The Death of Implicit Memory

**Daniel B. Willingham** & Laura Preuss  
Department of Psychology  
102 Gilmer Hall  
University of Virginia  
Charlottesville, VA 22903-2477  
U.S.A.

[dbw8m@virginia.edu & lap2c@virginia.edu]

Copyright (c) Daniel B. Willingham & Laura Preuss 1995

PSYCHE, 2(15), October 1995  

KEYWORDS: classification, consciousness, explicit/implicit processes, memory, specificity.

**Abstract**

The thesis of this article is that implicit memory does not exist. Implicit memory phenomena are distinct from explicit memory phenomena at a neural and information processing level, but there is such variety among the implicit memory phenomena that nothing holds them together in a common category. Other researchers have distinguished among different types of implicit memory, but have retained the superordinate category. Extant data is evaluated in light of how classification systems should be developed, and it is concluded that there is currently not a reason to retain the construct "implicit memory."

**1. What Implicit Memory Might Be**

1.1 As defined by Graf & Schacter (1985) implicit memory refers not to a memory system, but to a set of memory tasks. Their distinction is based on characteristics of the experimental situation. Implicit memory tasks make no reference to the initial encoding episode and are not necessarily associated with awareness of engaging in recall. Explicit memory tasks make reference to the initial encoding episode, and are associated with subjective awareness of engaging in recall. The terms implicit/explicit have in some sense been hijacked by memory researchers, and now are often used to refer to hypothetical memory systems; this practice has become widespread enough that Schacter has felt compelled to comment on the topic (Schacter & Tulving, 1994). With apologies to Schacter, this article uses the terms implicit/explicit in this common (but strictly speaking, incorrect) sense. The arguments made here apply to many bicameral memory
1.2 All of these classification schemes rely on two sets of data for support: neuroanatomic and cognitive. The neuroanatomic data demonstrate dissociations between implicit and explicit memory based on brain lesion or brain imaging data. Thus, the two types of memory are separate because they rely on distinct anatomic structures for normal operation. For example, patients with global amnesia perform poorly on tasks that rely primarily on explicit memory such as recognition and recall tasks (Scoville & Milner, 1957) but these same subjects perform normally on tasks that rely primarily on implicit memory, such as motor skills (e.g., Corkin, 1968) and repetition priming (e.g., Warrington & Weiskrantz, 1970). The second data set lending support to the implicit/explicit distinction is purely cognitive. Implicit and explicit memory are thought to use different processes and/or representations in their normal operation. For example, the level of performance in explicit memory tasks is strongly influenced by the depth of processing at encoding, but is not influenced by the match of the perceptual characteristics of the stimuli (e.g. type font of written words) at encoding and retrieval. Measures of implicit memory, on the other hand, are often insensitive to level of encoding, at least in some paradigms, but are sensitive to physical characteristics of the stimuli such as type font (e.g., Graf, Mandler, & Haden, 1982; Jacoby & Hayman, 1987).

1.4 The answer, then, to "What is implicit memory?" can take two forms. First, the answer might be that "implicit memory is a memory system supported by particular brain structures." The second answer is that "implicit memory is a memory system using particular processes and representations." Naturally a third answer is that both of the previous answers are true because both distinctions apply. We will argue that neither distinction applies. Implicit memories have nothing in common; implicit memory does not exist as a coherent system.

1.5 The point is this. A number of different tasks purportedly offer relatively pure measures of implicit memory (although it is accepted that tasks are not process-pure, Jacoby, 1991): motor skill learning (Nissen & Bullemer, 1987), perceptual learning (Cohen & Squire, 1980), perceptual after-effects (Savoy & Gabrieli, 1991) and classical conditioning (Weiskrantz & Warrington, 1979). These tasks may be differentiable on a number of dimensions, but if they are all going to be called implicit, they must share some characteristic. What is the need for a superordinate category called "implicit memory" that ties these tasks together? To examine that question more closely, it is worth considering how classification systems in general are created.

2. Creating Classification Systems of Memory

2.1 Creating a classification system of anything (including memories) entails creating categories of objects based on their attributes. Objects have more attributes than will be
used to create categories. How then can one select which attributes are the important ones? There are, in fact, many, many possible classification systems for any group of objects, and these different classification systems might serve different functions. For example, someone concerned with safety will develop a different classification system for automobiles than someone concerned with economy and these two classification systems will use different attributes of automobiles. The function of classification systems of memory is different; the goal is to create a classification system that "carves nature at its joints"--a system that captures important divisions in the nature of memory itself.

2.2 That desire leads in turn to the question "What important attribute must be captured?" The answer to this question might be called the conceptual basis of a classification system (Ridley, 1986). For example, one might develop a system based on neuroscientific criteria. The attributes used to differentiate memories will, in that case, be their neuroanatomic or neuropharmacologic basis, for example. Or one might select information processing as a conceptual basis; memories might be identified as different (or the same) depending on the processes and representations they share. These conceptual bases may lead to different classification systems. For example, using a neuroscientific representation, one might differentiate between two types of memories with nonoverlapping neuroanatomic bases, but if there was good reason to think that the cognitive processes and representations supporting the memories were identical despite the diverse anatomic substrates, a classification system using an information processing conceptual basis would not differentiate the memories.

2.3 Once a conceptual basis has been selected, one must select attributes that will differentiate memories. For example, a memory classification system using a neuroscientific conceptual basis might use the attribute "effect of basal forebrain acetylcholine (ACh) on the memory." Naturally, the conceptual basis limits the possible attributes, but does not dictate them.

2.4 After the attributes have been selected, cut-points of the attributes that determine category membership must be developed. For example, given that "effect of basal forebrain ACh on the memory" is an important attribute, how is that attribute used? Should there be two levels, say "unaffected," and "affected?" Perhaps there will be three levels of this attribute, "impaired," "enhanced," and "unaffected."

2.5 The point of this section is to highlight that classification systems represent choices about what the researcher takes to be important, defining evidence for differences among memories. The classification systems that have been developed thus far use neuroscientific and information processing conceptual bases. The validity of the category "implicit memory" should be evaluated within the same framework used to develop classification systems.
3. Implicit Memory as a Neuroscientific Concept

3.1 Within neuroscientific conceptual bases, the one most often used is neuroanatomy. Although there has been some work on distinguishing implicit and explicit memories based on neuropharmacology (e.g., Nissen, Knopman, & Schacter, 1987), the bulk of the work with human subjects has used anatomy. The primary finding is that a set of interconnected structures in the medial temporal lobe, basal forebrain, and diencephalon support the formation of new explicit memories, but do not appear to contribute to the formation of new implicit memories (Squire & Zola-Morgan, 1991). If one or more of these sites is damaged, explicit memory is impaired, but implicit memory is not. This finding allows one to assign coherence to explicit memory because explicit memory is anatomically localized–remember, neuroanatomy is the conceptual basis. While there has been some debate about possible differences in the amnesic syndromes caused by damage to the medial temporal lobe vs. that caused by diencephalic damage, it now appears that there is not an appreciable difference between them (Squire, Knowlton, & Musen, 1993). This finding does not, however, allow one to assign coherence to what is spared in amnesia. What is spared in amnesia is simply "memory not in the medial temporal lobe, etc." For it to merit the term "implicit memory" there must be coherence in these type of memories within the conceptual basis of the system, in other words, anatomic consistency. Although there were some suggestions that implicit memory might reside in a single anatomic system (Mishkin, Malamut, & Bachevalier, 1984), that view has not been popular in recent years, for reasons described in the next section.

3.2 There does not appear to be anatomic consistency within implicit memory. Different implicit memory tasks are differentially affected by brain damage, i.e., they are dissociable. Tulving and Schacter (1990) and Keane, Gabrieli, Fennema, Growdon and Corkin (1991) have argued that repetition priming in fact represents two types of priming: a perceptual representation system in the occipital and posterior temporal cortices, and a semantic system in the anterior temporal cortices. Heindel, Salmon, Shults, Walicke, & Butters (1989) and Harrington, Haaland, Yeo and Marder (1990) have shown double dissociations between motor skill learning and other implicit tasks in striatal patients. Patients with Huntington's disease (HD) or Parkinson's disease (PD) do not learn the pursuit rotor motor skill normally, but HD patients show normal repetition priming (Heindel et al., 1989) and PD patients learn the mirror reading perceptual skill normally (Harrington et al., 1990). Willingham and Koroshetz (1993) have argued that different types of motor skill are differentially impaired in HD.

3.3 No one has proposed any manner in which to tie together the apparent neural separability of these different tasks. That is, one might argue that the implicit memory system is in fact a coherent system. Although it is subserved by a number of structures, those structures are interconnected, and to the extent that one sees dissociations with the system, that simply reflects the differential effects of damage to isolated components. The system, nevertheless, is coherent. Such an argument is theoretically possible, but has not been made, to our knowledge, nor does one seem likely. The multiplicity of putative sites
of implicit memory tasks (cerebellum, basal ganglia, a number of cortical areas) make it seem difficult to tie them all into one system, although it is dangerous, as always, to accept the null hypothesis.

4. Implicit Memory as an Information Processing Concept - Consciousness

4.1 What then, of the other conceptual basis? If it is hard to argue that implicit memory is a valuable superordinate category using a neuroanatomic conceptual basis, can one make that argument using an information processing basis? Is there some aspect of information processing that ties implicit memory in its various forms together? There are two characteristics of memories that have been offered as unifying implicit memory: consciousness and specificity. We address each in turn.

4.2 One way in which implicit memory appears to have some unity is in the role of consciousness. In all of the tasks typically identified as implicit, the subject need not be conscious of having learned, or indeed be conscious that learning is being expressed at the time of test. Note that this is not to say that the subject needs to be unaware of the stimulus materials at encoding or test; rather, the subject need not be aware that his or her performance is being affected by an earlier experience (although the performance measure indicates that it is). For example, in the case of the serial reaction time (SRT) task, the subject performs a four-choice response time task. The subject is not told that the stimuli appear in a 12-unit repeating sequence. After some training, subjects' RTs are faster if the sequence is present than if the stimuli appear randomly. Naturally the subject is aware of the stimuli as he or she performs the task; but the subject is unaware of the contingencies among the stimuli. If tested again a week later, the subject will again be aware of the stimuli, but will not know that his or her performance is affected by experience with the sequence in the previous session. As is true of many implicit memory tasks, subjects can learn the sequence explicitly in the SRT task, and explicit knowledge improves performance (Curran & Keele, 1993).

4.3 One might say, then, that consciousness is the attribute that all implicit memory tasks have in common, and the value of that attribute is "awareness is unnecessary to learning." That conclusion is controversial (Shanks & St. John, 1994) but suppose for a moment that it is true. There is still cause for concern in using consciousness as an attribute in an information-processing classification system of memory; what is consciousness, and how will knowing something about consciousness and memory aid in developing models of memory? Consciousness is not simply poorly understood; there is virtually no agreement over what it is.

4.4 That consciousness is so poorly understood might give one pause. But is that reason enough to preclude its use in a memory taxonomy? Farber & Churchland (1994) argue that, tempting as it is to focus on defining consciousness before making a serious attempt
to study it, this strategy is unwise. Rather, one should study consciousness even before one is sure of how to define it, and refine the definition as it becomes better understood.

4.5 We believe that argument is true for the study of consciousness, but the subject matter under study is memory, not consciousness, and we believe that the argument holds true for memory only to a point. The critical difference is that memory researchers seek to use the construct of consciousness as a tool. Researchers of consciousness have no choice but to make do with an incomplete understanding and a provisional definition of consciousness. Memory researchers, on the other hand, should ask a different question: does the construct of consciousness add to our understanding of memory?

4.6 How does a classification system aid our understanding of memory? By making clear that two types of memory are fundamentally different and will require different explanations. For example, suppose one evaluated many putatively different types of memory and noticed that all memories are more difficult to retrieve as time passes, except motor skills. This fact might be taken as evidence that motor skills are fundamentally different than other types of memory because this difference in the rate of forgetting implies necessary differences in how motor skills are retained. One might hypothesize that most memories decay, but motor skills do not, or that most memories are subject to interference, but motor skills are not.

4.7 We propose that a classification system based on consciousness alone does not make clear that two types of memory are fundamentally different. How do we know that consciousness is a critical attribute separating types of memories? Why is this attribute so important, more important than, say, the emotional content of the memory? The fact is, the main reason to be sanguine about the usefulness of consciousness is that it is consonant with neuroscientific data (Willingham, 1994). Memories that are always associated with consciousness are those that rely on the integrity of the medial temporal lobe and diencephalon. Memories that do not rely on these structures need not be associated with consciousness. Thus, consciousness provides some additional support for the separation of implicit and explicit memory. In the absence of the neuroscientific data, there would be no reason to think that consciousness was an important attribute; it could just as easily be an accidental attribute, no more important than the attribute "the memory is/is not about chocolate bars."

4.8 This is true exactly because we do not know the function of consciousness. To return to the earlier example, we might have some confidence that the rate of forgetting is an important characteristic that sets motor skills apart from other types of memory because it has implications for how these two hypothetical systems would operate and, what is more important, how their operation would differ. In other words, motor skills appear to be learned and remembered in a way that is different than other memories because they are not forgotten; perhaps motor skills are so different that they are handled by an altogether different memory system. Because we do not know what consciousness does, it does not imply anything about ways in which conscious and unconscious memory systems would
need to be different. This argument is similar to that offered by Sherry and Schacter (1987), who proposed that separate memory systems could be inferred when two memory functions were so computationally incompatible that it was unlikely or impossible that both functions could be performed by the same system. Consciousness, because we know nothing about it, implies nothing about why memories that have or do not have it might be different.

4.9 The role of consciousness in memory parallels the neuroscientific data in another way. Lack of awareness as a characteristic does not really unify implicit memory; rather, awareness sets explicit memory apart from the others. Similarly, not relying on the hippocampus does not unify implicit memories; rather, relying on the hippocampus sets explicit memory apart.

5. Implicit Memory as an Information Processing Concept - Specificity

5.1 A second characteristic has been suggested as one that ties together the various subtypes of implicit memory: specificity. Several researchers have argued that implicit memory is specific, meaning that it is more sensitive to changes between encoding and retrieval conditions, whereas explicit memory is more flexible, meaning explicit memories may be accessed and used in contexts quite different from those under which it was encoded. (Berry & Dienes, 1993; Cohen & Eichenbaum, 1993; Squire, 1992) The interpretation of this fact has varied, however. Squire (1992; 1994) has merely remarked that specificity seems to be a common characteristic of implicit memory. Berry and Dienes and Cohen and Eichenbaum have been more forceful in offering specificity as a common characteristic of disparate implicit memory tasks, and Cohen and Eichenbaum have specifically addressed the issue of memory taxonomy from this perspective. We will first lay out the supporting evidence and then point out some problems with this position.

5.2 There is good evidence supporting this position in the animal literature. In both rats and monkeys, an animal with medial temporal lesions has difficulty using knowledge in a new situation. For example, if a rat is trained with a series of paired associates in an olfactory discrimination task, and then is transferred to novel pairings of the items, the animal treats it as a new problem, even though the proper response is predictable from training. For example, the rat might be given the pairs A and C during training, and then transferred to the pair A. A sham-operated rat will perform well on this task; a rat with fornix lesions (a major output pathway of the hippocampus) will not (Eichenbaum, Fagan, Mathews, & Cohen, 1989).

5.3 Conceptually similar results have been obtained with rhesus monkeys. Monkeys with medial temporal lesions (different combinations of lesioned structures were examined) and control monkeys were trained on a complex object-object association task. Object A was rewarded only if B was present, but not if C was present. Object D was rewarded
only if C was present, but not if B was present. Operated and control monkeys acquired this task equally well, but operated monkeys had trouble expressing this knowledge in a new testing situation. An object was placed before the monkey (e.g., object A), and two objects placed on either side of it one of which was rewarded in the presence of A (object B) and one of which was not (object C). The monkey was to pick one of the objects (B or C) that was rewarded in the presence of the centrally presented object (A). Control monkeys performed well; operated monkeys did not (Saunders & Weiskrantz, 1989).

5.4 Glisky and Schacter (Glisky & Schacter, 1987; Glisky, Schacter, & Tulving, 1986) taught amnesic patients some computer-programming terms. This task would normally be considered an explicit memory task, and the patients were indeed quite impaired in learning the terms. They required much more practice than control subjects. The interesting finding, however, was that their knowledge, once acquired, appeared to be qualitatively different than that of control subjects because it was, in Glisky's term, "hyperspecific." Subjects could only reproduce what they had learned if the recall conditions matched the encoding conditions quite precisely. Still, other researchers have reported that amnesic patients may successfully transfer explicit knowledge to other learning situations (e.g., Shumamura & Squire, 1988) 5.5 The evidence from non-amnesic human subjects seems weaker. Some studies show that implicit measures are sensitive to structural changes in the stimuli, but explicit measures are not. For example Graf, Shimamura, and Squire, (1985) showed that repetition priming is reduced when the modality of word presentation (visual or auditory) is changed between encoding and retrieval, but recall is not effected. Many other studies support this basic finding: when the physical characteristics of stimuli are changed, priming is reduced out of proportion to any reduction in explicit memory under similar circumstances (Graf et al., 1985; Jacoby & Hayman, 1987; Weldon & Roediger, 1987).

5.6 There are instances, however, in which the precise stimulus need not be repeated for priming to occur at normal levels (Biederman & Cooper, 1991; Cooper, Schacter, Ballesteros, & Moore, 1992). Even if there were not exceptions to the effect, there would be a problem. These studies change the structural characteristics of stimuli and priming suffers--but priming is based on structural characteristics of stimuli! Naturally changing them affects priming (see Kolers & Roediger, 1984, for a general discussion of related points). This point was made forcefully by Graf and Ryan (1990). They found that attention to structural characteristics at encoding mediates whether changing the format at retrieval influences repetition priming. If subjects were asked to rate the readability of a word at encoding, there is an effect of changing the font at test; if subjects rate the pleasantness at encoding, changing the font at test does not affect priming.

5.7 Explicit memory, on the other hand, very likely maintains semantic but not structural information (e.g., Bransford & Franks, 1971) so it is natural that changing structural characteristics of the stimulus will have little effect on explicit memory. A comparable change for explicit memory is to bias subjects towards different meanings of the stimuli at encoding and retrieval. Such experiments have been conducted, and under those
circumstances it is explicit memory that appears "hyperspecific." If subjects are cued at encoding to think of a particular attribute of the to-be-remembered (TBR) word and then at retrieval are cued with a different attribute of the TBR word, explicit memory performance is reduced. For example, subjects see "the man tuned the PIANO" or "the man lifted the PIANO" (the word in capitals being the TBR word). At recall, the retrieval cue "music-producing" is more effective for the first orienting sentence than the second, but the opposite is true for the cue "heavy." (Barclay, Bransford, Franks, McCarrell & Nitsch, 1974). This is not a small or marginal effect; subjects recalled an average of 4.4 target words when the retrieval cue matched the encoding sense, and 0 target words when it did not.

5.8 If it's true that flexibility is not absolute, but is a function of whether the important aspect of the representation is changed between study and test, then implicit memory should not be affected when semantic sense is changed between encoding and retrieval. That seems to be true, at least for repetition priming. One way to change the semantic interpretation of a word is to present the word with a second word at encoding, ask subjects to relate the two words, and then present the target word at retrieval with a different, third word. That paradigm was used by Graf & Schacter (1985) and they reported reduced priming in amnesic patients when the context was changed. Subsequent work, however, showed that the effect was difficult to replicate in amnesic patients (Cermak, Bleich, & Blackford, 1988; Musen & Squire, 1993; Shimamura & Squire, 1989) and that, in normal subjects, the context effect was mediated by explicit, not implicit, memory (Bowers & Schacter, 1990). Thus implicit memory is flexible in that it is not affected by changes in semantic context.

5.9 What about implicit measures other than priming? Do they show specificity? Cohen & Eichenbaum (1993) argue they do. In our view, this evidence is mixed. It is true that when learning to read transformed text subjects perform best on the materials used during training (Kolers, 1976) although there is not hyperspecificity to the point that new text materials are read as though the subject had no experience reading transformed text. More dramatic is a study showing that if the training is limited to certain letters of the alphabet, there is no transfer to new letters (Masson, 1986). It is noteworthy, however, that precisely the same manipulation--changing the letters used during training--has no impact at all on another implicit task, learning finite state grammars. In finite state grammar learning, subjects are shown letter strings (e.g., RTVVVE) and are asked to memorize them. These strings are all generated by a finite state grammar. Later, subjects are informed of this fact and are shown other strings generated by the grammar, as well as some strings that violate the rules of the grammar. Subjects must judge the grammaticality of the strings. Normal subjects can do so in the absence of awareness of the formal rules governing the grammar, and amnesic patients can learn this task normally (see Reber, 1989 for a review). Thus, this learning appears to be implicit. Most important for this discussion, the letters used on the test stimuli can be changed, and subjects show transfer (Matthews et al., 1989). Thus, the same manipulation (changing letters at training and test) has a profound negative impact on one implicit task (reading transformed text) and less impact on another (grammaticality judgements). It appears to us that one cannot make a flat statement about the specificity of implicit memory.
5.10 Indeed, other evidence indicates that some forms of implicit show quite a bit of flexibility. A hallmark of motor skills is that they can be applied to conditions that were not part of the training regimen. It is certainly true that some motor skills transfer to different effectors; for example, a person's handwriting is recognizably his or her own, even if the effector used is the non-dominant hand, or if the pen is held in the mouth (Glencross, 1980). In part because of this work, many researchers subscribe to the idea that motor skill is represented as a generalized motor program—a representation that can be adjusted to particular stimulus situations by changing parameters that in turn influence the behavior produced. This theory was specifically formulated to account for the adaptability of motor skill to new situations (see Schmidt, 1988, for a review). The flexibility of implicit memory is not limited to purely motor skills. Bedford (1989) trained subjects to point to specific targets when wearing wedge-prism spectacles that offset the visual world thirty degrees, and found perfect transfer to novel targets.

5.11 In sum, there is evidence that human and animals with lesions to the medial temporal lobe have difficulty expressing what they can learn in novel situations. The data on this point are rather limited in these populations, however. In tests of normal human subjects, there is data for all possible outcomes: specificity and flexibility for both implicit and explicit memory, making this dimension a poor candidate for use in a classification system of memory. In our judgement, the reason that data from human amnesics appear to support the hypothesis that implicit memory is hyperspecific is because there has been limited testing of the hypothesis in those populations. If amnesic patients were taught a novel motor skill, it seems likely it would be as flexible for patients as for control subjects.

6. The Death of Implicit Memory

6.1 There are two points of this article. First, the construct "implicit memory" is only useful if some attribute can be found that ties together motor skill learning, priming, conditioning, habituation, and so on. Second, there is not an attribute that usefully does so. It should be borne in mind that in this article "implicit memory" refers to a memory system classification, not a task-based classification.

6.2 Two conceptual bases have been used to distinguish explicit and implicit memory--neurobiology (specifically neuroanatomy) and information processing. These two conceptual bases were examined for evidence that the construct "implicit memory" should be retained. There is already ample evidence that the various components of implicit memory are subserved by distinct neural structures. It may be that there is an aspect of information processing that all of the components of implicit memory have in common. Specificity has been proposed. It appears that explicit memory is equally specific when the semantic interpretation of stimuli is changed between study and test. It also appears that implicit memory can be flexible when characteristics of stimuli that are not important to implicit memory are changed.
6.3 The best candidate for such a common information-processing attribute is consciousness. Although we believe the data are consistent with the idea that implicit memory tasks are those that can be learned in the absence of awareness, this attribute does not aid the theorist, although consciousness is of some use because it is in agreement with the neuroscientific data.

References


