory trace, perceptual analysis, and depth of processing. And note the central idea that *"the memory trace can be understood as a byproduct of perceptual analysis."* No explicit mention of encoding, anywhere, because it is not needed. Perceptual analysis automatically produces the memory trace. Perceived items can be analyzed shallowly, or deeply, as well as at levels between the two extremes, and the resultant trace has a strength, or *persistence*, that reflects the level of processing.

Gus Craik had, at the time of writing the LOP paper, already conducted some preliminary experiments (Craik, 1973) of the kind that later were published in a longer report (Craik & Tulving, 1975) on the effects on recall and recognition of experimentally manipulated *orienting tasks*. Others had experimented with orienting tasks (Postman & Adams, 1956; Saltzman, 1956). In the absence of an appropriate framework for thinking about the data, however, the findings, historically speaking, amounted to nothing. More influential were earlier studies on the effects of orienting tasks done by Jenkins and his students

(Hyde & Jenkins, 1969; Johnston & Jenkins, 1971) that paved the way for the LOP steamroller. The results of all these studies vividly illustrated the previously unknown and unimagined power of the orienting tasks. Perhaps even more important was their demonstration that the intention to learn, a sine qua non of learning in the Ebbinghausian tradition and one of the mainstays in verbal learning (McGeoch & Irion, 1952), was largely if not entirely irrelevant for subsequent recall or recognition. It was the realization of that fact, shown by both the Jenkins and the LOP experiments, that turned the verbal learning world, as it operated at the time, upside down.

Craik and Lockhart (1972) did not, of course, deny the existence of a hypothetical process whose product was the memory trace. In that respect, they were securely in the camp of the information processing orientation that had recently arrived on the scene (Tulving & Madigan, 1970). What they were arguing against, however, was a special, dedicated, critical process that transcended perception, or *perceptual analysis*, and that represented a property of the faculty or mental ability of memory. Indeed, our friend went so far as to deny the existence of any such special faculty. In an important paper (Craik, 1983) in which he reviewed the fate of the LOP ideas in their first 10 years, he wrote:

> In any event, the central idea is that memory is not a separate faculty in any sense, but is a reflection of processing carried out primarily for the purposes of perception and comprehension, with certain types of processing (typically, richer, more elaborate, and more meaningful encodings) being associated with higher levels of retention. (p. 343)

For other observers of the memory scene, this kind of thinking was getting serious. Here was a statement whose implications went far beyond the earlier claims. For a card-carrying information processor it may have been possible to live without encoding, by simply assuming that Craik's perceptual analysis was simply a different term for the same concept that others labelled encoding. But the denial of the status of a separate faculty to the most important part of the human mind was bordering on treason! We were told, in essence, that not only is there no encoding, and no need whatsoever to postulate it, there was no such thing as memory! What we think is memory is nothing more nor less than a byproduct of perception. Among other things, if our friend was right, then those of us who thought we had spent our lifetimes (or good chunks of it) in studying memory had to change our self-concepts and ruefully admit that we had done no such thing. In reality, we had been doing nothing more than dabbling with byproducts of perception!

Was this kind of seditious thinking just a matter of Gus's romantic indiscretion? Evidence shows that while he may have shown ambivalence on the issue—one can, if one tries hard, detect signs of wavering now and then—the fact is that as late as 1999 he lovingly quoted the earlier lines:

> In our 1972 paper Lockhart and I stressed that the "levels" we were talking about were not "levels of memory" but rather levels of general cognitive processing, running from early data-driven sensory analysis to later conceptually-driven analysis of meaning and implication. By this view there is no self-contained "memory module": rather, memory is a byproduct of the general processes of perception and comprehension. Attention and memory are intimately linked in this model, since it is attentional processes that largely determine how deeply a stimulus is processed." (Craik, 1999, p. 101)

At this point he could have informed the reader that these views were those he (and Lockhart) held in 1972, that in the meantime the world had changed, and that his early romantic views no longer seemed reasonable. But there was no such renunciation. The flames of early loves burn, if not forever, then at least for a very long, long time.

If Gus Craik were completely alone in his views, we might try to solve the problem by simply ignoring him whenever he departs from the party line. But he is not. To pick an example not entirely at random: Gus's friend and colleague Boris Velichkovsky has recently expressed enthusiastic agreement with Craik and Lockhart's position that "memory performance is a byproduct of perceptual-cognitive processing" (Velichkovsky, 1999, p. 203). And the mutinous sentiments are spreading to the younger generation, some of whom seem to have embraced the "parsimony-at-any-cost" vision of science. For example, Randy Buckner and colleagues, who are well known for their work in functional imaging of memory, approvingly mention the byproduct theory of trace formation, although they may have developed the concept independently of Craik and Lockhart (Buckner, Kelley, & Peterson, 1999). In the context of our present discussion, and to remind our friend, however, the important point is that functional neuroimaging data, too, point to the existence of special, postperceptual, memory-related encoding processes.

In the present context, too, it is worth noting that Gus Craik's one-time student, Michael Watkins, has gone on record arguing against not just encoding, but all other postulated, "reified" happenings in the mind, including memory traces (Watkins, 2001). Children, in their rebellion against the parents, frequently take the road of greatly exaggerating what they perceive as parental shortcomings. Craik would not accept any special memory-related process at the input side, such as encoding, but otherwise has always adhered to the standard cognitive-psychology line on memory. Watkins goes further and would banish from the psychological science of memory every term that hints at any kind of contents of the black box of the mind. For him, the mind ("consciousness," he would say) is the object of study of psychology, and because psychology is a sovereign science, and not other natural sciences' handmaiden, phenomena of memory are to be described independently of anything that is not mind and not psychology. The worst thing that students of memory can do is to try to explain their findings in terms of hypothetical constructs or neural correlates, because these things are not part of psychology (Watkins, 2001).

We will return to the byproduct theory, to see whether it has any merit, later in the chapter. Next, however, let us consider an alternative to it.

## ☐ General Abstract Processing System: GAPS

Some time ago I described "a conceptual framework for the study and understanding of episodic memory" (Tulving, 1983, p. 129) that I called the general abstract processing system, or GAPS for short. Its purpose was to help to integrate the diverse research findings, suggest theoretical interpretations of the data, and relate the laboratory work on memory to remembering as experienced by people in real life.

Some people think GAPS was "Tulving's theory," but it was neither mine nor a theory. It was my way of summarizing the broad features of the experimental doings and theoretical thinkings of the practitioners in the information processing approach to memory.

The GAPS framework of 1983 held that an act of memory begins with encoding processes and ends with the use, or retrieval, of the encoded information. In total it consisted of 13 conceptual components, or elements: four observables, four hypothetical processes, and five hypothetical states. One of the central elements was encoding, defined as the process that converts an event into an engram, or memory trace. The other central concepts were engram, ecphory, and ecphoric information. Encoding was defined as the process that converts the event-information (provided by perception or thought) into an engram (memory trace); ecphory is the process that combines the information in the engram and the retrieval cue into ecphoric information. Ecphoric information determines recollective experience, the end product of an act of cognitive memory. Encoding and engram are the principal components of storage of information in memory; ecphory and ecphoric information are the principal components of retrieval of the stored information.

Note that the sequentially cumulative organization of the elements of (episodic) memory in GAPS is meant to parallel the same kind of organization in real-life memory situations, in which any given process necessarily reflects or depends on something that has happened earlier. Encoding depends on the perceived event and the cognitive environment, the original engram depends on the event and its encoding, and so on through the totality of an act of remembering, ending up, for cognitively important purposes, with conscious recollection of the event that depends on everything that has preceded it in the given sequence.

At any rate, whereas in the byproduct theory memory traces result from perceptual analysis, in GAPS they result from the encoding process. The question, then, is whether there is a difference, other than semantic, between the two hypotheses, and if so, exactly wherein it lies.

The difficulty in answering this question lies in the fact that both theories are highly similar in that both do postulate a hypothetical (not directly observable) process that begins with the act of perception and ends with the formation of the memory trace. They just name the process differently: perceptual analysis versus encoding. Is that all there is to it?

One can, of course, try to differentiate the two ideas by pointing to specific empirical facts and say, "explain this in terms of your ideas," or even better, "tell me again how you can explain this without invoking my ideas," and watch the opponent tumble. This procedure sounds useful in principle—what is more noble than adjudicating scientific disputes by facts?—but it does not work in practice, as all veteran scientists know.

One difference between the ideas of perceptual analysis and encoding is that the latter postulates the existence of a process that is more than, or transcends, "mere" perceptual analysis. That is, perceptual analysis may be necessary for trace formation, but it is not sufficient. Trace formation requires something like an act of perception (the perceived object or event, of course, could be either external or internal), plus something else that follows it or is superimposed on it, namely encoding. Perception, or perceptual analysis, of an event may result in trace formation (when it leads to encoding), but it need not (when it does not).

We will now turn to some examples of phenomena of memory that would be difficult to make sense of without postulating some process like encoding that intervenes between the act of perception of an item or an event and the formation of the trace of that item or event. Needless to say, many more such examples could be given, otherwise the concept of encoding would not be as widely used as it actually is. For the purposes of the exercise that follows, we ignore the retrieval side of memory processes, although retrieval processes are at least as important, if not more so, than encoding processes (Roediger & McDermott, 2000). Here we are concerned with the issue of trace formation and the roles that perception (comprehension, attention) alone versus perception (comprehension, attention) plus encoding play in it.

## Empirical Difficulties for the Byproduct Theory?

### Interrupted Encoding

I reported a study some time ago that I suggested demonstrated *retrograde amnesia* in normal healthy university students (Tulving, 1969). (Replications and extensions of the study have been described by Schulz, 1971 and Guynn & Roediger, 1995.) Subjects studied lists of 15 items. Experimental lists consisted of 14 familiar words and 1 name of a famous person (such as Freud, Churchill, or Beethoven). Control lists contained familiar words only, and no 'high-priority events' occurred. A large number of these two kinds of lists were presented in haphazard order. Subjects were tested for immediate free recall after each list. They had been instructed to be on the lookout for the names of famous persons. Whenever one occurred in a list, it was to be produced first in recall, before any other words.

The results in the condition in which the high-priority event appeared in the middle of the list, in serial position 8, are shown in Figure 2.1. The level of recall of the high-priority event was, as expected, very high. But the interesting finding was that the presentation and recall of the high-priority event was associated with rather low recall of the one or two words that had immediately preceded the high-priority event in the presentation sequence. I used the somewhat whimsical appellation of retrograde amnesia to describe this outcome: A highly conspicuous and memorable event seems to have suppressed the memory for events immediately preceding it.

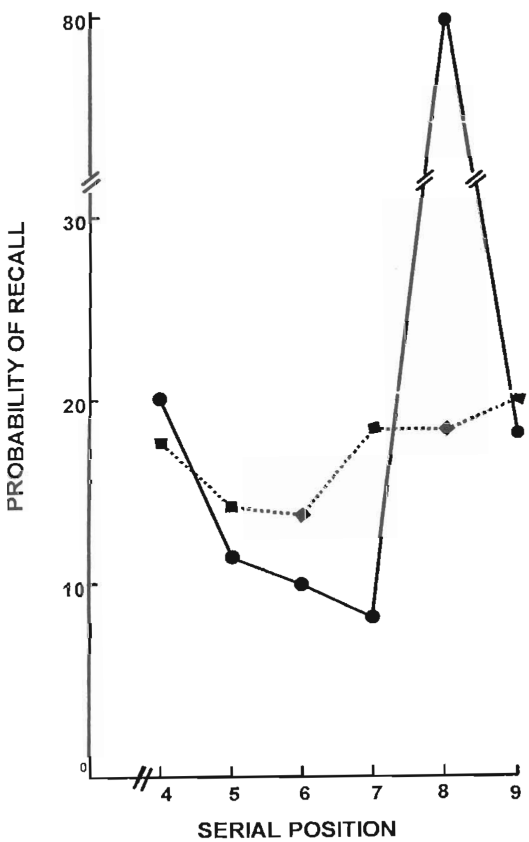


**FIGURE 2.1.** Probability of recall of words as a function of serial position in a single-trial free-recall task in which a high-priority event (e.g., the name of a famous person) occurred in serial position 8. (Data reported by Tulving, 1969.)

An additional finding was that the results as described were obtained only with very fast rates of presentations of the words, namely .5 seconds or 1 second per word. When the rate of presentation was slowed to 2 seconds per word, the effect disappeared (Tulving, 1969).

How are we to interpret, or make sense of, these miniature retrograde amnesia effects. The interpretation offered at the time still seems viable: Some kind of an encoding or consolidation-like process occurs for each perceived (read) item. When the presentation rate is slow, these processes run their normal course, and, together with other processes of the kind specified in GAPS, determine the item's subsequent recall (and recognition, as in Schultz's study). With the fast presentation rate, however, the encoding processes for successive items overlap: Processing of one item is not yet completed when the next item is presented. When a high-priority event occurs, it is immediately recognized as such, and the relevant processing resources (a nice Craikian term!) are withdrawn from items "on line," and devoted to the encoding of the high-priority event. In other words, the onset of the high-priority item prematurely terminates the encoding of the immediately preceding item and therefore impairs its trace formation.

Can one argue that the retrograde amnesia results from inadequate perception of items? The fact that the effect disappeared with slower rates of presentation, which was also observed by Guynn and Roediger (1995), could mean that the subjects needed more time than 1 second per word for the words' full perception and comprehension. One can, of course, come up with this kind of an explanation, especially if one is really desperate, but I do not think that Gus would take this low road.

## Remembering Events That Never Happened

Consider another difference between the byproduct and encoding theories. If perceptual analysis of an object or an event determines the trace of the object or event, it is easy enough to see how forgetting occurs, or even distortions in what is being remembered. But it would be more difficult to see how it is ever possible for anyone to remember an event that never happened and therefore could not have been originally perceived, because a nonperceived event could not have any byproduct; that is, it would leave no trace in the memory system. If retrieval, even of events that never happened, requires a trace, absence of the trace would make retrieval impossible. Thus it looks as if the byproduct theory could not say anything very useful about those aspects of *constructive memory* that consist in memory distortions, memory illusions, and false memories, on which a rich literature is now available, ably reviewed by Roediger and McDermott (2000). The problem again is that trace formation is too rigidly tied to perception.

The GAPS scheme is different. Perception plays an important role in trace formation, of course, but there is more to trace formation than just perception and comprehension of what is being perceived at input. At retrieval, the relevant trace information must be available, otherwise retrieval could not occur, as it could not occur in the absence of relevant retrieval cues. But the trace information need not have been originally perceived, just encoded as a part of the study task. Thus it is that GAPS even allows remembering something that never happened (Tulving, 1983, p. 183). In the early 1980s, psychologists were not greatly interested in false remembering, but the mere existence of errors such as false alarms in recognition and intrusions in recall made it necessary to postulate retrieval of originally nonperceived items.

Could one argue that perception and, especially, comprehension of lists like those that are used these days to demonstrate false remembering in the Deese–Roediger–McDermott (DRM) paradigm (Roediger & McDermott, 1995) are broad enough concepts to include the kinds of inferential processes that probably underlie remembering of events that never happened in the DRM paradigm and comparable situations? Of course one could. Psychology is a flexible science. The question is how far one is willing to extend the definition of "perception and comprehension," and why the extension is preferable to the idea of postperceptual encoding.

## Anterograde Amnesia

Direct evidence against the byproduct theory is provided by brain damaged patients with anterograde amnesia (Mayes, 2000; Squire, 1992). The brain damage that causes amnesia may involve different structures and pathways: medial temporal lobe and diencephalic, as well as other limbic-system regions (Markowitsch, 2000). In anterograde amnesia—a pathological impairment in the ability to remember ongoing happenings and to learn new information—perception occurs, but there is no memory byproduct. Why not? What is missing in the amnesic brain that is present in the normal one? It would be difficult to argue that what is missing is perception, or perceptual analysis, or comprehension, because typical amnesic patients have no problems in these spheres. Indeed, the clinical definition of amnesia is two-pronged: It specifies what is impaired (ability to remember ongoing experiences and to acquire new knowledge) and what is preserved (general intelligence, perceptual and conceptual abilities, language, reasoning, and even short-term memory). Neurological patients with memory impairments who also have other mental deficits are classified as demented, not as amnesic (Hodges, 2000).

Many students of amnesia today accept the idea that anterograde amnesia can be attributed to a faulty process of encoding information into long-term memory, that is, long-term memory trace formation (Milner, 1966; Squire, 1992). To the extent that this is so, we can think of anterograde amnesia as evidence that it is not faulty perception that is responsible for the failure of the establishment of memory traces.

## Psychopharmacology of Encoding

Our final example, although it does not exhaust the reservoir of possible examples, is provided by experiments on the effects of various amnestic drugs, such as benzodiazepines (BZs), on memory performance. In these experiments, subjects study lists of words, have no difficulty perceiving and comprehending the words, but subsequently show serious impairment in explicit retrieval of the words. These kinds of effects suggest that the drugs adversely affect some necessary postperceptual processing. Is it encoding? Or perhaps retrieval?

In a typical study of the kind that helps us to identify the source of the amnestic effects of the BZ drugs, subjects study two lists of words. Between the two study periods they are administered either the drug or a placebo. Then they are tested on two occasions: (a) while still under the influence of the drug, and (b) when the drug effects have worn off. The results of such studies show that (a) administering the drug before the learning of the list (severely) reduces recall performance, both in the same state (tested under the influence) and in the different state (tested sober), and (b) administering the test while the subjects are under the influence of the drug does not reduce the performance in comparison with the placebo condition. Conclusion: The drug's amnestic influence operates through its effect on some part of the encoding process. (For reviews, see Curran, 1991; Ghoneim & Mewaldt, 1990; Polster, 1993.)

Thus, here we have another case where it would be difficult to argue that trace formation is a byproduct of perceptual analysis. BZ drugs have little deleterious influence on perception, working memory, or retrieval of previously acquired long-term information, but do impair encoding, the process that converts the perceived information into the memory trace.

We have considered four examples: Interrupted processing reduces memory after perception, perfectly normal people remember events they did not perceive, brain-damaged people do not remember things they do perceive, and amnestic drugs knock out encoding of perceived items while leaving retrieval intact. How would a byproduct theorist deal with such apparently troublesome facts?

Gus Craik has not said anything about interrupted processing, or false memory, or drug effects on encoding. But he did briefly discuss the problem of anterograde amnesia. In the abstract of a paper based on a talk he gave at The Royal Society discussion meeting on memory he wrote:

> It is argued that memory is largely a function of depth and elaboration of the initial encoding, and that the memory deficits found in elderly people and under conditions of divided attention reflect impaired comprehension of the material. On the other hand, amnesic patients exhibit adequate comprehension yet poor memory, suggesting that some physiological process of consolidation may also be involved in normal learning and remembering. (Craik, 1983, p. 341)

And in the text of the paper, he wrote:

> Most cases of clinical amnesia are not well described by the present account, since these patients show at least adequate comprehension yet extremely poor subsequent recollection. (Craik, 1983, p. 356)

How did our friend deal with the trouble? Three devices are possible. Perhaps what manifests itself as amnesia is rooted in a deficit of conscious recollection. Perhaps it is rapid forgetting. And if neither, it may be a function of the breakdown of some physiological process, such as consolidation, but if so, it still leaves the psychological theory intact! In 1983 most of us psychologists did not pay much attention to physiology.

Regrettably, there is little support for the first of these two possible escape routes. Although Warrington and Weiskrantz (1970, 1974) did consider the idea that anterograde amnesia is a condition in which encoding is intact and retrieval is impaired, they themselves (Warrington & Weiskrantz, 1978) renounced the theory because it was not supported by data. Since no other champions of the retrieval theory of amnesia have emerged, Brenda Milner's (1966) original idea still holds: Amnesics such as HM show the horrendous memory deficit because of their inability to encode the information appropriately. (In the terminology of that time: They cannot get the information from the short-term memory store into the long-term store, according to the then-prevailing "modal models" [Murdock, 1967].)

The idea of amnesia as rapid forgetting would save the appearance of the byproduct theory, because it would mean that the requisite trace information was formed correctly, as a consequence of perception, but then lost. The obstacle here, however, is that there is little or no evidence for rapid forgetting in amnesia. Indeed, the evidence generated by studies directly designed to examine the issue shows that once amnesics have acquired the material to the same level as their controls, the subsequent forgetting is usually indistinguishable for the two populations. In one dramatic case, Freed and colleagues (Freed, Corkin, & Cohen, 1987) equated HM's initial learning of pictures of natural scenes with the initial level of acquisition of the same pictures by normal controls, by the simple device of allowing HM much more time to study the pictures, and then measured the retention performance over intervals extending to 1 week. The results showed that HM's retention performance was equivalent to that of the controls throughout the retention interval.

The last attempt by our friend to avoid facing the full consequences of anterograde amnesia for the byproduct theory, namely a firm separation between psychology and physiology, took the following form. In discussing the implications of anterograde amnesia for his theory, Gus Craik, acting in form as a real scientist, readily admitted that:

> Some "consolidation" process is necessary to "fix" in a permanently accessible fashion the changes in neural circuitry induced by the original event. . . . (Craik, 1983, p. 356)

But then he immediately continued:

> The evidence supporting consolidation is still incomplete, but there seems to be no reason in principle that a psychological account of remembering in terms of depth and elaboration of processing during initial perception and comprehension could not coexist with a more clearly physiological process of altering the underlying neural mechanism in some permanent fashion. (Craik, 1983, p. 356)

This was in 1983. Psychology and physiology were in separate domains, and it was natural for psychologists to ignore physiology then. All of us in experimental psychology did so. The "ignore the brain" stratagem may have worked back then, especially when the Zeitgeist of psychology allowed us to ignore physiological findings about memory, but it would be a less admired way to save a psychological theory today, because today physiology, under the cloak of functional neuroimaging, has become highly relevant to what we do as students of memory.

## Functional Brain Measures of Encoding

Everyone knows today that techniques of functional neuroimaging, positron emission tomography (PET), and functional magnetic resonance imaging (fMRI) can be used to localize cognitive function in the brain. Not everyone in cognitive psychology cares much about such localization, because knowledge of neuroanatomical sites that are involved does not help much with cognitive problems and cognitive theories. What difference does it make if, say, encoding is associated with neural activity in the left or right prefrontal cortex, or the frontal operculum, or the cerebellum, or some combination of these or many other areas?

Perhaps localization does not matter much, indeed. But functional neuroimaging does much more than just provide information on where in the brain "something happens." Among other things, it has contributed enormously to making consciousness, the central psychological "problem," scientifically respectable. It can also illuminate purely theoretical problems and help to settle them. With respect to the issue under scrutiny in this paper, functional neuroimaging research has turned out to be invaluable.

Let us go back to one of the basic premises of the present discussion. Everyone, including our friend Gus Craik, accepts the fact that encoding exists as an idea in many theorist's minds. It is a genuine part of the mental world, or World 2 as conceptualized by Popper and Eccles (1977). The uncertainty, and the ensuing debate, has to do with the question of whether encoding as a process "exists" or occurs in the world outside scientists' thought—the physical world, or Popper and Eccles's World 1? If so, is it possible to get evidence for such existence that is more direct than mere inferences from observed behavior? We know, of course, that we will never be able to get direct sensory evidence, because encoding processes are no more visible or tangible than entities such as seeing and feeling. Hence the question, Can we get *more direct* evidence than that suggested by observed behavior? The answer is yes, and such more direct evidence is provided by looking at what the brain does when some information is encoded into memory.

The original evidence of brain activity correlated with encoding did not have to wait until the advent of PET and fMRI. The very first time that the encoding process was identified at the level of brain activity was the event-related potentials (ERPs) experiment reported by Sanquist, Rohrbaugh, Syndulko, and Lindsley (1980). These investigators recorded ERP signals (scalp-recorded EEG voltage changes time-locked to stimulus events) in an explicit recognition experiment and found systematic differences between encoding of words that were subsequently recognized and encoding of those that were not. Since then, many others have replicated and extended these results (see Rugg, 1995, for review). Although ERPs provide fine information about temporal changes in neural activity, their spatial resolution is poor. From the point of view of localization, better evidence is provided by the findings from functional neuroimaging studies, and especially from those using the event-related fMRI technique (Buckner et al., 1996; Dale & Buckner, 1997). This technique, like ERPs, allows the experimenter to sort individual stimulus events (at study, or at test, or both) into categories according to their behavioral or cognitive fate, on the basis of any one of a variety of criteria, such as "recognized," "highly confidently recognized," "not recognized," "correctly rejected," or even "recognized but not remembered," and then examine the neural signatures of such categories, averaged over all stimulus events in the category.

The first event-related fMRI study of encoding, designed to compare the neural signatures of encoding of items that were subsequently remembered versus those that were subsequently forgotten, was reported by Wagner et al. (1998). In this study, subjects made abstract or concrete judgments about successively presented words that later appeared in a yes/no recognition test in which confidence ratings were also taken. Brewer, Zhao, Glover, and Gabrieli (1998) conducted a similar study, but used pictures (indoor and outdoor scenes) and examined only selected brain slices, whereas Wagner et al. (1998) looked at the whole brain. Wagner et al. (1998) found that neural activity *at encoding*, in specific brain regions, specifically left prefrontal and temporal cortices, was higher for studied words that were subsequently confidently recognized than for studied words that were not recognized. The activity at study in these regions, therefore, can be said to predict what happens to the studied items in the recognition test.

These studies thus have provided neural evidence of the existence of processes at the time of study that (a) determine the subsequent retrievability of the studied items and (b) are differentiated at the level of brain activity in specific cortical and subcortical regions. Does such different neural activity reflect differences in perception or differences in encoding? The data from a single experiment never answers broad questions of this sort, because a single experiment cannot control or manipulate all possibly relevant variables. But it seems more reasonable to attribute the differences to (postperceptual) encoding rather than to perception, for at least two reasons. First, there was no behavioral or cognitive evidence of any differences in perception or comprehension of the subsequently remembered and not remembered words. Second, and more important—perhaps even more interesting, in that it shows how 'localization of function' might turn out to be relevant even to purely cognitive issues, after all—was the finding that at least some of the regions of the brain in which the neural differences between remembered and nonremembered items were observed, such as the left prefrontal cortex, are better known as regions related to memory than to perception. Neuropsychological studies have shown that damage to frontal lobes can impair memory performance (for a review, see Wheeler, Stuss, & Tulving, 1995), yet few studies have shown that such damage impairs perception or comprehension. Functional neuroimaging studies, too, have consistently implicated the prefrontal cortex in memory processes, including encoding, but not in perception or comprehension (for an extensive review of relevant neuroimaging findings, see Cabeza & Nyberg, 2000).

Although the initial results of this kind of research, showing that circumscribed regional brain activity at study "predicts" the subsequent memorial fate of TBR items, seem to be more amenable to the interpretation in terms of encoding (that is, perception plus something else) than perception (and comprehension) alone, as in our other examples, they do not force an encoding interpretation. Nor do they totally exclude a perceptual byproduct kind of interpretation. The main significance of these results lies in the direction to which they point by way of future research. They imply that it should be possible to systematically control and manipulate relevant perceptual and attentional variables, as well as variables related to different aspects of postperceptual encoding processes, and observe the consequences of such manipulations at retrieval.

## ☐ Novelty Encoding Hypothesis

We have just seen how functional neuroimaging data can illuminate one of the most difficult problems that memory theorists have faced from the earliest days of memory research: What determines which particular items from a studied lot are and which ones are not remembered? The answer is that, partly at least, the memorial fate of an item depends on how it is dealt with by the brain at encoding—on its neural processing. But why are different items processed differently in the brain?

Many properties of items, verbal or otherwise, are known to be correlated with their memorability. Textbooks of learning and memory discuss them routinely. Here, in the last section of this chapter, let us consider an item variable whose effect on memorability is less well known. The variable is the item's novelty or familiarity to the rememberer (or just novelty, for easier reference). The story of novelty fits well into the present context, because it dissociates, in a sense, perception and memory: When a particular item is presented repeatedly, its perception is facilitated while its subsequent retrieval, at least in the recognition test, is reduced. Such a state of affairs has obvious implications for the byproduct theory and its rivals.

The relevant data come from experiments in which all properties of TBR items are held constant except one: situation-specific novelty. The general procedures are as follows: (a) Subjects are exposed to some stimulus items in the "prefamiliarization" phase to produce "familiar" items. Such prefamiliarization may take different forms, and items may be presented once or more times. (b) The subjects are then given a list of items to study for a subsequent memory test. The study list consists of both familiar items, as defined, and novel items of the same general class—items that the subjects have not encountered in the experiment. (c) They are then tested for their memory for the study list items. Typically this test is one of recognition. If so, in the recognition test there are both novel and familiar studied items, and novel and familiar distractors—items not encountered in the study list.

The typical results from experiments of this kind is that novel items are recognized more readily than familiar items. A summary of data from 13 different conditions in six different experiments is given in Table 1. These data come from experiments by Kinsbourne and George (1974), Tulving and Kroll (1995), Habib (2000), and Kormi-Nouri, Nilsson, and Ohta (2000). Stimulus items in all six experiments were single words, but otherwise the experiments differed with respect to details of the procedure. Thus, for example, Kinsbourne and George (1974) used both high and low background frequency words presented visually; Habib (2000) did the experiment in both the visual and auditory modality, and also varied the encoding task (deep versus shallow processing, as in a typical LOP study); Kormi-Nouri et al. (2000) prefamiliarized their subjects with the TBR materials, in different experiments, in three different kinds of subject-performed-tasks procedures,

**TABLE 2.2. Hit rates and false alarm rates in recognition of novel and familiar items from 13 experimental conditions in 4 experiments. See text for details.**

| Experiment | Condition | | Hits | | FAs | |
|---|---|---|---|---|---|---|
| | | | Nov | Fam | Nov | Fam |
| Kinsbourne | Hi F | | .64 | .54 | .36 | .46 |
| | Lo F | | .80 | .70 | .20 | .30 |
| Tulving | | | .77 | .67 | .21 | .47 |
| Kormi-Nouri | Exp 1 | V | .71 | .58 | .15 | .26 |
| | | N | .84 | .76 | .10 | .23 |
| | Exp 2 | V | .75 | .57 | .16 | .23 |
| | | N | .88 | .77 | .09 | .19 |
| | Exp 3 | V | .78 | .70 | .09 | .21 |
| | | N | .88 | .81 | .04 | .13 |
| Habib | Vis | D | .91 | .85 | .04 | .07 |
| | | S | .65 | .59 | .08 | .13 |
| | Aud | D | .90 | .90 | .03 | .09 |
| | | S | .80 | .70 | .10 | .14 |
| Average | | | .79 | .70 | .13 | .22 |

*Notes:* Hi F = high-frequency words; Lo F = low-frequency words; V = verbs; N = nouns; Vis = visually presented; Aud = auditorily presented; D = deep encoding; S = shallow encoding.

and measured recognition for both verbs and nouns that had occurred in the verbal descriptions of tasks performed (Nilsson, 2000).

In Table 2.2, hit rates (proportions of studied items that are correctly recognized) and false alarm rates (proportions of distractors that are incorrectly identified as studied items) are presented for each of the 13 conditions. The data are remarkably stable. Almost without exception, hit rates are higher for novel than familiar items, and false alarms are higher for familiar than novel items. The overall measure of goodness of recognition, the difference between hit rates and false alarm rates calculated over the means listed in Table 2.2 (and ignoring the differences in the numbers of individual observations that went into the means) is sizable: .66 for novel items and .48 for familiar items. In this kind of experiment at least, people find it easier to remember novel information than familiar information.

How could these kinds of data be interpreted in terms of the byproduct theory? The perception and comprehension of repeated items is presumably facilitated by repetition. Such facilitation is well known and has been thoroughly studied under the label of priming (Roediger & McDermott, 1993; Schacter, 1987; Tulving & Schacter, 1990). Yet they are less well recognized as items presented in a particular study list. Why? Why should a readily perceived event produce a less viable memory trace? I submit that the answer is not obvious.

It turns out that the answer is not obvious even if we postulate encoding as a postperceptual process that converts the perceived information into the memory trace. At the time when GAPS was proposed (Tulving, 1983), students of memory knew little about the effect of situational novelty on recognition. The fact that infrequent words are more readily identified as old in memory experiments than are frequent words was, of course, a well-known fact even then (Gregg, 1976), but this is a fact about a correlation between two variables, and correlations are usually open to a myriad of explanations. The importance of the single experiment that had demonstrated a causal effect of novelty on recognition memory (Kinsbourne & George, 1974) was not appreciated.

However, the encoding theory seems easier to revise in light of the facts regarding novelty than the byproduct theory does. It requires the postulation of an early stage of the encoding process that has to do with the assessment of novelty of incoming or on-line information. If the information is novel, it is transmitted for further encoding. If not, it is not further processed. Thus, instead of the perception–encoding sequence of processes in GAPS, we assume a perception–novelty-assessment–encoding sequence in the revised GAPS. (Note that because novelty assessment necessarily requires a comparison of incoming information with information already available in the memory store, the novelty-assessment process involves retrieval of some kind. If so, the revision of GAPS is a bit more complex than just the insertion of another process. But this is another story, to be told some other day.)

The idea that the brain assesses incoming information for its novelty or familiarity makes good sense, because it provides a simple, but I believe convincing, answer to a question frequently asked of and by students of memory: Are all the experiences that a person has stored in memory? Despite the fact that, logically speaking, the question cannot be answered (at least not negatively) by the methods of science, the problem has been repeatedly discussed (e.g., Loftus & Loftus, 1980). Although it is logically impossible to disprove the possibility of all perception becoming potentially retrievable from long-term memory, rational reflection, aided by what we know about biological evolution, suggests that the answer to the question is highly likely to be negative. It is, of course, clearly advantageous to a creature to have all kinds of knowledge available for use "on line," and generally it may also be true that the more such knowledge there is, the better off the creature is in surviving long enough to make more of its kind. The problem is that long-term storage of everything that ever happened requires enormous storage capacity and is very costly. The brain does have enormous capacity, but probably not enough to record, in detail, an individual's lifetime experiences and all the inferences that can be drawn from them (Dudai, 1997). Another problem with the idea of storing everything is that it is dumb, because it would mean storing loads and loads of information that is already available in the long-term store. One thing we know about nature is that it is not dumb.

One possible alternative is that everything is indeed stored in long-term memory, but then, from time to time—perhaps when the owner of the memory store is asleep—in order to alleviate the storage problem, the "contents" of the store are purged of redundant information or information that has not been used for some time. This idea, however, looks like a nonstarter. It reminds one of Rube Goldberg rather than nature: It is horrendously and unnecessarily complicated. There exists a much simpler solution. It is called *novelty assessment.*

The idea, also discussed under the rubric of *novelty detection,* is this: All incoming (perceptual and on-line) information is filtered through novelty assessment circuits of the brain and is selected for long-term encoding depending upon its novelty/familiarity status. The idea has been called the *novelty encoding hypothesis* (Tulving, Markowitsch, Kapur, Habib, & Houle, 1994). It holds that the (neural) novelty assessment mechanism is a component of the (neural) encoding process. This means, among other things, that the encoding process is not unitary, as it was assumed to be in GAPS in 1983, but consists of at least two sets of concatenated subprocesses. The first is novelty assessment, and the second involves higher level encoding operations. The end product of the concatenation is the engram or memory trace. Thus, on-line information, including that provided by the sensory and perceptual systems, is transmitted for higher level encoding depending upon its novelty: novel information receives preferential treatment over familiar information. Completely redundant information of a perceived and comprehended stimulus event is screened out from further processing, is not encoded, and hence leaves no long-term trace, a state of affairs that I alluded to earlier when I talked about the difference between "mere" perceptual analysis and encoding.

Although the concept of novelty is not simple—there are presumably many different kinds and forms of novelty—the idea of novelty encoding as a basic component of the encoding process itself is. At Toronto, we stumbled onto it when we started doing PET studies of memory (Tulving et al., 1994; Tulving, Markowitsch, Craik, Habib, & Houle, 1996), but similar ideas had been discussed earlier by other researchers and modellers (e.g., Knight, 1984; Kohonen, Oja, & Lehtio (1989); Metcalfe (1993); Siddle, Packer, Donchin, & Fabiani (1991); and Sokolov (1963); among others).

Several other functional neuroimaging studies have identified brain regions that are more active for novel than familiar information of various kinds: scenes (Constable et al., 2000; Stern et al., 1996), words (Saykin et al., 1999), or both scenes and words (Kirchhoff, Wagner, Maril, and Stern, 2000). These regions include the hippocampal formation in the medial temporal lobes, long known to be critically involved in memory processes (Milner, 1966; Squire, 1992), as well as other regions in the extended limbic system of which the hippocampal formation is a part (Tulving et al., 1996). Still other functional imaging studies have reported and focused on similar novelty regions in the left inferior prefrontal cortex but interpreted the data in terms of the concept of priming rather than novelty detection or novelty assessment (Buckner, Koutstaal, Schacter, & Rosen, 2000; Demb et al., 1995; Koutstaal et al., 2000; Wagner, Desmond, Demb, Glover, & Gabrieli, 1997; for a review of these studies see Habib, 2001).

The point, then, is this: Functional neuroimaging data suggest that the brain can respond differentially to novel as compared with familiar information. Why should it do so? One possibility is that such novelty detection is a part of the brain process that controls the encoding of information into long-term memory: Perceived information that is already available in the memory store is less likely to be encoded (converted into long-term traces) than information that is not yet available. The data summarized in Table 2.2 reflect this novelty encoding hypothesis, although there are many other "purely psychological" facts that do likewise and that can also be interpreted in terms of the evolved properties of the brain (Tulving et al., 1996).

## Conclusion

I began by praising Gus Craik as a brilliant explorer of the human mind, one who has fundamentally changed the way we think about memory, but also noting that he may have been wrong on one issue: the role of encoding in memory. "Craik's error" consisted in the idea that memory traces are a byproduct of perceptual processing of stimuli, and that it is not necessary to postulate the existence of an additional, separate, memory-related process of encoding.

This essay presents some reasons for believing that the concept of encoding corresponds to something in the physical or biological world. First, there are experimental findings that cannot readily be (or can only awkwardly be) explained without invoking processes that are "switched on" after perception and that, when they run their full course, produce a memory trace. Second, there are clinical observations of intelligent but brain-damaged amnesic patients, who perceive the world around them normally but who cannot remember later on what they perceived. Third, there are findings showing that certain psychoactive drugs, such as benzodiazepines, selectively impair the long-term encoding process while leaving perception, working memory, and even long-term retrieval intact. Fourth, there are functional brain imaging findings of encoding-related brain activity in brain regions not known for their involvement in perceptual processes. Finally, there is the question, Why should nature have evolved organisms that store in their long-term memory everything that they perceive? The answer here is that it has not. Organisms, including human beings, selectively encode and store incoming information on the basis of its novelty or familiarity. The strong form of this novelty encoding hypothesis holds that only nonredundant information is encoded and stored.

The evidence thus converges on a dissociation between perception and encoding. Some events are indeed perceived and encoded, or not perceived and therefore not encoded, as the byproduct theory would hold. Other events are perceived but not encoded, as in the case of anterograde amnesia. Or they are not perceived but nevertheless remembered, as in the case of false memory. Or their memorability may be inversely related to their perceivability, as in the case of novelty/familiarity of the TBR information. These and other similar dissociations are difficult to account for in terms of a single underlying process, and therefore pose problems for the byproduct theory. A theory that postulates perception and encoding as separate but related processes can make better sense of the facts.

Where does all this leave our friend and his encoding-less theorizing? In the long run, of course, it does not matter. Science as a self-correcting enterprise has a habit of proving all ideas that scientists have about what they observe wrong. It is only a question of time before both the byproduct theory and the GAPS framework, even if revised, become hopelessly inadequate. Moreover, even now it is not clear that the byproduct theory is wrong and the encoding theory right, or at least more so. When it comes to ontological problems (Does X exist even though we do not see it and cannot touch it?), there can be no decisive experiments producing incontrovertible data. There can be only decisive beliefs about controvertible data.

Therefore, the discussion here leaves our friend and esteemed colleague in the happy position of an exceptionally successful scientist who has the satisfaction of knowing that he was instrumental in bringing into our world one of the most influential systematic conceptual frameworks within which problems of memory can be raised and investigated and within which many findings can be accommodated and integrated. If the LOP framework does not cover *everything* about memory under the sun, and if there are aspects of memory to which it does not pertain, it simply means that LOP is like any other theory or framework: a part of the normal world of science. So be it.

## Acknowledgments

Preparation of this chapter was supported by an endowment by Anne and Max Tanenbaum in support of research in cognitive neuroscience and by a grant from the Natural Sciences and Engineering Research Council of Canada.

## References

Bartlett, F. C. (1932). *Remembering: A study in experimental and social psychology.* Cambridge, England: Cambridge University Press.

Bower, G. H. (1972). Stimulus-sampling theory of encoding variability. In A. W. Melton & E. Martin

(Eds.), *Coding processes in human memory* (pp. 85–124). Washington, DC: V. H. Winston & Sons.

Bower, G. H. (2000). A brief history of memory research. In E. Tulving & F. I. M. Craik (Eds.), *The Oxford handbook of memory* (pp. 3–32). New York: Oxford University Press.

Brewer, J. B., Zhao, Z., Glover, G. H., & Gabrieli, J. D. E. (1998). Making memories: Brain activity that predicts whether visual experiences will be remembered or forgotten. *Science, 281*, 1185–1187.

Brown, J. (1958). Some tests of the decay theory of immediate memory. *Quarterly Journal of Experimental Psychology, 10*, 12–21.

Buckner, R. L., Bandettini, P. A., O'Craven, K. M., Savoy, R. L., Petersen, S. E., Raichle, M. E., & Rosen, B. R. (1996). Detection of cortical activation during averaged single trials of a cognitive task using functional magnetic resonance imaging. *Proceedings of the National Academy of Science USA, 93*, 14878–14883.

Buckner, R. L., Kelley, W. H., & Peterson, S. E. (1999). Frontal cortex contributes to human memory formation. *Nature Neuroscience, 2*, 311–314.

Buckner, R. L., Koutstaal, W., Schacter, D. L., & Rosen, B. R. (2000). Functional MRI evidence for a role of frontal and inferior temporal cortex in amodal components of priming. *Brain, 123*, 620–640.

Cabeza, R., & Nyberg, L. (2000). Imaging cognition II: An empirical review of 275 PET and FMRI studies. *Journal of Cognitive Neuroscience, 12*, 1–47.

Cofer, C. N. (Ed.) (1961). *Verbal learning and verbal behavior*. New York: McGraw-Hill.

Cofer, C. N., & Musgrave, B. S. (Eds.). (1963). *Verbal behavior and learning*. New York: McGraw-Hill.

Constable, R. T., Carpentier, A., Pugh, K., Westerveld, M., Oszunar, Y., & Spencer, D. D. (2000). Investigation of the human hippocampal formation using a randomized event-related paradigm and z-shimmed functional MRI. *Neuroimage, 12*, 55–62.

Craik, F. I. M. (1973). A 'Levels of Analysis' view of memory. In P. Pliner, L. Krames, & T. M. Alloway (Eds.), *Communication and affect: Language and thought*. New York: Academic Press.

Craik, F. I. M. (1983). On the transfer of information from temporary to permanent memory. *Philosophical Transactions of the Royal Society of London, B302*, 341–359.

Craik, F. I. M. (1999). Levels of encoding and retrieval. In B. H. Challis & B. M. Velichkovsky (Eds.), *Stratification in cognition and consciousness* (pp. 97–101) Amsterdam: John Benjamins.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11*, 671–684.

Craik, F. I. M., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General, 104*, 268–294.

Curran, H. V. (1991). Benzodiazepines, memory and mood: A review. *Psychopharmacology, 105*, 1–8.

Dale, A. M., & Buckner, R. L. (1997). Selective averaging of rapidly presented individual trials using fMRI. *Human Brain Mapping, 5*, 329–340.

Demb, J. B., Desmond, J. E., Wagner, A. D., Vaidya, C. J., Glover, G. H., & Gabrieli, J. D. E. (1995). Semantic encoding and retrieval in the left inferior prefrontal cortex: A functional MRI study of task difficulty and process specificity. *Journal of Neuroscience, 15*, 5870–5878.

Dudai, Y. (1997). How big is human memory, or on being just useful enough. *Learning & Memory, 3*, 341–365.

Ebbinghaus, H. (1885). *Über das Gedächtnis*. Leipzig: Duncker & Humblot.

Freed, D. M., Corkin, S., & Cohen, N. J. (1987). Forgetting in H. M.: A second look. *Neuropsychologia, 25*, 461–471.

Ghoneim, M. M., & Mewaldt, S. P. (1990) Benzodiazepines and human memory: A review. *Anesthesiology, 72*, 926–938.

Gregg, V. (1976). Word frequency, recognition and recall. In J. Brown (Ed.), *Recall and recognition* (pp. 183–216). London: Wiley.

Guynn, M. J., & Roediger, H. L. (1995). High priority event instructions affect implicit and explicit memory tests. *Psychological Research—Psychologische Forschung, 57*, 192–202.

Habib, R. (2001). On the relation between conceptual priming, neural priming, and novelty assessment. *Scandinavian Journal of Psychology, 42*, 187–195.

Habib, R. (2000). *Human novelty assessment systems: Positron emission tomography evidence*. Unpublished doctoral dissertation. University of Toronto.

Hodges, J. R. (2000). Memory in the dementias. In E. Tulving & F. I. M. Craik (Eds.), *The Oxford handbook of memory* (pp. 441–459). New York: Oxford University Press.

Hyde, T. S., & Jenkins, J. J. (1969). Differential effects of incidental tasks on the organization of recall of a list of highly associated words. *Journal of Experimental Psychology, 82,* 472–481.

Johnston, J. D., & Jenkins, J. J. (1971). Two more incidental tasks that differentially affect associative clustering in recall. *Journal of Experimental Psychology, 89,* 92–95.

Kinsbourne, M., & George, J. (1974). The mechanism of the word-frequency effect on recognition memory. *Journal of Verbal Learning and Verbal Behavior, 13,* 63–69.

Kirchhoff, B. A., Wagner, A. D., Maril, A., & Stern, C. E. (2000). Prefrontal-temporal circuitry for episodic encoding and subsequent memory. *Journal of Neuroscience, 20,* 6173–6180.

Knight, R. T. (1984). Decreased response to novel stimuli after prefrontal lesions in man. *Electroencephalography and Clinical Neurophysiology, 59,* 9–20.

Kohonen, T., Oja, E., & Lehtio, P. (1989). Storage and processing of information in distributed associative memory systems. In G. E. Hinton & J. A. Anderson (Eds.), *Parallel models of associative memory* (pp. 129–167). Hillsdale, NJ: Erlbaum.

Kormi-Nouri, R., Nilsson, L.-G., & Ohta, N. (2000, July). *The novelty effect: Support for the novelty/encoding hypothesis.* Paper presented at the XXVII International Congress of Psychology. Stockholm, Sweden.

Koutstaal, W., Wagner, A. D., Rotte, M., Maril, A., Buckner, R. L., & Schacter, D. L. (2000). Perceptual specificity in visual object priming: Functional magnetic resonance evidence for a laterality difference in fusiform cortex. *Neuropsychologia, 39,* 184–199.

Loftus, E. F., & Loftus, G. R. (1980). On the permanence of stored information in the human brain. *American Psychologist, 35,* 409–420.

Markowitsch, H. J. (2000). Neuroanatomy of memory. In E. Tulving & F. I. M. Craik (Eds.), *The Oxford handbook of memory* (pp. 465–484). New York: Oxford University Press.

Mayes, A. R. (2000). Selective memory disorders. In E. Tulving & F. I. M. Craik (Eds.), *The Oxford handbook of memory* (pp. 427–440). New York: Oxford University Press.

McGeoch, J. A., & Irion, A. L. (1952). *The psychology of human learning.* New York: Longmans, Green.

Melton, A. W. (1963). Discussion of Professor Deese's paper. In C. N. Cofer & B. S. Musgrave (Eds.), *Verbal behavior and learning* (pp. 355–361). New York: McGraw-Hill.

Melton, A. W., & Martin, E. (Eds.). (1972). *Coding processes in human memory.* Washington, DC: V. H. Winston & Sons.

Metcalfe, J. (1993). Novelty monitoring, metacognition, and control in a composite holographic associative recall model: Implications for Korsakoff amnesia. *Psychological Review, 100,* 3–22.

Milner, B. (1966). Amnesia following operation on the temporal lobes. In C. W. M. Whitty & O. L. Zangwill (Eds.), *Amnesia* (pp. 109–133). London: Butterworth.

Murdock, B. B., Jr. (1967). Recent developments in short-term memory. *British Journal of Psychology, 58,* 421–433.

Nilsson, L.-G. (2000). Remembering actions and words. In E. Tulving & F. I. M. Craik (Eds.), *The Oxford handbook of memory* (pp. 137–148). New York: Oxford University Press.

Peterson, L. R., & Peterson, M. J. (1959). Short-term retention of individual verbal items. *Journal of Experimental Psychology, 58,* 193–198.

Polster, M. R. (1993) Drug-induced amnesia: implications for cognitive neuropsychological investigations of memory. *Psychological Bulletin, 114,* 477–493.

Popper, K. R., & Eccles, J. C. (1977). *The self and its brain.* Berlin: Springer International.

Postman, L., & Adams, P. A. (1956). Studies in incidental learning IV. The interaction of orienting tasks and stimulus materials. *Journal of Experimental Psychology, 51,* 329–333.

Roediger, H. L., & McDermott, K. B. (1993). Implicit memory in normal human subjects. In H. Spinnler & F. Boller (Eds.), *Handbook of neuropsychology* (pp. 63–131). Amsterdam: Elsevier.

Roediger, H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 803–814.

Roediger, H. L., & McDermott, K. B. (2000). Distortions of memory. In E. Tulving & F. I. M. Craik (Eds.), *The Oxford handbook of memory* (pp. 149–162). New York: Oxford University Press.

Rugg, M. D. (1995). Event-related potential studies of human memory. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 789–801). Cambridge, MA: MIT Press.

Saltzman, I. J. (1956). Comparisons of incidental and intentional learning with different orienting tasks. *American Journal of Psychology, 69,* 274–277.

Sanquist, T. F., Rohrbaugh, J. W., Syndulko, K., & Lindsley, D. B. (1980). Electrocortical signs of levels of processing: Perceptual analysis and recognition memory. *Psychophysiology, 17*, 568–576.

Saykin, A. J., Johnson, S. C., Flashman, L. A., McAllister, T. W., Sparling, M., Darcey, T. M., Moritz, C. H., Guerin, S. J., Weaver, J., & Mamourian, A. (1999). Functional differentiation of medial temporal and frontal regions involved in processing novel and familiar words: An fMRI study. *Brain, 122*, 1963–1971.

Schacter, D. L. (1987). Implicit memory: History and current status. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13*, 501–518.

Schulz, L. S. (1971). Effects of high-priority events on recall and recognition of other events. *Journal of Verbal Learning and Verbal Behavior, 10*, 322–330.

Siddle, D. A. T., Packer, J. S., Donchin, E., & Fabiani, M. (1991). Mnemonic information processing. In J. R. Jennings & M. G. H. Coles (Eds), *Handbook of cognitive psychophysiology: Central and autonomic nervous system approaches* (pp. 449–510). Chichester, England: Wiley.

Sokolov, E. N. (1963). Higher nervous functions: The orienting reflex. *Annual Review of Physiology, 25*, 545–580.

Squire, L. R. (1992). Memory and the hippocampus: A synthesis from findings with rats, monkeys, and humans. *Psychological Review, 99*, 195–231.

Stern, C. E., Corkin, S., Gonzalez, R. G., Guimaraes, A. R., Baker, J. R., Jennings, P. J., Carr, C. A., Sugiura, R. M., Vedantham, V., & Rosen, B. R. (1996). The hippocampal formation participates in novel picture encoding: Evidence from functional magnetic resonance imaging. *Proceedings of the National Academy of Science USA, 93*, 8660–8665.

Tulving, E. (1969). Retrograde amnesia in free recall. *Science, 164*, 88–90.

Tulving, E. (1983). *Elements of episodic memory*. Oxford, England: Clarendon Press.

Tulving, E. (1993). Human memory. In P. Andersen, O. Hvalby, O. Paulsen, & B. Hökfelt (Eds.), *Memory concepts - 1993: Basic and clinical aspects* (pp. 27–45). Amsterdam: Elsevier.

Tulving, E., & Kroll, N. E. A. (1995). Novelty assessment in the brain and long-term memory encoding. *Psychonomic Bulletin and Review, 2*, 387–390.

Tulving, E., & Madigan, S. A. (1970). Memory and verbal learning. *Annual Review of Psychology, 21*, 437–484.

Tulving, E., Markowitsch, H. J., Craik, F. I. M., Habib, R., & Houle, S. (1996). Novelty and familiarity activations in PET studies of memory encoding and retrieval. *Cerebral Cortex, 6*, 71–79.

Tulving, E., Markowitsch, H. J., Kapur, S., Habib, R., & Houle, S. (1994). Novelty encoding networks in the human brain: Positron emission tomography data. *Neuroreport, 5*, 2525–2528.

Tulving, E., & Schacter, D. L. (1990). Priming and human memory systems. *Science, 247*, 301–306.

Velichkovsky, B. M. (1999). From levels of processing to stratification of cognition. In B. H. Challis & B. M. Velichkovsky (Eds.), *Stratification in cognition and consciousness* (pp. 203–226). Amsterdam: John Benjamins.

Wagner, A. D., Desmond, J. E., Demb, J. B., Glover, G. H., & Gabrieli, J. D. E. (1997). Semantic repetition priming for verbal and pictorial knowledge: A functional MRI study of left inferior prefrontal cortex. *Journal of Cognitive Neuroscience, 9*, 714–726.

Wagner, A. D., Schacter, D. L., Rotte, M., Koutstaal, W., Maril, A., Dale, A. M., Rosen, B. R., & Buckner, R. L. (1998). Verbal memory encoding: Brain activity predicts subsequent remembering and forgetting. *Science, 281*, 1188–1191.

Warrington, E. K., & Weiskrantz, L. (1970). Amnesic syndrome: Consolidation or retrieval? *Nature, 228*, 628–630.

Warrington, E. K., and Weiskrantz, L. (1974). The effect of prior learning on subsequent retention in amnesic patients. *Neuropsychologia, 12*, 419–428.

Warrington, E. K., & Weiskrantz, L. (1978). Further analysis of the prior learning effect in amnesic patients. *Neuropsychologia, 16*, 169–177.

Watkins, M.J. (2001). The modality effect and the gentle law of speech ascendancy: Protesting the tyranny of reified memory. In H. L. Roediger, J. S. Nairne, I. Neath, & A. M. Suprenant (Eds.), *The nature of remembering: Essays in honor of Robert G. Crowder*. Washington, DC: American Psychological Association.

Wheeler, M., Stuss, D. T., & Tulving, E. (1995). Frontal lobes and memory impairment. *Journal of the International Neuropsychological Society, 1*, 525–536.