41

Implicit and Explicit Memory and Learning

JOHN F. KIHLSTROM, JENNIFER DORFMAN, AND LILLIAN PARK

Learning and memory are inextricably intertwined. The capacity for learning presupposes an ability to retain the knowledge acquired through experience, while memory stores the background knowledge against which new learning takes place. During the dark years of radical behaviorism, when the concept of memory was deemed too mentalistic to be a proper subject of scientific study, research on human memory took the form of research on verbal learning (Schwartz & Reisberg 1991; Anderson 2000).

Explicit and Implicit Memory

In the earliest years of scientific psychology, research focused on immediate conscious experience, in the form of sensations and percepts analyzed first by psychophysicists such as Weber and Fechner and then by structuralists such as Wundt and Titchener. Wundt believed that "higher" mental processes were not amenable to experimental study. But Hermann von Ebbinghaus proved him wrong in 1885: by counting repetitions and calculating savings in relearning, Ebbinghaus invented the verbal-learning paradigm that has dominated the scientific study of memory ever since (Tulving & Craik 2000).

Principles of Conscious Recollection

For most of the century following Ebbinghaus, the psychology of memory was concerned with conscious recollection – with the ability to recall or recognize events that had occurred in the past. From this research has emerged a small set of principles that largely govern how human memory operates (Kihlstrom 1996).

- *stage analysis*: memories are analogous to books in a library, or the information contained within them: mental representations of events are encoded as memory traces, which are retained in memory storage and retrieved in the course of ongoing experience, thought, and action.
- *elaboration*: memory for an event is a function of the degree to which that event is related to pre-existing knowledge at the time of encoding.

- *organization*: memory is also a function of the degree to which events are related to each other.
- *time-dependency*: memory fades with time, mostly by virtue of interference among competing memory traces.
- *availability vs. accessibility*: Encoded memories, available in memory storage, may not be accessible when retrieval is attempted.
- *cue-dependency*: the probability of retrieving an event is a function of the informational value of cues provided at the time of retrieval.
- *encoding specificity*: retrieval is most effective when cues present at the time of retrieval match those processed at the time of encoding.
- *schematic processing*: events that are relevant to currently active beliefs, expectations, and attitudes are remembered better than those that are irrelevant; events that are incongruent with these mental schemata are remembered better than those that are congruent.
- *reconstruction*: memory reflects a mix of information contained in the memory trace and knowledge derived from other sources; in the final analysis, memories are beliefs, and remembering an event is more like writing a story from fragmentary notes than reading it from a book.
- *interpersonal*: remembering is an act of interpersonal communication as well as of information retrieval, and so memories are shaped by the social context in which remembering occurs.

Taxonomy of Memory and Knowledge

These principles apply to so-called "long-term" memory (William James called it *secondary memory*). But the domain of memory also includes modality-specific *sensory registers* (e.g., *iconic* and *echoic* memory), and *primary memory* (also known as *short-term or working* memory), which may operate on somewhat different principles. For example, the time-dependency characteristic of forgetting from the sensory registers is produced by decay and displacement rather than interference. Primary memory is sometimes viewed as a separate memory system from secondary memory; in other theories, primary memory is identified with representations stored in long-term memory that are currently in a state of activation. Primary or working memory is closely identified with consciousness.

The knowledge stored in long-term memory comes in two broad forms. *Declarative knowledge* constitutes our fund of factual knowledge, and can be represented by sentence-like propositions. *Procedural knowledge* consists of our cognitive repertoire of rules and skills, and can be represented by "if-then" structures known as productions. Within the domain of declarative knowledge, we can distinguish *episodic* memory, or autobiographical memory for events that have occurred in our personal past, and *semantic* memory, a sort of impersonal mental dictionary. Procedural knowledge can be further classified into motoric and perceptual-cognitive skills. The declarative–procedural distinction has its immediate origins in computer science and artificial intelligence, but can be traced back to Ryle's distinction (in *The Concept of Mind*, 1949) between "knowing that" and "knowing how," and Bergson's assertion (in *Matter and Memory*, 1911) that "the past survives in two forms" – as recollections and as habits. Episodic memory is what most people mean by "memory," as opposed to "knowledge."

Priming Effects in Amnesia and Normal Memory

For most of its history, the scientific study of episodic memory was concerned mostly with conscious recollection, to the extent that it was concerned with consciousness at all, and the notion of unconscious memory was relegated mostly to the Freudian fantasyland. But beginning in the 1960s, research began to suggest that the notion of unconscious memories was valid after all – if not in the Freudian form. Of particular interest were studies of patients with the *amnesic syndrome* associated with bilateral damage to the hippocampus and related structures in the medial temporal lobe, or to the mammillary bodies and related structures in the diencephalon. In 1968, Warrington and Weiskrantz reported an experiment in which amnesic patients were asked to study a list of familiar words. Compared with control subjects, the patients performed very poorly on standard tests of recall and recognition. However, when they were presented with three-letter stems or fragments, and asked simply to complete the cues with the first word that came to mind, amnesics and controls were equally likely to complete the cues with items from the studied list.

This is a *priming effect*, in which the processing of one item influences the processing of another item. In positive priming, the prime facilitates processing of the target; in negative priming, the prime inhibits processing of the target. In this instance, the priming effect indicates that the studied items were encoded in memory, retained in storage, and influenced performance on the completion test. The fact that equivalent levels of priming occurred in neurologically intact subjects, who remembered the priming episode normally, and amnesic patients, who had very poor memory, indicates that priming can be dissociated from conscious recollection. On the basis of such evidence as this, Schacter distinguished between two expressions of episodic memory: explicit and implicit (Schacter 1987). Explicit memory refers to conscious recollection of a past event, as exemplified by performance on recall and recognition tests. By contrast, implicit memory refers to any effect of an event on subsequent experience, thought, or action. Priming is, of course, just such an effect. The dissociation between priming and recall in amnesic patients indicates implicit memory can persist in the absence of explicit memory.

Spared priming in amnesic patients has now been confirmed in a host of studies, and has been extended to a wide variety of other forms of amnesia:

- anterograde and retrograde amnesia occurring as a consequence of electroconvulsive therapy (ECT) for depression;
- anterograde amnesia produced by general anesthesia administered to surgical patients (see also Kihlstrom & Cork, chapter 49);
- anterograde amnesia associated with conscious sedation in outpatient surgery (see also Kihlstrom & Cork, chapter 49);
- memory disorders observed in dementia, including Alzheimer's disease, as well as those encountered in normal aging;
- hypnotic and posthypnotic amnesia following appropriate suggestions to hypnotizable subjects;
- "functional" or "psychogenic" amnesias encountered in genuine cases of dissociative disorder, including dissociative amnesia, dissociative fugue, and the interpersonality amnesia of dissociative identity disorder (also known as multiple personality disorder).

In each of these cases, the memory disorder primarily impairs explicit memory and spares implicit, which is either wholly or relatively intact. It is in this sense that implicit memory persists in the absence of explicit memory. However, implicit memory can be observed in individuals with normal memory functions as well. For example, normal subjects show significant savings in relearning for items that they can neither recall nor recognize. And while elaboration is an important determinant of explicit memory, "depth of processing" has relatively little impact on many priming effects. In nonamnesic individuals implicit memory may be said to be independent of explicit memory, in that priming does not depend on whether the prime is consciously remembered.

The Vocabulary of Implicit Memory

In general, dissociations between explicit and implicit memory come in several forms (Richardson-Klavehn, Gardiner, & Java 1996). In *population dissociations*, a condition like amnesia or aging affects explicit memory, but not implicit memory. In *functional dissociations*, an experimental variable (like depth of processing) affects explicit memory but not implicit memory. Of course, the dissociations can also go the other way: shifting from auditory presentation at time of study to visual presentation at time of test can have a big effect on implicit memory, but relatively little effect on explicit memory. In *single dissociations*, a single variable effects one expression of memory, explicit or implicit, but not the other. In *double dissociations*, a single variable has opposite effects on explicit and implicit memory. Double dissociations are the "Holy Grail" of cognitive neuropsychology, because they provide compelling evidence that two functions, such as explicit and implicit memory, are mediated by separate cognitive modules or brain systems. But they are also exceedingly rare. Many ostensible double dissociations are more like *twin dissociations*, in which one variable affects explicit but not implicit memory, while another variable affects implicit but not explicit memory.

Implicit memory is usually tested with a priming task, but priming comes in a number of different forms. Most research has focused on *repetition priming*, in which the target of the priming test is a recapitulation, in whole or in part, of the prime itself. For example, subjects might study a word like *doctor* and then be asked to complete the stem *doc-* or the fragment *d-c-o-* with the first word that comes to mind, to identify the word *doctor* when presented against a noisy background, or to decide whether the letter string *doctor* is a legal word. But *semantic priming* effects can also be observed when subjects who have studied a word like *doctor* are asked to give free associations to cues like *nurse*, or to generate instances of categories like *occupations*. Repetition priming can be mediated by a *perception-based representation* that is limited to the physical attributes of the prime and its configuration in space and time, but semantic priming requires a *meaning-based* representation that includes information about the semantic and conceptual features of the prime. Semantic priming can be studied with the same tasks normally used to measure repetition priming, such as perceptual identification and lexical decision, provided that the target does not recapitulate the prime.

Explicit and implicit memory are sometimes referred to as "declarative" and "procedural" memory, or "declarative" and "nondeclarative" memory (Squire, Knowlton, & Musen 1993), respectively. The declarative-procedural distinction was initially based on the view that preserved learning in amnesia was limited to procedural knowledge such as cognitive and motor skills, and an interpretation of priming as procedural (*if-then*) in nature. But amnesic patients can acquire new declarative knowledge as well, provided that they do not have to remember the circumstances in which they learned it; and semantic priming is arguably mediated by semantic memory, which is a component of declarative knowledge. While some implicit expressions of memory may be mediated by procedural or nondeclarative knowledge, the declarative-nondeclarative distinction risks confusing the interpretation of explicit memories as representations that can be consciously "declared" with the propositional format in which declarative knowledge is represented.

Similarly, tests of explicit and implicit memory are sometimes referred to as "direct" and "indirect." That is to say, recall tests memory directly, while savings or priming test memory indirectly. It should be understood, though, that the direct–indirect distinction applies to memory tests and not to expressions of memory. In principle, priming could be used to assess consciously accessible memories that the subject declines to report, much as psychophysiological measures are used in forensic lie-detection. Along the same lines, explicit and implicit memory are sometimes referred to as "intentional" and "incidental" respectively. That is to say, in recall tests subjects are instructed to intentionally remember some past event, while priming occurs incidentally when the subject is performing some non-memory task. The intentional–incidental distinction reminds us that there are two aspects of consciousness relevant to memory: conscious awareness and conscious control (Butler & Berry 2001). A conscious memory might well emerge, unintentionally and inadvertently, in the course of a priming test – a situation that is sometimes referred to as "involuntary explicit memory." Involuntary conscious recollection has been a topic of literary discussion at least since the time of Proust (Salaman 1970), and of scientific investigation since the time of Galton (Crovitz 1970), but it should be distinguished from implicit memories that are inaccessible to conscious recollection in the first place.

In the final analysis, both the "direct–indirect" and "intentional–incidental" dichotomies fail to capture the essence of the explicit–implicit distinction – which is that explicit memory is conscious recollection, and implicit memory is unconscious memory, of the past. But if implicit memory is unconscious memory, why not simply call it that? The answer is more likely to be found in sociology than psychology, as those who would make a science of unconscious mental life have sought to avoid the taint of Freudian psychoanalysis. Even without the specter of Freud looming over their shoulders, the topic of consciousness still makes some psychologists nervous (Flanagan 1992). Still, what makes implicit memory interesting is not that implicit tests provide indirect, possibly surreptitious, assessments of what a person remembers; nor that implicit expressions of memory occur involuntarily. What makes implicit memory interesting is that it represents the dynamic influence of memory in the absence of conscious recollection.

Theories of Implicit Memory

Based on the "modularity" view popular in cognitive neuroscience, a number of theorists have suggested that explicit and implicit memory reflect the performance of separate memory systems in the brain (Schacter & Tulving 1994). For example, Squire has identified explicit memory with a medial temporal-lobe memory system including the hippocampus and related structures (Squire & Zola-Morgan 1991). Damage to this system will impair explicit memory but spare implicit memory, which is mediated by other brain systems, presumably cortical in nature. At the other end, Tulving and Schacter have proposed that repetition priming is mediated by a set of *perceptual representation systems* that store representations of the physical structure of the prime, but not its meaning (Tulving & Schacter 1990). For example, a *visual word form system* associated with the extrastriate cortex mediates visual stem-completion, while an *auditory word-form system* mediates auditory perceptual identification. Semantic priming, in turn, is held to be mediated by a separate semantic memory system.

Another take on the multiple-systems view is provided by Bowers and Marsolek (Bowers & Marsolek 2003). Instead of invoking multiple memory systems, they propose that implicit memory is a byproduct of brain systems that are devoted to perceptual pattern recognition, conceptual processing, and motor behavior, rather than memory per se. On their view, implicit memory is a byproduct of the learning capability of these systems. These systems have memory, in that they are capable of encoding and recognizing information, but they are not memory systems. Although Bowers and Marsolek's approach is based on contemporary theories of object recognition, psycholinguistics, and concept formation, it has its deeper roots in a proposal by Ewald Hering, the nineteenth-century sensory physiologist, that memory is "a universal function of all organized matter" (Hering 1870/1880, p. 63). Hering's ideas, in turn, were promoted by Samuel Butler, author of *Erewhon* (1872) and *The Way of All Flesh* (1903), in a ground-breaking book on *Unconscious Memory* (1880) that actually predated Ebbinghaus. Unconscious memory, on Hering's and Butler's view, may be likened to the "memory" of a paper clip – which, when once bent, is easier to bend again in the same direction. Paper clips do not have memory systems, but they do have a physical structure that allows them to retain traces of stimulation. Bowers and Marsolek do not have much to say about explicit memory, which presumably *is* mediated by a dedicated brain system.

By contrast with the multiple-systems view, other theories hold that explicit and implicit expressions of memory are the products of a single memory system. For example, Mandler's *activation view* argues that priming in all of its forms is mediated by the automatic activation and integration, at the time of encoding, of pre-existing knowledge structures corresponding to the prime; explicit memory, by contrast, requires effortful elaboration to establish new relations among activated structures (Mandler 1980). But activation, integration, and elaboration all take place within a single memory system. Roediger's *transfer-appropriate processing* view (Roediger & McDermott 1993) holds that most implicit memory tasks, such as repetition priming, are "perceptually driven," in that they require access only to surface features of an object; by contrast, explicit memory tasks are "conceptually driven," in that they require access to semantic or contextual information associated with the studied item. In this view, dissociations occur because explicit memory depends on "top-down" or "symbolic" processing, while implicit memory depends on "bottom-up" or "data-driven" processing.

Yet a third single-systems view invokes Jacoby's *process dissociation framework* to explain dissociations between explicit and implicit memory (Jacoby 1991). In this view, explicit memory is largely a product of conscious, controlled, effortful, deliberate processing, while implicit memory is largely a product of unconscious, automatic, effortless, involuntary processing. Jacoby has also introduced a method, the *process dissociation procedure* (PDP), which measures the relative contributions of automatic and controlled processing to any task by pitting them against each other in the "method of opposition" (MOP). A typical result of the PDP is to confirm that the performance of normal subjects on a memory task is mediated by a mix of controlled and automatic processes, while the performance of amnesic patients is largely supported by automatic processes.

Testing the Theories

Each of these views has its strengths and weaknesses, not least because they evolved in different research contexts. Multiple-systems theories are based largely on work with neurological patients, while single-system theories emerged mostly from work on neurologically intact subjects. The multiple-systems views bask in the reflected glory of cognitive neuroscience, but are bedeviled by the temptation to invoke a new memory system to explain every new dissociation revealed by research. The activation view gives a plausible account of priming results, but finds it difficult to explain how activation could persist for days or months – as it is sometimes observed to do. The transfer-appropriate processing view can explain not only dissociations between explicit and implicit memory, but also those that occur between two explicit or two implicit memory tasks (one perceptual, the other conceptual in nature), but has some difficulty explaining dissociations between semantic priming and explicit memory, both of which are, in its terms, conceptually driven. A further question is whether it is appropriate to term explicit memory as conceptually driven in the first place.

The PDP view, for its part, offers a way to reconcile single-system and multiple-system views, on the assumption that automatic and controlled processes are based on separate processing modules that operate on a single memory store. At the very least, it has provided an increasingly popular technique for measuring the contributions of automatic and controlled processes to task performance, and offers a way to reconcile single-system and multiple-system views. However, the mathematics of the PDP requires the troubling assumption that these processes are independent of each other. An alternative view, also consistent with a single-system view of memory, describes automatic processes as embedded in, and thus redundant with, controlled ones. For example, Mandler has proposed that the automatic activation and integration of stored information, which is the basis for implicit memory, precedes the effortful elaboration of these knowledge structures, which is the basis for explicit memory. In such a system, explicit memory entails implicit memory, even if the reverse is not the case.

One area where the various theories make competing predictions is with respect to implicit memory for novel, unfamiliar information. Activation theories would seem to suggest that this is not possible, because there is – by definition – no pre-existing knowledge structure stored in memory to be activated, or modified, by perceptual input. By contrast, the other theories are, at least in principle, open to the acquisition of new information. In fact, there is considerable evidence for priming of novel nonverbal items such as dot patterns and novel objects – though not, apparently, for line drawings of "impossible" objects that cannot exist in three-dimensional space (much like the drawings of the Swiss artist M. C. Escher). Although interpretation of these findings remains somewhat controversial, priming for novel stimuli would seem to be inconsistent with activation/modification views of implicit memory. Instead, they appear to support the multiple-systems view that repetition priming, at least, is the product of perceptual representation systems that encode and preserve structural descriptions of stimulus events. Priming does not occur for impossible objects because the perceptual representation system cannot form a structural description of objects that cannot exist in three-dimensional space.

The situation with respect to priming for verbal materials is more complicated. Early results, which showed priming for words such as *candy* and *number* (which have pre-existing representations in semantic memory) but not for pseudowords like *canber* and *numdy* (which do not), are consistent with the activation view of implicit memory. Bowers found priming for words (e.g., *kite*), nonwords that followed the rules of English orthography (e.g., *kers*) and for illegal nonwords (e.g., *xyks*), again contradicting the activation view. However, as Bowers himself noted, the priming he obtained for illegal nonwords may have been contaminated by explicit memory, which softens the blow somewhat. On the other hand, Dorfman found priming for pseudowords made up of familiar morphemes (e.g., *genvive*) and familiar syllables (e.g., *fasney*), but not for pseudosyllabic pseudowords (e.g., *erktofe*) made up of elements that are neither morphemes nor syllables in English. These results are consistent with the view that priming of novel (and familiar) words results from the activation and integration of pre-existing sublexical components stored in memory. Priming cannot occur where there are no such components to be activated.

The theoretical debate continues back and forth, but theoretical development is hampered by the fact that experimental research on implicit memory is narrowly focused on a single experimental paradigm – namely, repetition priming. It has been estimated that some 80 percent of implicit memory tests are perceptual in nature, involving variants on repetition priming. Viewed in this light, it is not surprising to find theorists proposing that implicit memory is the product of a perceptual representation system, or of perceptually based processing. But if implicit memory includes semantic priming, as well as repetition priming, such theories are too limited to account for the entire phenomenon. Repetition priming may be independent of depth of processing – although a more accurate statement would be that it is only *relatively* independent; but this is unlikely to be the case for semantic priming. Repetition priming may be modality specific – though not *hyperspecific*; but again, this is unlikely to be the case for semantic priming. Research on implicit memory must move beyond repetition priming if we are ever to determine its true nature.

Interactions Between Explicit and Implicit Memory

Owing largely to the hegemony of cognitive neuroscience, the most popular theory of implicit memory remains some version of the multiple-systems view. For example, Schacter and Tulving have proposed that while repetition priming is mediated by a perceptual representation system, semantic priming is mediated by a semantic memory system and spared procedural learning (see "Explicit and Implicit Learning," below) by a procedural memory system (Schacter & Tulving 1994). Even so, claims for a strict separation of these memory systems should not be made too strongly. If these various memory modules were truly independent of each other, we would expect to see neurological cases where explicit memory is spared and implicit memory impaired. The reverse, of course, is what is commonly observed in amnesia. In fact, only one case has been reported in which implicit memory is impaired and explicit memory intact – and that one is uncertain. The patient, known as M. S., who has an extensive scotoma secondary to brain surgery, performed poorly on a visual test of repetition priming, but normally on test of recognition – a reversal of the usual finding in amnesic patients. However, M. S. also showed normal performance on a test of conceptual priming, so it can hardly be said that he lacks implicit memory.

Whatever their underlying basis, the interaction between explicit and implicit memory can also be observed in other ways. Subjects who consciously recognize that the items on a perceptual-identification test come from a previously studied wordlist may develop a mental set that actually enhances their priming performance – which is why researchers in this area take care to assess "test awareness" in their subjects, and why Jacoby's "process dissociation" procedure has become so popular. Amnesic patients are not able to take advantage of explicit memory, of course, but that does not mean that conscious recollection cannot influence priming in other circumstances.

Moreover, there is considerable evidence that subjects can take strategic advantage of implicit memory to enhance their performance on tests of explicit memory. Although free recall epitomizes conscious recollection, both Mandler and Jacoby have argued that recognition judgments can be mediated by either conscious recollection of the test item, or by a feeling of familiarity that might be based on priming. If so, then when implicit memory is spared, subjects can strategically capitalize on the priming-based feeling of familiarity to enhance their performance on recognition tests. We know that, as a rule, recognition is superior to recall in normal subjects and this is also true for neurological patients with the amnesic syndrome, depressed patients receiving ECT, demented patients suffering from Alzheimer's disease, and normal subjects with posthypnotic amnesia. In addition, studies of recollective experience indicate that amnesic recognition is typically accompanied by intuitive feelings of familiarity, rather than full-fledged remembering.

Accordingly, it seems reasonable to suggest that successful recognition in amnesia can be mediated by spared implicit memory. This claim has been vigorously debated by Squire and his colleagues, who insist that priming is inaccessible to conscious awareness, and so cannot serve as a basis for recognition. Despite methodological issues cutting this way and that, studies employing the process-dissociation procedure clearly indicate that amnesic recognition is mediated by a priming-based feeling of familiarity – as theory suggests they might be, and as the subjects themselves say they are. It may be that recollection and familiarity are governed by separate memory systems; but against a further proliferation of memory systems, it may be more parsimonious to conclude that explicit and implicit memory interact after all.

Explicit and Implicit Learning

Traditionally, learning has been defined as a relatively permanent change in behavior that occurs as a result of experience. Early investigators – Pavlov and Thorndike, Watson and Skinner – construed learning as conditioning – the formation of associations between environmental stimuli and an organism's responses to them (Bower & Hilgard 1981; Schwartz & Reisberg 1991). However, the cognitive revolution in psychology has led to a reconstrual of learning as a relatively permanent change in *knowledge* that occurs as a result of experience – declarative and procedural knowledge that the organism will subsequently use for its own purposes in predicting and controlling environmental events. Thus, in classical conditioning the organism forms expectations concerning the likely consequences of events, and in instrumental conditioning the organism forms expectations concerning the likely consequences of its own behaviors. How this knowledge translates into behavior is another matter.

In addition to classical and instrumental conditioning, researchers have studied *perceptual learning*, involving long-lasting changes in perception or perceptual-motor coordination, as in the case of prism adaptation to inverted or distorted images; and *conceptual learning*, by which individuals induce abstract concepts from encounters with specific instances – not to mention language learning, and especially the learning of a second language. Although these forms of learning seem to be mediated by direct experience, Bandura has described *social learning*, also known as vicarious or observational learning, in which the individual gains knowledge by observing other people. Social learning comes in two broad forms: by example, through imitation and modeling; and by precept, through sponsored teaching.

At least in the case of humans (and certainly other primates, probably other mammals, perhaps other vertebrates, and maybe some invertebrates), this cognitive emphasis on individuals acquiring knowledge to help them predict and control events in the world implied that learning was a conscious activity. This is also true of perceptual learning, which occurs more rapidly with active than passive movements of the observer. And its emphasis on the role of observing, modeling, and teaching – not to mention the fact that civilization has created institutions to support these activities – marks social learning, too, as a conscious act of mind. Nevertheless, it is also true that some organisms, such as the sea mollusk *aplysia*, can learn even though they probably do not have enough neurons, much less a cerebral cortex, to support consciousness. Even in humans, who have a capacity for consciousness, it has long been evident that some learning can take place unconsciously (Adams 1957; Razran 1961).

Implicit Learning

The concept of *implicit learning* was introduced into the psychological literature well before that of implicit memory. In a pioneering series of experiments published in 1967, Reber asked subjects to memorize lists of letter strings, each of which had been generated by a Markov-process artificial grammar – a set of rules that specified what letters could appear in the string, and in what order. Over trials, the subjects found it easier to memorize grammatical strings, compared to random strings, indicating that their learning was exploiting the grammatical structure. Moreover, when presented with new strings, subjects were able to distinguish between grammatical and nongrammatical strings at levels significantly better than chance, indicating that they had acquired some knowledge of the grammar. Yet when queried, the subjects were unable to specify the grammatical rule itself. They had learned the grammar, and this knowledge had guided their behavior, but they were not aware that they had learned anything, and they were not aware of what they had learned (Reber 1993).

At roughly the same time, neuropsychologists noticed that, over trials, amnesic patients improved their performance on such tasks as maze learning, pursuit-rotor learning, and mirror-reversed learning. Clearly, then, amnesic patients had the capacity to acquire new skills, but they did not recognize the tasks, nor did they remember the learning experiences; moreover, they seemed to have no conscious awareness of their newly acquired knowledge. Later studies showed that amnesic patients could learn artificial grammars, just as neurologically intact individuals do.

By analogy with memory, we can define explicit learning as a relatively permanent change in knowledge or behavior that is accompanied by conscious awareness of what has been learned. Implicit learning, then, refers to a relatively permanent change in knowledge or behavior in the absence of conscious awareness of what has been learned. Sometimes evidence for implicit learning is taken as evidence for implicit memory, but implicit memory is more narrowly restricted to the learning *episode* itself, while implicit learning covers the knowledge acquired in that episode. In a famous case published in 1911, Claparede described an amnesic patient who forgot an episode in which he pricked her hand with a pin while greeting her, but who was consciously aware that "Sometimes people hide pins in their hands" (Kihlstrom 1995). This patient was conscious of what she had learned, but displayed *source amnesia* (also known as cryptomnesia, or unconscious plagiarism), a concept more closely related to implicit memory. Implicit learning goes beyond the formation of simple associations, as in classical or instrumental conditioning, and involves the acquisition of knowledge of some complexity, at some level of abstraction.

Varieties of Implicit Learning

Implicit learning has been studied in a wide variety of experimental paradigms, in addition to artificial grammars and motor learning (for comprehensive reviews, see Berry & Dienes 1993; Seger 1994; Stadler & Frensch 1998; Frensch & Runger 2003):

- Concepts: In a paradigm somewhat similar to artificial grammar learning, subjects learn to identify instances of novel concepts, such as patterns of dots that vary around a prototype, without being able to describe the defining or characteristic features of the concepts themselves.
- Covariation detection: subjects learn the association between two features, such as hair length and personality, but cannot identify the basis for their predictions.
- Sequence learning: subjects learn the sequence in which certain stimuli will occur – for example, the appearance of a target in a particular location on a computer screen – without being able to specify the sequence itself.
- Dynamic systems: subjects learn to control the output of a complex system by manipulating an input variable, without being able to specify the relationship between the two.

In each of these cases, subjects demonstrate, by performance measures such as accuracy or response latency in judgment, or prediction or control of behavior, that they have acquired knowledge from experience; yet they are unable to provide an accurate account of the methods by which they achieve these results. They have learned something new, but they do not know what they know.

What is Learned in Implicit Learning?

Observations of preserved learning capacity in amnesic patients led Cohen and Squire to propose that amnesia impaired declarative memory ("knowing *that*"), but spared procedural memory ("knowing how"; Cohen & Squire 1980). Perceptual and motor skills, such as covariation detection, sequence learning, and motor learning, can certainly be represented as systems of *if-then* productions, but it is not at all clear that all forms of implicit learning are procedural in nature. For example, there is evidence that amnesic patients can acquire new conceptual knowledge in amnesia – knowledge that, ordinarily, would be represented in the propositional format characteristic of declarative memory.

In this respect, artificial grammar learning is an interesting case. The structure of a Markov-process finite-state grammar lends itself easily to translation into a production system: *If* the first letter is a P, *then* the next letter must be a T or a V; *If* the first letter is a T, *then* the next letter must be an S or an X; and so on. However, subjects memorizing grammatical strings might simply abstract what a "prototypical" grammatical string looks like. When making grammaticality judgments, subjects could then compare test items to this stored prototype – or, perhaps, to the specific instances stored in memory during the memorization phase of the experiment. In either case, the unconscious knowledge acquired through implicit learning would more closely resemble declarative than procedural knowledge – yet another reason not to use "declarative" instead of "explicit" to label conscious memory.

In principle, both prototypes and rule systems are abstract knowledge representations that go beyond the specific instances encountered in the study set. Accordingly, an important question concerns the degree to which implicit learning is generalizable beyond the specific. That is, can subjects apply a grammar learned from strings of Ps, Vs, and Ts to test strings composed of Ls, Bs, and Ys? Reber reported that this was the case, although subsequent research has often found that transfer is substantially degraded. Studies of transfer in other domains have also yielded mixed results. Of course, the degree of transfer will depend on the degree of initial learning. In the artificial grammar experiments, classification performance typically ranges between 60 and 80 percent correct, with the more frequent outcomes at the lower end of this range. Viewed against a base rate of 50 percent correct, a performance at 65 percent may be statistically significant, but may not leave a lot of room to show incomplete transfer.

Is Implicit Learning Really Unconscious?

Implicit learning is distinct from mere *incidental learning*, where knowledge is acquired in the absence of instructions or intention to learn, but the person is conscious of what he or she has learned (Eysenck 1982). The critical feature of implicit learning is that it is unconscious, in the sense that the subjects are unaware of what they have learned. Documenting dissociations between explicit and implicit learning, then, is a somewhat tricky business. Many studies do not give a great deal of detail about the methods by which subjects' conscious knowledge was assessed, but it is probably not enough merely to ask subjects in the artificial-grammar experiments to describe the rule that governs the letter strings, and count them as "unconscious" when they fail to do so. In the first place, unless the test stimuli are very carefully constructed, even partial awareness of the rule – that the first letter must be either a *P* or a *T*, for example, may be enough to permit subjects to discriminate between grammatical and ungrammatical strings at better than chance levels. Although investigators of explicit and implicit memory have developed rigorous standards for matching explicit and implicit tasks, similar standards are generally lacking in studies of implicit learning.

The argument that implicit learning is really unconscious is sometimes bolstered by the fact that amnesic patients show preserved implicit learning. Of course, amnesics also forget the learning *episode* as well, confusing implicit learning with implicit memory. In this regard, it is somewhat disconcerting to note that subjects can show significant implicit "learning" even in the absence of any learning experience! That is to say, in some experimental procedures involving classification performance, it is possible for subjects to intuit the structure of the target category from test instances, even when they were denied an opportunity to learn the category during a prior study phase.

However, implicit learning is not always, or entirely, spared in amnesic patients. For example, amnesics show normal levels of perceptual learning in a visual search task, but impaired learning when contextual cues are added to the procedure. Even intact implicit learning by amnesic patients does not mean that explicit and implicit learning must be mediated by different brain systems. Dissociations between recognition and concept learning can be simulated in a computational model of exemplar memory that has only one system for storing memory, with different thresholds for recognition and classification.

It is sometimes claimed that implicit learning, precisely because it is automatic and unconscious, is a very powerful (as well as more primitive) form of learning – more powerful than conscious forms of learning that emerged more recently in evolutionary history (Reber 1993). While it does seem amazing that subjects can pick up knowledge of something as complex as an artificial grammar or a dynamic system automatically, and apply it unconsciously, claims for the superiority of unconscious processing sometimes seem to reflect a Romantic notion of the unconscious that goes back to von Hartmann (Hartmann 1868/1931, p. 40), who wrote that the unconscious "can really outdo all the performances of conscious reason." Unfortunately, enthusiasts of implicit learning have not always compared implicit learning to conscious, deliberate knowledge acquisition. How well would subjects perform if we actually showed them the finite-state grammar, or if we gave them feedback about their classification performance? What if we simply told subjects the sequence of quadrants in which the target would appear?

The Implicit and the Unconscious

Together with the concept of automaticity, research on implicit learning and memory constituted psychology's first steps toward a revival of interest in unconscious mental life (Kihlstrom 1987). Although the psychological unconscious suffered much in the twentieth century from taint by Freudian psychoanalysis the concepts and methods employed to study implicit learning and memory have now been extended to other domains, such as perception and even thinking – and beyond cognition to emotion and motivation (Kihlstrom 1999). In this way, the study of implicit learning and memory offers a new, non-Freudian perspective on unconscious mental life – and, in turn, on consciousness itself.

See also 38 Attention and consciousness; 40 Preconscious processing; 49 Consciousness and anesthesia.

Further Readings

Anderson, J. R. (2000) *Learning and Memory: An Integrated Approach*, 2nd edn. New York: Wiley.

Bower, G. H. and Hilgard, E. R. (1981) *Theories of Learning*, 5th edn. Englewood Cliffs, NJ: Prentice Hall.

McConkey, J. (ed.) (1996) *The Anatomy of Memory: An Anthology*. Oxford: Oxford University Press.

Tulving, E. and Craik, F. I. M. (eds.) (2000) *Oxford Handbook of Memory*. Oxford: Oxford University Press.

References

Adams, J. K. (1957) Laboratory studies of behavior without awareness. *Psychological Bulletin* 54, 383–405.

Anderson, J. R. (2000) *Learning and Memory: An Integrated Approach*, 2nd edn. New York: John Wiley and Sons.

Berry, D. C. and Dienes, Z. (1993) *Implicit Learning: Theoretical and Empirical Issues*. Hove: Erlbaum.

Bower, G. H. and Hilgard, E. R. (1981) *Theories of Learning*, 5th edn. Englewood Cliffs, NJ: Prentice Hall.

Bowers, J. S. and Marsolek, C. J. (eds.). (2003) *Rethinking Implicit Memory*. Oxford: Oxford University Press.

Butler, L. T. and Berry, D. C. (2001) Implicit memory: intention and awareness revisited. *Trends in Cognitive Sciences* 5, 192–7.

Cohen, N. J. and Squire, L. R. (1980) Preserved learning and retention of pattern analyzing skill in amnesia: dissociation of knowing how and knowing that. *Science* 210, 207–10.

Crovitz, H. F. (1970) *Galton's Walk: Methods for the Analysis of Thinking, Intelligence and Creativity*. New York: Harper and Row.

Eysenck, M. W. (1982) Incidental learning and orienting tasks. In C. R. Puff (ed.), *Handbook of Research Methods in Human Memory and Cognition*, 197–238. New York: Academic.

Flanagan, O. (1992) *Consciousness Reconsidered*. Cambridge, MA: MIT Press.

Frensch, P. A. and Runger, D. (2003) Implicit learning. *Current Directions in Psychological Science* 12: 1, 13–17.

Hartmann, E. V. (1868/1931) *Philosophy of the Unconscious: Speculative Results according to the Inductive Method of Physical Science*. London: Routledge and Kegan Paul.

Hering, E. (1870/1880) On memory. In S. Butler (ed.), *Unconscious Memory* (new edn., 1910), 63–86. London: Fifield.

Jacoby, L. L. (1991) A process dissociaton framework: separating automatic from intentional uses of memory. *Journal of Memory and Language* 30, 513–41.

Kihlstrom, J. F. (1987) The cognitive unconscious. *Science* 237: 4821, 1445–52.

Kihlstrom, J. F. (1995) Memory and consciousness: an appreciation of Claparede and *Recognition et Moïeté. Consciousness and Cognition: An International Journal* 4: 4, 379–86.

Kihlstrom, J. F. (1996) Memory research: the convergence of theory and practice. In D. Hermann, C. McEvoy, C. Hertzog, P. Hertel, and M. K. Johnson (eds.), *Basic and Applied Memory Research: Theory in* Context, vol. 1, 5–25. Mahwah, NJ: Erlbaum.

Kihlstrom, J. F. (1999) Conscious and unconscious cognition. In R. J. Sternberg (ed.), *The Nature of Cognition*, 173–204. Cambridge, MA: MIT Press.

Mandler, G. (1980) Recognizing: the judgment of previous occurrence. *Psychological Review* 87: 3, 252–71.

Razran, G. (1961) The observable unconscious and the inferable conscious in current Soviet psychophysiology: interoceptive conditioning, semantic conditioning, and the orienting reflex. *Psychological Review* 68, 81–147.

Reber, A. S. (1993) *Implicit Learning and Tacit Knowledge: An Essay on the Cognitive Unconscious*. Oxford: Oxford University Press.

Richardson-Klavehn, A., Gardiner, J. M., and Java, R. I. (1996) Memory: task dissociations, process dissociations, and dissociations of consciousness. In G. Underwood (ed.), *Implicit Cognition*, 85–158. Oxford: Oxford University Press.

Roediger, H. L. and McDermott, K. B. (1993) Implicit memory in normal human subjects. In F. Boller and J. Grafman (eds.), *Handbook of Neuropsychology*, 63–131. Amsterdam: Elsevier Science Publishers.

Salaman, E. (1970) *A Collection of Moments: A Study of Involuntary Memories*. London: Longman.

Schacter, D. L. (1987) Implicit memory: history and current status. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 13, 501–18.

Schacter, D. L. and Tulving, E. (eds.) (1994) *Memory Systems*. Cambridge, MA: MIT Press.

Schwartz, B. and Reisberg, D. (1991) *Learning and Memory*, 1st edn. New York: Norton.

Seger, C. A. (1994) Implicit learning. *Psychological Bulletin* 115, 163–96.

Squire, L. R. and Zola-Morgan, S. (1991) The medial temporal lobe memory system. *Science* 253, 1380–6.

Squire, L. R., Knowlton, B., and Musen, G. (1993) The structure and organization of memory. *Annual Review of Psychology* 44, 453–95.

Stadler, M. A. and Frensch, P. A. (eds.) (1998) *Handbook of Implicit Learning*. Thousand Oaks, CA: Sage.

Tulving, E. and Craik, F. I. M. (eds.) (2000) *The Oxford Handbook of Memory*. New York: Oxford University Press.

Tulving, E. and Schacter, D. L. (1990) Priming and human memory systems. *Science* 247: 4940, 301–6.

THE BLACKWELL COMPANION TO CONSCIOUSNESS

Max Velmans
Science Editor

Susan Schneider
Philosophy Editor

Advisory Editors

Science of Consciousness: Jeffrey Gray, John Kihlstrom, Phil Merikle, Stevan Harnad

Philosophy of Consciousness: Ned Block, David Chalmers, José Bermúdez, Brian McLaughlin, George Graham

Blackwell
Publishing

BLACKWELL PUBLISHING
350 Main Street, Malden, MA 02148-5020, USA
9600 Garsington Road, Oxford OX4 2DQ, UK
550 Swanston Street, Carlton, Victoria 3053, Australia

First published 2007 by Blackwell Publishing Ltd

1 2007

Library of Congress Cataloging-in-Publication Data

The Blackwell companion to consciousness / edited by Max Velmans, science editor and Susan Schneider, philosophy editor
p. cm.
Includes bibliographical references and index
ISBN-13: 978-1-4051-2019-7 (hardback)
ISBN-13: 978-1-4051-6000-1 (pbk.)
1. Consciousness. I. Velmans, Max, 1942- II. Schneider, Susan, 1968- III. Companion to consciousness.
BF311.B5348 2006
153--dc22 2006026291

A catalogue record for this title is available from the British Library.

Set in 10/12.5 pt Minion Pro
by The Running Head Limited, Cambridge, www.therunninghead.com
Printed and bound in Great Britain
by TJ International Ltd, Padstow, Cornwall

The publisher's policy is to use permanent paper from mills that operate a sustainable forestry policy, and which has been manufactured from pulp processed using acid-free and elementary chlorine-free practices. Furthermore, the publisher ensures that the text paper and cover board used have met acceptable environmental accreditation standards.

For further information on
Blackwell Publishing, visit our website:
www.blackwellpublishing.com