HANDBOOK OF NEUROPSYCHOLOGY

Series Editors

FRANÇOIS BOLLER
Unit 324 INSERM, Centre Paul Broca, Paris, France

JORDAN GRAFMAN
Cognitive Neuroscience Section, Medical Neurology Branch, NINDS, Bethesda, MD, U.S.A.

Editorial Board:

G. BERLUCCI (Verona), G.W. BRUYN (Leiden), S. CORKIN (Cambridge, MA),
A. DAMASIO (Iowa City), E. DE RENZI (Modena), G. GAINOTTI (Rome),
H. GARDNER (Boston), H. GOODGLASS (Boston), R. KATZMAN (La Jolla),
R.D. NEBES (Pittsburgh), I. RAPIN (New York), G. RIZZOLATTI (Parma),
S. SEGALOWITZ (St. Catharines, Ont.), H. SPINNLER (Milan),
L.R. SQUIRE (La Jolla)

ELSEVIER
Amsterdam – London – New York – Tokyo
CHAPTER 2

Implicit memory in normal human subjects

Henry L. Roediger III and Kathleen B. McDermott

Department of Psychology, Rice University, Houston, TX, USA

Introduction

Implicit memory is a relatively new term in the cognitive psychologist’s lexicon, having been introduced by Graf and Schacter in 1985. Even so, the rubric has served as the impetus for a surprising volume of debate, research, and controversy. In the process, the term has been used in a variety of ways. In his influential review article, Schacter (1987) stated that ‘Implicit memory is revealed when previous experiences facilitate performance on a task that does not require conscious or intentional recollection of those experiences’ (p. 501). This form of memory is contrasted with explicit memory, which ‘is revealed when performance on a task requires conscious recollection of previous experiences’ (p. 501). Despite these straightforward verbal descriptions of the terms, debate has swirled about their appropriate use (e.g., Richardson-Klavehn and Bjork, 1988), because they have been used to refer to both a form of memory and to a type of psychological experience. Graf and Schacter (1985) said that ‘For descriptive purposes, we use the terms implicit memory and explicit memory to distinguish between these forms of memory’ (p. 501). Schacter (1987) also advocated descriptive use of the terms, but stated that they were meant to be ‘descriptive concepts that are primarily concerned with a person’s psychological experience at the time of retrieval’ (p. 501). These differing descriptions may lead to confusion, as we discuss below.

In many ways the primary paradigm for studying one form of implicit memory is, on the surface, quite similar to the customary way of studying explicit memory in laboratory experiments. In both cases, subjects are exposed to material (usually a list of discrete words or pictures) in a first phase. In a later, test phase, subjects in an explicit memory task are asked to recall words or pictures or to recognize them amongst plausible lures; in either case, they are asked to consciously retrieve the prior experiences.

In an implicit memory task, subjects are not directed to intentionally recollect the items presented earlier; rather, they are given some ostensibly unrelated task, such as identifying fragmented forms of the pictures or words or identifying them from brief displays. The finding is that subjects can name the fragmented forms of the stimuli better if they have recently studied them than if not. This facilitation is termed priming1 and is the index of implicit memory in the sense that Schacter (1987) used the term. Priming on the implicit memory tests can be dissociated from recall or recognition on explicit tests, as this review will document.

The original reason that cognitive psychologists became excited about the study of priming on implicit memory tests was the observation that performance on these tests seems resistant to the dramatic forgetting sometimes seen on explicit tests. Densely

---

1 Priming is used in many senses by cognitive psychologists. The semantic priming paradigm using the lexical decision or naming task — in which a prime and target occur in close temporal contiguity — is perhaps the dominant use of the term priming (e.g., Neely, 1977). Priming on implicit memory tests occurs over longer intervals.
amnesic patients, who perform very poorly on standard recall or recognition tests, were found to show normal patterns of priming on tasks that today would be called implicit memory tests (Graf, Squire and Mandler, 1984; Warrington and Weiskrantz, 1968, 1970). This pattern of performance has been regularly obtained, as reviewed in the next Chapter by Moscovitch, Vriezen, and Gottstein (this Volume), and has led to the conclusion that implicit tests of retention measure a form of knowledge different from that tapped by standard, explicit tests. The logic for this claim is that although amnesic patients are not able to recollect their experiences, they nonetheless show intact retention on implicit tests; therefore, the tests must assess different processes, information, or systems. This conclusion is bolstered by other reports in which researchers have shown intact priming in groups of subjects or under experimental conditions in which conscious recollection is impaired. For example, although subjects under the influence of alcohol have been found to perform poorly on explicit tests of retention, they show levels of priming equivalent to those of sober subjects on implicit tests (Hashtroudi, Parker, Delisi et al., 1984). Similarly, both the aged (Howard, 1991; Light and Singh, 1987) and children (Carroll, Byrne and Kirsner, 1985; Greenbaum and Graf, 1989; Naito, 1990; Parkin and Streete, 1988) show this same pattern of impaired explicit but relatively intact implicit test performance. Subjects exposed to information when under anesthesia (and undergoing surgery) have also shown priming on implicit memory tests when they have shown absolutely no ability to recall or to recognize the material (Jelicic, Bonke, Wolters et al., 1992). Finally, depressed subjects often show intact priming but impaired performance on explicit tests (see Roediger and McDermott, 1992). For these reasons, and others, certain implicit memory tests seem to reveal a form of retention that is quite different from that measured on explicit memory tests such as recall or recognition.

The purpose of this review is to examine what is known about priming phenomena in normal, young adult subjects under a variety of experimental conditions. We cannot take space to review the interesting phenomena alluded to in the previous paragraph, but the succeeding review by Moscovitch et al. covers work with brain-damaged subjects. In addition, we focus primarily on priming tasks involving words and pictures and only occasionally consider research on other phenomena that meet the criteria of implicit memory tests. Indeed, considered at its broadest level, most of the dependent measures in psychology would be encompassed by the burgeoning area of implicit memory. That is, most experimental areas of psychology study how changing environmental conditions later affect behavior without subjects being told explicitly to remember the earlier experiences that are affecting their behavior. Studies of classical and operant conditioning, of observational learning, of motor skill learning, and of perceptual adaptation effects (among others) would classify as implicit memory tests under the broadest description. Thus, any review of implicit memory phenomena must delineate its purview, and we have chosen priming phenomena as a relatively self-contained area of inquiry.

The present review will be organized as follows. In the first Section, we consider some issues of terminology and then in the next part provide an overview of the types of implicit tests that will be covered. In the third Section, we turn to the vexing issue of how to determine if an implicit test is reflecting unintentional retrieval and consider several possible criteria. The main part of the review is devoted to an examination of the primary classes of experimental variables that have been examined for their influence on priming tests in normal subjects. The major classes of variables are types of materials, surface features of materials, presentation time and repetition, instructional manipulations at study, retention interval, and finally a few miscellaneous variables. Following this review of the empirical evidence, we then consider some of the theories that have been proposed to explain dissociations between explicit and implicit memory tests (and between implicit tests themselves). Finally, we provide a few concluding remarks.
Inchoate terminology

Research on implicit memory tests is developing rapidly and, as occurs in all situations of rapid change, the terms used to describe the new ideas, paradigms, findings, and theories are in flux (Roediger, 1990b). Gardiner and Java (1993) have complained about the confusing use of terms that has occurred in this area, particularly use of the same term to refer to memory tasks, to underlying hypothetical constructs, and to states of awareness that might accompany performance on the tests. Although Graf and Schacter (1985) and Schacter (1987) intended for the terms implicit and explicit to be descriptive, the above quotes from Schacter’s (1987) paper show that he used the terms to describe both a form of memory and a type of conscious experience. The same is true of other terms; episodic memory has been used to name a memory system and a type of task. Gardiner and Java (1993) suggest that terms not be used in this confusing way but, although we agree in principle, it seems unlikely that researchers will agree upon any set of terms in the near future.

In this review, we generally use the terms explicit and implicit to refer to different tasks, distinguished operationally by the instructions subjects are given during the test. On explicit tests, subjects are asked to recall or to recognize events from certain times and places in their past experience; therefore, the most common usage of the term overlaps with Tulving’s (1972) concept of episodic memory. Others have preferred to call these sorts of tests direct memory tests (Johnson and Hasher, 1987; Segal, 1966), intentional memory tests (Jacoby, 1984) or those displaying controlled uses of memory (Jacoby, 1991).

On priming tests of implicit memory, subjects are not directed to recollect past events. Rather, they are asked to complete some task as well as possible; retention is inferred from how the task has been affected by recent experiences. This effect is determined by comparing task performance on primed items to performance on unprimed items, the control condition. Therefore, the test does not require conscious recollection of recent experiences for its completion, although the task may be affected by conscious recollection of such events. Some researchers prefer the term indirect tests (Johnson and Hasher, 1987; Segal, 1966), as memory is assessed indirectly, or incidental tests (Jacoby, 1984), because the reflection of memory is incidental (from the subject’s viewpoint) to the purpose of the task. More recently, Jacoby (1991) has suggested that these tests (in the pure case) may assess automatic uses of memory.

As is clear from the previous paragraphs, numerous terms exist to describe the phenomena under review here. At the present stage, little harm — and perhaps even great good — may come from having terms compete in the marketplace of ideas, to see which ones researchers will find most useful in the long run. We have chosen to orient our review around the implicit/explicit contrast, and we could mount a defense for this choice (see Roediger, 1990b; Schacter, 1987, 1990, for some reasons), but others clearly disagree (e.g., Richardson-Klavehn and Bjork, 1988). At this point we will not pause to consider relative merits of the various terms. We suspect that the meanings of the main contrasts are highly correlated, if not entirely overlapping, so that researchers can communicate with one another reasonably well (if not perfectly). We intend the distinction between explicit and implicit memory tests to refer only to types of tasks, not to different forms of memory or types of conscious experience, although the usual implication is that these tasks are reflecting different underlying properties of retrieval. The terms are meant to be neutral theoretically, except in the sense that all terms carry tacit theories with them.

The contrast between explicit and implicit memory tests does at least endorse the notion that the terms refer to a fundamental contrast. The evidence for this assertion is the striking dissociations between performance on implicit and explicit tests of memory that are often found for pathological states (as mentioned above) and in normals (as reviewed below). However, a problem for the study of these issues, especially in normal subjects, is that subjects
may not follow directions given for the task. Performance on implicit memory tests may be affected by intentional recollection; similarly, explicit memory tests may be affected by priming or unintentional retrieval. These considerations provide thorny problems for researchers studying priming and its relation to conscious recollection. We consider these problems and what can be done about them below, but first we survey the main forms of implicit memory tests that assess priming and extend our discussion of the terminology used to describe them.

Tests used to assess priming

Tests that can be used to assess implicit memory are potentially numerous, limited only by experimenters’ ingenuity. To date, about a dozen or so have been used relatively widely so that they contribute enough literature to be worth reviewing. Table 1 provides a list of 12 tasks, although some represent relatively minor variations on a theme, and others could be added (e.g., clarification procedures where stimuli gradually appear). As noted, priming experiments involve at least two phases. In the first phase, subjects are exposed to material (usually words or pictures); often, subjects are asked to perform some orienting task as they examine the material. Following this first phase, subjects may be given various filler tasks before being given the critical task, the implicit memory test, which is presented as just another task the subject is to complete; no instruction is given for subjects to recollect or to remember information from the prior phase to aid in performing the current task. (Some researchers inform subjects that some items produced may be ones encountered in an earlier phase in the experiment). Again, interest centers on priming, or the difference in the speed or accuracy with which subjects perform the task after relevant prior study compared to performance when no relevant experiences have occurred.

If subjects study a token of the same item they are to produce on an implicit test, then the situation is typically referred to as one measuring direct priming (e.g., study the word elephant in a list and later complete the fragment e- - p- - n- ). If the studied item is more indirectly related to the target, then the situation is one of indirect priming (e.g., study task or pachyderm and complete e- - p- - n- ). This distinction (owing to Cofer, 1967) is helpful and worth preserving, but complications can arise. If subjects study a picture of an elephant and then are asked to complete a fragment or stem referring to the name of the beast, is this direct or indirect priming? (The fact that direct and indirect priming already have well-established meanings in this literature is another reason to favor the terms explicit and implicit in describing the two forms of test).

In most cases, the form of the item given at study (say, an intact word) is different from the form of the test cue (say, a fragmented form of the word). We refer to this as direct priming, but not repetition priming. Similarly, the response subjects make often changes between study and test; during study they may make pleasantness judgments to intact words and during the test try to name fragmented words. In our lexicon, the term repetition priming will be reserved for the case in which both the stimulus presented and the response required during the study phase are equivalent to the stimulus and response at test, as sometimes occurs when subjects make decisions about whether strings of letters are words or nonwords during both the study and test phases. Direct priming will refer to the more typical case in which the same type of item is presented at study and test, although the exact form changes (e.g., study a word and then later complete its fragmented form). We include cross-modal priming (hear elephant, complete the visual fragment later) as an instance of direct priming because the same lexical item occurs on both occasions. Indirect priming refers to cases in which the study item and test cue are related in some other way: by appearance (element), by sound (sycophant), or by meaning (pachyderm). Finally, we use the term cross-form priming when the studied item and target are in different symbol systems (e.g., a picture of an elephant might be studied, or its name in another language for bilingual subjects, before its name in a
TABLE 1
Twelve tests used to assess implicit memory

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Word (Perceptual) Identification</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Word Stem Completion</td>
<td>c_e_c_h_e_t</td>
<td></td>
</tr>
<tr>
<td>Word Fragment Completion</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Degraded Word Naming</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anagram Solution</td>
<td>lepanthe</td>
<td></td>
</tr>
<tr>
<td>Lexical Decision</td>
<td>word/nonword decision</td>
<td></td>
</tr>
<tr>
<td>Nonverbal</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Picture Fragment Naming</td>
<td>Degraded picture</td>
<td>Srinivas (1993)</td>
</tr>
<tr>
<td></td>
<td>identification</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Object/nonobject decision</td>
<td>Kroll and Potter (1984)</td>
</tr>
<tr>
<td></td>
<td>Possible/impossible</td>
<td>Schacter, Cooper and Delaney (1990)</td>
</tr>
<tr>
<td></td>
<td>object decision</td>
<td></td>
</tr>
<tr>
<td>Conceptual tests</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Category Instance Generation</td>
<td>animals- ?</td>
<td>Srinivas and Roediger (1990)</td>
</tr>
<tr>
<td></td>
<td>Hannibal use to help</td>
<td></td>
</tr>
<tr>
<td></td>
<td>him cross the Alps in</td>
<td></td>
</tr>
<tr>
<td></td>
<td>his attack on Rome?</td>
<td></td>
</tr>
</tbody>
</table>

This set of terms is not incontestable and certainly does not apply to all situations that can be devised, but seems a good starting place for classifying relations between study and test. We find it useful to distinguish among various types of implicit memory tests, as shown in Table 1 (Roediger, Weldon and Challis, 1989b). A major contrast is between perceptual and conceptual tests; additionally, perceptual tests may be verbal or nonverbal. The perceptual implicit tests challenge one of the perceptual systems (usually vision) by presenting stimuli either rapidly or in fragmented form. The test cue is a perceptually degraded form of the target that the subject is trying to decipher. (We include the lexical decision task because subjects make speeded responses, even if the test stimulus is not degraded.) As we shall see, these tasks are often greatly affected by the perceptual similarity between study and test stimulus forms. In the case of conceptual tests, there is no perceptual resemblance between the priming event (study of the word *elephant*, say) and the test cue (answering a general knowledge question such as ‘What animal did Hannibal use to help him cross the Alps in his attack on Rome?’). The perceptual similarity between study and test events (e.g., both being in the same modality) does not affect priming on conceptual tests, but manipulations of meaning do. The contrast between tests that are primarily perceptual and those that are primarily conceptual is not theoretically neutral, but is based on theories that suppose that these reflect different cognitive operations (Roediger, 1990a; Roediger and Blax-
ton, 1987b) and different neural systems (Schacter, 1990, 1992; Tulving and Schacter, 1990). Some (e.g., Hunt and Toth, 1990; Tenpenny and Shoben, 1991) do not find the distinction helpful because of empirical difficulties. We consider some of these difficulties later, but for us the distinction between perceptual and conceptual forms of priming helps to make sense of and to organize much of the literature on priming in normal subjects (hence its use here) and in brain-damaged patients (see Schacter, 1990, 1992).

Our review is guided by an approach that emphasizes the mental procedures that underlie task performance rather than structural properties of hypothetrical memory traces or systems (Kolers and Roediger, 1984). (However, most claims we make can be readily translated into alternate theoretical languages; Graf and Ryan, 1990; Schacter, 1990.) In particular, we orient the review around a general transfer appropriate processing approach to explaining differences among memory tests (Transferd, Franks, Morris, Stein, 1979), as augmented by several assumptions. The guiding assumptions of the approach, stated briefly, are as follows:

1. Performance on memory tests benefits to the extent that the cognitive operations involved in the test recapitulate or overlap those engaged during initial learning.
2. Implicit memory tests usually require different mental processes (or access different forms of information) than do explicit tests and consequently benefit from different types of processing during learning.
3. The most commonly used implicit memory tests draw primarily on perceptual processes used in word recognition and object recognition. We refer to such tests as perceptual implicit tests.
4. Most typical explicit memory tests, involving recall and recognition, draw largely on conceptual processing (elaboration, organization, meaningful processing and the like).
5. The perceptual/conceptual classification of tests is not meant as a dichotomy; rather, it is more profitable to think of these as two separate dimensions that do not necessarily trade off against one another.
6. The perceptual/conceptual contrast is not coextensive with the implicit/explicit distinction; one can create conceptual implicit tests and perceptual explicit tests.

These assumptions, stated too boldly here and without qualification, are fleshed out in several places, but especially in Blaxton (1989), Roediger (1990a), Roediger et al. (1989b) and Weldon (1991). We use this framework to some extent to organize our review and, to anticipate, many findings fit into this approach quite comfortably. On the other hand, we encounter several problems in accounting for specific findings. In the section on Theoretical Approaches near the end of the Chapter, we consider the relative merits of the transfer appropriate processing approach compared to others.

We turn now to consider tests in Table 1. By far, the most popular implicit memory tests are verbal tests that seem to invoke lexical processes or to tap perceptual word form representations (Kirshner, Dunn and Standen, 1989b; Schacter, 1992; Weldon, 1991). At the top of Table 1 appear the most frequently used forms of these tests, although the bulk of the studies have used word (or perceptual) identification, word stem completion, or word fragment completion. On the right side of Table 1 are references that include good procedures and (usually) materials for the various tests that might be useful. The entries in the perceptual verbal section of the table could be expanded, because an auditory form of most implicit tests could be used. For example, subjects could be required to try to identify words in white noise for word identification (Schacter and Church, 1992), to listen to the first parts of words and then complete them orally for word stem completion (Bassili, Smith and MacLeod, 1989), or to perform an auditory lexical decision task (e.g., Slowiaczek, Nusbaum and Pisoni, 1987).

Surprisingly, pictorial and object priming have been used less extensively than verbal priming tests, although this practice may be changing. Tasks such as naming fragmented pictures have been long used...
to study perception (e.g., Leeper, 1935; Gollin, 1960) and were used to test memory-impaired patients early in the century (Schneider, 1912) and in the modern era (Warrington and Weiskrantz, 1968). Picture fragment naming tasks are now frequently used as implicit memory tests in normal subjects, too (e.g., Snodgrass and Feenan, 1990; Srinivas, 1993; Weldon and Roediger, 1987). Kroll and Potter (1984) developed an object decision task in which subjects decide whether presented objects exist in the real world, and Schacter, Cooper and their colleagues have used a similar task in which subjects decide whether unfamiliar line drawings represent possible or impossible objects (i.e., the impossible objects are Escher-like figures that provide conflicting depth cues; see Schacter, Cooper and Delaney, 1990). These tasks all seem to tap perceptual processes or a structural description system that represents objects (Schacter, 1992).

Conceptual implicit memory tests have been even less thoroughly studied than the nonverbal perceptual tests. In conceptual tests, subjects are given a cue word or query that is meaningfully (conceptually) related to the target item that was studied. In the case of word association, they are asked to free associate to the cue (e.g. tusk --? after studying elephant). In the case of category instance production, they are asked to produce members of the relevant category (e.g., animals). Answering general knowledge questions, as the name implies, assesses priming on items like those used in the popular Trivial Pursuit game (see Nelson and Narens, 1980, for a listing of materials).

The list of implicit memory tests provided in Table 1 is certainly not exhaustive. By the broadest definition of implicit memory tests, every sort of judgment or test that is (a) affected by past experience, and (b) given under conditions in which subjects are not explicitly instructed to remember earlier events, would qualify. Other tasks that have been used include reading inverted text (Cohen and Squire, 1980; Kolers, 1976), spelling homographs (Jacoby and Witherspoon, 1982), making judgments of preference or liking (e.g., Johnson, Kim and Risse, 1985), naming pictures or words (e.g., Mitchell and Brown, 1988), learning artificial grammars (Reber, 1967), learning sequences of key presses (e.g., Lewicki, Czyzewska and Hoffman, 1987; Stadler, 1989), making decisions about whether movement of dots is natural or not (Nilsson, Olofsson and Nyberg, 1992), making judgments of fame (Jacoby and Whitehouse, 1989) or judgments of brightness and darkness or level of noise (Jacoby, Allan, Collins et al., 1988; Mandler, Nakamura and Van Zandt, 1987), or rating faces for familiarity (Ellis, Young, Flude et al., 1987). We could also include perceptual adaptation effects (Benzing and Squire, 1989; Savoy and Gabrieli, 1991), implicit memory for new associations (Graf and Schacter, 1985), and event related potentials (Bentin and Moscovitch, 1990). Even this list omits motor learning and conditioning tests. Given the broad criteria, the list of tasks used as implicit memory tests will surely grow longer. In the remainder of the review, we concentrate on the 12 tasks in Table 1 because a reasonably self-contained literature has developed about them. Mention of other tasks will occur as needed.

Determining the phenomenological status of implicit memory tasks

Implicit memory tests are thought to tap an automatic form of memory, or incidental retrieval. Yet, especially in the case of normal subjects, the possibility exists that subjects will augment their responding to test cues in the implicit test by resorting to controlled or intentional uses of memory. Suppose subjects participated in an experiment in which, during the first phase, they were asked to answer Trivial Pursuit types of questions ("What is the average age of U.S. Presidents when they are elected to office?") and then given the answer (54). After a short delay they are given more questions of the same sort during the transfer phase. No mention is made of the original learning phase during the transfer phase, and subjects are now just told to answer the questions as well as possible with no feedback given. If subjects were again given the question about U.S. Presidents' average age when elected, we
suspect they would answer much better than originally (or than a group of subjects who did not get the question originally). Although this hypothetical experiment might be conducted under conditions ostensibly used for repetition priming — same task, cue and response given at study and test with no instruction to use controlled memory processes to answer the question — we suspect most readers would share our intuition that intentional retrieval would likely be involved in answering the question the second time (unless the retention interval were very long). The nature of the cue given during the test would be such that it should serve as a powerful reminder that subjects had just been given the question and answer previously.

Of course, no one has done experiments with normals exactly like the hypothetical one outlined here², but the question remains as to how one can know if implicit or indirect tests are relatively pure assessments of the incidental form of memory that they are designed to assess. Even if the cue is not so powerful — say the fragment _ _ bu _ ta _ for a studied word — how could a researcher be sure that subjects did not use the cue to try to recollect presentation of the word (rebuttal) from the list?

This thorny issue of awareness will exercise the field for some time to come; Schacter, Bowers and Booker (1989) present an excellent orientation to the problem, noting that conscious processes can be invoked at several stages during an implicit memory test, even if subjects do not intend to use test items as recall cues. For example, in one likely scenario, while performing the task of completing fragmented words, a subject may notice after producing a word that it was studied in the list. Such an occurrence — which might be dubbed unintentional conscious recollection — would not be troubling and would not compromise the implicit test, unless this awareness caused the subject to change his or her strategy and adopt intentional retrieval processes on upcoming items.

In this Section we consider possible criteria that may help decide if implicit tests are tapping a mode of retrieval different from standard explicit tests. Following Jacoby (1984) and Schacter et al. (1989), we distinguish intentional from unintentional retrieval and assume that explicit memory tests primarily invoke the former and implicit memory tests mostly involve the latter. (Retrieval in this context is used broadly — accessing information to respond to demands of the test.) The critical issue then becomes determining if subjects are predominantly employing one or the other mode of retrieval in responding to cues on a test.

The defining criterion of the distinction between explicit and implicit tests is the instructions subjects are given at the time of test: on an explicit test they are told to recollect events from an earlier time; on an implicit test they are told simply to perform a task as well as possible, with no mention made of recollecting earlier experiences in performing the task. In some cases subjects are told that words they produce on the test may be from the study list, but that they should always respond with the first word that comes to mind (Tulving, Schacter and Stark, 1982).

Although a necessary condition for instantiating the two types of test, a simple instructional manipulation may not be sufficient to do so. Subjects in the implicit test condition may, for whatever reason, invoke recollective strategies to solve the problems they are given. This contamination of implicit tests by the intentional use of memory has been the primary concern of researchers, but in our opinion the converse problem — incidental or automatic retrieval affecting performance on explicit tests — is at least as likely (indeed, we believe more likely under certain circumstances) and just as troublesome (Gardiner, 1988; Jacoby, 1993). In general, we believe that contamination of implicit tests by intentional retrieval is more likely when test cues are meaningful items such as general knowledge questions (as in the hypothetical example above) or even single words (e.g., associates), but less likely when the test cues are data-limited items.

² Experiments like this one have been conducted with memory-impaired patients, who do show evidence of intact priming without conscious recollection of prior episodes (Schacter, Harbluk and Mclachlan, 1984). Answers by normals in the transfer phase might also be due to automatic priming. The issue at stake here is how we can decide.
such as word fragments. Conversely, influences of incidental retrieval on explicit tests may be more likely when data-limited cues are used and less likely under conditions of semantic cues or free recall. However, these suppositions are merely hypotheses to be tested at this point.

Thus, although instructions given to subjects at test nominally establish the retrieval orientation (to use the phrase of Nelson, Keeler and Negrao, 1989), additional criteria must be satisfied to establish that the tests involve incidental retrieval. We consider several criteria suggested in the literature: performance of memory-impaired subjects, dissociations between tests, questionnaires given to assess awareness, the retrieval intentionality criterion, the process dissociation procedure, and the method of triangulation and stochastic independence. Instructions are critical in several of these techniques, such as the retrieval intentionality criterion and the process dissociation procedure, discussed below, but they are combined with other experimental features to provide more powerful inferences in these techniques.

Evidence from memory-impaired patients

Memory-impaired patients, by definition, are those who show deficits on explicit memory tests. Cognitive psychologists’ current fascination with studying implicit memory tests originated in the finding of preserved priming in amnesics on many different implicit tests (see Moscovitch et al., this Volume; Shimamura, 1986). Thus, another defining characteristic of implicit memory tests might be that amnesics and other memory-impaired patients show preserved priming on them. Although preserved priming is a useful guide, there are difficulties in using this criterion to judge a task as involving only incidental retrieval. First, memory impairments may arise from many sorts of brain damage and it may be that some implicit tests are affected by such damage, as are explicit tests. Heindel, Salmon, Shults et al. (1989) and Blaxton (1992), for example, present cases of patients impaired on some implicit tests but spared on others. Second, even if a test passes the criterion of being preserved in memory-impaired patients, there is no guarantee that intentional retrieval processes will not be introduced when the task is used with normal subjects. Therefore, although one might be leery of deeming a memory test as reflecting incidental retrieval unless some memory-impaired subjects showed intact priming on it, the converse conclusion does not hold: finding intact priming in memory-impaired patients provides no guarantee that, when used on normals, the test will measure incidental retrieval.

Questionnaire techniques for assessment

One way to attempt to determine if an implicit test is contaminated by intentional retrieval strategies is to query the subjects, after the test, about the strategies they used in completing the task. Bowers and Schacter (1990) have used this technique with success to separate ‘test aware’ from ‘test-unaware’ subjects (see too McAndrews and Moscovitch, 1990; Richardson-Klavehn, Gardner and Java, 1993). In Experiment 1, Bowers and Schacter used a standard word stem completion test, giving subjects a long list of words and later asking them to complete word stems with the first word that came to mind. They varied (a) whether subjects expected a memory test, (b) whether they expected to complete word stems with studied words, and (c) whether the orienting tasks during study engaged structural (shallow) processing or semantic (deep) processing. Following the test phase, subjects’

**TABLE 2**

Results from Bowers and Schacter (1990): Mean proportions of word stems completed with target words as a function of awareness

<table>
<thead>
<tr>
<th></th>
<th>Self-report</th>
<th>Level of encoding</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Semantic</td>
<td>Structural</td>
</tr>
<tr>
<td>Test aware</td>
<td>0.43</td>
<td>0.23</td>
<td>0.33</td>
</tr>
<tr>
<td>Test unaware</td>
<td>0.33</td>
<td>0.28</td>
<td>0.31</td>
</tr>
<tr>
<td>M</td>
<td>0.38</td>
<td>0.26</td>
<td></td>
</tr>
</tbody>
</table>
awareness of the relation between study and test was assessed by a questionnaire. Test-aware subjects were those who said that, at some point during the experiment, they had realized that some items they had produced on the test had appeared on the list studied earlier. Test-unaware subjects denied such awareness. Interestingly, both groups of subjects showed equivalent priming effects (see Table 2), but test-aware subjects showed larger levels of processing effect than did test-unaware subjects. The aware subjects produced reliably more responses following the semantic condition than did the unaware subjects, indicating that responding may have been augmented by intentional retrieval.

Although separating subjects into test-aware and unaware conditions did not produce major changes in the pattern of results in the standard word stem completion test, Bowers and Schacter (1990) reported more dramatic differences in a paradigm testing for implicit memory for new associations (Graf and Schacter, 1985). Therefore, separating subjects into different groups on postexperimental questionnaires may represent a useful technique for assessing awareness and eliminating subjects who use intentional strategies on tests designed to assess unintentional retrieval. However, such questionnaire techniques only assess awareness retrospectively and are subject to the complaints often raised about such methods, such as forgetting of the true mental state at the time of completion, compliance with the experimenter’s demands, etc.

In addition, one must be careful in defining awareness on these tests. Subjects may become aware of the relation between their responding and the original study list in several ways (see Schacter et al., 1989). If subjects become aware of the relation only after producing a response, then the test may still assess unintentional retrieval. The greater priming that may occur on the test may lead to greater opportunities for such retrospective awareness; in one scenario, greater priming (via unintentional retrieval) may lead to more ‘aware’ subjects. Questionnaires are unlikely to capture these complex issues. Others have tried to develop more refined procedures to assess subjects’ states of awareness during recollection (e.g., Gardiner, 1988; Rajaram, 1993; Tulving, 1985b), although the judgments in these cases are made during explicit tests (see Gardiner and Java, 1993, for a review).

Despite these problems, Bowers and Schacter (1990) have shown that postexperimental questionnaires can be used with some success, as have others (McAndrews and Moscovitch, 1990; Richardson-Klavehn et al., 1993). Interestingly, the problems from subjects becoming test-aware and then using intentional retrieval on implicit tests seem relatively slight in the standard word stem completion test, based on their results.

Dissociations between tests

Probably the most common way that experimental psychologists have tried to show that a test is implicit is to dissociate it from an explicit test within an experiment. For example, Jacoby and Dallas (1981) manipulated levels of processing of words during study and then tested one group of subjects on a word identification test and a separate group of subjects for explicit recognition of the words. Levels of processing had a large effect on the explicit recognition test, but no effect on primed word identification. The latter test was not simply insensitive, because reliable priming above the nonstudied baseline was obtained, and other variables (such as presentation modality) affected priming. One conclusion to be drawn from such dissociations is that the two tests are sensitive to different variables, and this case is strengthened when other variables can be shown to produce opposite effects on the two tests (Jacoby, 1983b). Such patterns of results certainly bolster the case that the two tests draw on different modes of retrieval (Roediger and Blaxton, 1987a), but dissociations between tests do not necessarily indicate that one is measuring intentional retrieval and the other incidental retrieval. The reason is that many variables can dissociate explicit tests (e.g., Tulving, 1983, Chapter 11) and evidence is growing that the same is true for implicit tests (Roediger, Srinivas and Weldon, 1989a). Therefore, simply showing that two tests are dissociated does not mean...
that they draw on different modes of retrieval, although in many cases this may indeed be the correct conclusion (Dunn and Kirsner, 1988).

Dissociations between explicit and implicit forms of test, such as word identification and recognition, certainly bolster the conclusion that the implicit test is measuring unintentional retrieval, but do not seal it. Typically, the test situations differ in other ways, especially when a test such as word fragment completion is contrasted with another test employing a very different set of test cues, or even without cues (i.e., free recall; Roediger and Challis, 1992; Weldon and Roediger, 1987). A contrast between free recall and primed word fragment completion (or some other perceptual implicit test) may be informative for those interested in distinguishing conceptual forms of memory from perceptual forms, but the conceptual/perceptual contrast is confounded with the explicit/implicit contrast and so cannot be taken for evidence of an effect of retrieval orientation (Blaxton, 1989). Still, dissociations between explicit and implicit tests are an important, if not unimpeachable, type of evidence supporting the distinction between intentional and unintentional retrieval. The most powerful form of such evidence comes from incorporating dissociations between tasks within the framework of the retrieval intentionality criterion, as discussed below.

The retrieval intentionality criterion

The retrieval intentionality criterion (Schacter et al., 1989) represents an extension of the test dissociation logic described above. Dissociations between tests may be caused by task differences that exist between (say) free recall relative to fragment completion and thus differences observed may be due to factors other than the mode of retrieval (intentional or incidental) required by the test. Therefore, Neely and Payne (1983) and Neely (1989) strongly advocated holding constant (as much as possible) extraneous variables that affect performance when comparing tasks that are believed to tap different memory systems or modes of retrieval.

The basic idea in the retrieval intentionality criterion is to hold constant all overt conditions affecting subjects at study and test except for the instructions given at the time of the test. For example, Graf and Mandler (1984) manipulated levels of processing at study by having subjects rate words for semantic features or for nonsemantic features. At the time of test, all subjects received word stems, but half were given implicit test instructions (say the first word that comes to mind beginning with these letters), whereas the other half were given cued recall instructions (recall an item from the earlier studied list that corresponds to the stem). Under these conditions, the only difference between the two groups is the instructions given to subjects before the test — a pure instantiation of the explicit/implicit retrieval orientation because nothing else is varied. They found that manipulating levels of processing at study affected the explicit form of the test, but had little effect on the implicit stem completion test (Graf and Mandler, 1984). This dissociation is like that of Jacoby and Dallas (1981) mentioned above, but now the test cues are held constant except for the critical instructions used to implement the tests.

The retrieval intentionality criterion has been used by Roediger, Weldon, Stadler et al. (1992) to determine that word fragment completion also represents a test of incidental retrieval, at least under their conditions. Level of processing affected intentional recall of words from word fragment cues, but not priming on the implicit test. Graf and Schacter (1987, 1989) have also used this logic for determining that memory for new associations can also be exhibited on implicit tests. However, this measure was deemed not a very pure measure of unintentional retrieval in Bowers and Schacter’s (1990) questionnaire study, and seems to fail the criterion of being preserved in patients with severe memory impairments (Cermak, Bleich and Blackford, 1988; Schacter and Graf, 1986b; Shimamura and Squire, 1989; but see Moscovitch, Winocur and MacLachlan, 1986) and in the aged (Howard, Fry and Brune, 1991).

The retrieval intentionality criterion requires a dissociation between explicit and implicit forms of a test when all conditions except instructions are
held constant. However, because test cues are themselves powerful causes of responding, it may be that the power of the cues will produce parallel patterns of performance on two tests even if one measures incidental and the other intentional retrieval, with little or no contamination. That is, parallel effects on two tests do not indicate that both are tapping intentional or unintentional retrieval, but merely that no strong claim can be made about whether they are the same or different. The researcher seeking to show that the tests differ in this way must keep searching for a variable to dissociate them. So far, researchers have succeeded when the test cues are data-limited word fragments, but we suspect that the situation may change when the retrieval intentionality criterion is used to distinguish explicit and implicit conceptual tests. Some theorists (e.g., Roediger, 1990a) propose that similar factors (elaboration, organization, categorization) are at work when subjects are engaged in both explicit and implicit tests with the same conceptual cues, such as category names. However, this is an empirical question and at least some researchers find different patterns of results even with conceptual cues on explicit and implicit versions of a test (e.g., Rappold and Hashtroudi, 1991). In addition, Challis and Sidhu (1993, Experiment 3) reported effects of differing magnitude from massed repetition on implicit and explicit conceptual tests. The retrieval intentionality criterion should prove one quite useful means of establishing whether a test draws upon unintentional or intentional retrieval processes, even if there are some situations in which this technique may not be useful, such as when semantic cues are given.

Merikle and Reingold (1991) have proposed an interesting extension to the retrieval intentionality logic of comparing explicit and implicit memory tests, one derived from their studies of unconscious perception (Reingold and Merikle, 1988). The basic idea is to show a greater effect of a variable on an implicit than on an explicit test, while holding constant all study and test conditions except for instructions. They argue that... Whenever an indirect measure indicates greater sensitivity than a comparable direct measure, it must reflect greater sensitivity of the indirect measure to unconscious, task relevant information’ (Merikle and Reingold, 1991, p. 225) because if the information were available for conscious decisions, it would be used on the direct or explicit tests. Note that the results of Graf and Mandler (1984) and Roediger et al. (1992), although satisfying the retrieval intentionality criterion, do not meet this more stringent criterion.

Several reports have appeared in which a variable has produced a greater effect on an implicit measure of retention than on an explicit measure (e.g., Kunst-Wilson and Zajonc, 1980; Mandler et al., 1987; Seamon, Marsh and Brody, 1984). All these experiments compared subjects’ old/new recognition judgments (an explicit test) to other sorts of judgments (preferences, brightness, darkness). Merikle and Reingold’s (1991) experiments showed that judgments of contrast were more sensitive, under certain conditions, to manipulation of independent variables than were recognition decisions.

Few experiments using more typical priming paradigms satisfy Merikle and Reingold’s (1991) criterion. Eich’s (1984) results using a homophone spelling task do, but homophone spelling has not been a widely used measure. The results described thus far satisfying this criterion represent an existence proof — it can be done — but no one has further explored the matter. Although this criterion is interesting and convincing, in our opinion the other means of satisfying the retrieval intentionality criterion (having a variable affect performance on an explicit test but not affect priming on an implicit test, with all factors but instructions held constant) is also a useful means of separating intentional and incidental retrieval modes.

We have discussed cases in which all conditions except instructions are held constant at study and test, and an independent variable affects either the explicit or the implicit test, while leaving the other unaffected. Are there situations in which a variable can be manipulated and be shown to have opposing effects on explicit and implicit tests while all else is held constant (except test instructions)? Such a case
TABLE 3

Results from Java (1993: Experiment 2) (Proportion correct on the explicit and implicit tests with words stems as cues. Data for the Read and Generate conditions of the implicit test are priming scores.)

<table>
<thead>
<tr>
<th>Test</th>
<th>Nonstudied</th>
<th>Read</th>
<th>Generate</th>
<th>G − R</th>
</tr>
</thead>
<tbody>
<tr>
<td>Explicit</td>
<td>0.04</td>
<td>0.22</td>
<td>0.41</td>
<td>0.19</td>
</tr>
<tr>
<td>Implicit</td>
<td>0.14/0.15</td>
<td>0.33</td>
<td>0.19</td>
<td>−0.14</td>
</tr>
</tbody>
</table>

would provide quite powerful evidence for different retrieval modes. The answer is yes: Java (1993) had subjects read words or generate them from conceptual clues during a study phase. She then tested them with word stems as cues under either explicit or implicit test instructions. Her results appear in Table 3. Generating produced better performance than reading on the explicit tests, whereas more priming was produced from reading than from generating on the implicit test. Thus Java’s results satisfy the retrieval intentionality criterion with a variable that produces opposing effects on the two tests.

One difficulty with the retrieval intentionality criterion is that, if a test fails to meet the criterion, then there is no solution to the problem other than to redo the experiment under different conditions. The next technique has the advantage of providing estimates of automatic and consciously controlled responding on any test, implicit or explicit.

**Process dissociation procedure**

The main difficulty in separating intentional retrieval processes from incidental processes is that both are expected to facilitate performance in the standard paradigms. That is, if subjects are given (say) a word stem completion test after studying a list of words, they could show facilitation in completing the word stems with a word from the list either from priming (incidental retrieval), from conscious recollection (intentional retrieval), or from some combination of the two. Because facilitation occurs from both sources, separating the influences is difficult. The process dissociation procedure developed by Jacoby (1991) accomplishes separation of these processes by comparing two test conditions in which cues are held constant but instructions to subjects differ.

The general logic is for subjects to receive a list of items (e.g., hotel), usually under different study conditions within the list. Then subjects are tested under one of two instructional sets. One group is instructed to respond to cues on the test with items from the list or, failing that, to produce any items that fit the cue. This is called an inclusion condition, so named because subjects try to include studied items that fit the cue. Responding in the inclusion condition could occur because subjects either intentionally or unintentionally retrieve words to the cues. A second group of subjects, tested in an exclusion condition, gets the same study phase and test cues, but is told to respond only with items that were not in the study list (i.e., to exclude items that occurred in the list and to respond only with items not in the studied set). The exclusion condition embodies the logic of opposition of processes: unintentional retrieval processes promote responding with studied items in the exclusion condition, whereas intentional retrieval processes inhibit responding. If subjects produce a studied item when they are trying not to do so (in the exclusion condition), then this is evidence that its production was automatic. Producing the studied response could not be under the controlled influence of explicit recollection; if subjects could recall that the item was on the list, then they should not produce it on the test.

Let us consider a recent experiment by Jacoby, Toth and Yonelinas (1993b) to illustrate the logic. In their Experiment 3, subjects either read intact words during the study phase (e.g., bowry) or solved them from anagrams (random rearrangements of the letters in the words but with two letters always in their correct places [e.g., yordw]) in each of two lists. Subjects were always given word stems during the test (dow _____) with inclusion or exclusion instructions. In the inclusion test condition, subjects were told to complete the stems with a word from the list (either anagrams or read words), or failing that, the
first word that came to mind. Subjects in the exclusion condition were told to complete the word stems with the first word that came to mind, but to exclude any words that had appeared in the prior lists. Therefore, intrusion errors under the exclusion condition instructions (i.e., writing down a word from the list) that occur with more frequency than in a nonstudied baseline condition indicate an automatic influence of memory.

Consider performance under the two test conditions. The probability of completing a stem with a studied word under the inclusion instruction condition (above baseline levels) can be expressed as the probability of recollecting it via intentional retrieval ($P_R$), plus the probability of its being produced via an automatic influence ($P_A$), with the assumption that these two sources act independently ($P_A(1 - P_R)$). Restated, the probability of producing a response in the inclusion condition is $P_{St, Incl} = P_R + P_A(1 - P_R)$. Consider now the exclusion condition: a studied word would be produced (above baseline) under exclusion instructions only in the case in which recollection failed but the automatic component exerted an influence. Thus, the probability of producing a studied word under exclusion conditions ($P_{St, Excl}$) can be expressed as $P_A(1 - P_R)$. Note that this latter term is included in the equation above; therefore, the probability of recollection can be solved as simply the difference between response production under inclusion and exclusion conditions, $P_R = P_{St, Incl} - P_{St, Excl}$. Once $P_R$ has been obtained, then $P_A$ can be calculated by algebra. The easiest way is $P_A = P_{St, Excl}$ divided by $(1 - P_R)$ or a failure in recollection. The details of the logic can be found in Jacoby (1991) as well as in Jacoby et al. (1993b).

Returning to the experiment, the results appear in Table 4. Under the inclusion response condition we see that probability of response is the same for words read and solved from anagrams. This is a bit of a surprise, because normally we would expect to see (under explicit test conditions) a generation effect, with better performance from anagrams than from read words (e.g., Allen and Jacoby, 1990). But the inclusion condition is not a normal explicit test condition because subjects are told to respond either on the basis of conscious recollection or with the first word that comes to mind (reflecting both intentional and incidental retrieval). As we see from the remainder of the table, the apparent null result actually arises because of competing forces from the two sources of responding. From the second column, we see that subjects were better able to exclude items if they had solved them from anagrams than if they had read them. That is, the probability of producing an old item when the instructions indicated that these items should not be produced was lower for those items that had been presented as anagrams, indicating that conscious recollection more successfully opposed automatic responding in the anagram condition. Subtracting $P_{St, Excl}$ from $P_{St, Incl}$ to estimate the probability of recollection ($P_R$) in the third column shows that recollection was more successful for items generated from anagrams than for read items. However, in solving for $P_{A1}$ in the fourth column, we now see that automatic responding is greater after reading the word than after solving the anagram. The $P_{A1}$ estimate includes responding from an automatic process and the baseline of completing fragments; the $P_{A2}$ measure, on the far right, subtracts out the baseline performance for a purer measure of automatic responding. Now we see that an automatic (presumably data-driven) component exists only for prior reading of words.

<table>
<thead>
<tr>
<th>Stimulus</th>
<th>$P_{St, Incl}$</th>
<th>$P_{St, Excl}$</th>
<th>$P_R$</th>
<th>$P_{A1}$</th>
<th>$P_{A2}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word</td>
<td>0.82</td>
<td>0.49</td>
<td>0.33</td>
<td>0.73</td>
<td>0.17</td>
</tr>
<tr>
<td>Anagram</td>
<td>0.82</td>
<td>0.25</td>
<td>0.57</td>
<td>0.59</td>
<td>0.03</td>
</tr>
</tbody>
</table>

$P_{St, Incl}$ refers to the probability of responding with a studied word in the inclusion condition, $P_{St, Excl}$ to the probability of responding with a studied word in the exclusion condition, $P_R$ to the probability of recollection, $P_{A1}$ to the probability of responding due to the automatic component and the baseline, and $P_{A2}$ to the probability of automatic responding with the baseline responding removed.
Thus we see that the simple manipulation of reading and generating has opposing effects on the two components of performance. Under most test conditions, responding is likely to be a mixture of automatic and controlled processes and therefore the process dissociation procedure is useful in disentangling the two bases of responding. Although the process dissociation procedure is new, it holds potential promise in separating intentional from incidental retrieval; the early findings are compelling (Jacoby, 1991; Jacoby et al., 1993b; Jacoby, Ste-Marie and Toth, 1993a).

Some difficulties, however, exist for the method. One is the assumption that recollection and the automatic component of responding are independent of one another, which needs to be evaluated. Also, the exclusion condition is potentially fraught with difficulties. Subjects may adopt different criteria for excluding responses depending on instructions and task demands, which will result in variable estimates of $P_R$ and $P_A$. More seriously, subjects may exclude responses on several distinguishable bases: (a) they may intentionally recollect the item and exclude it, or (b) they may unintentionally retrieve the word and then recognize it and exclude it (Richardson-Klavehn et al., 1993). If subjects covertly produce the response through an automatic process (or through incidental retrieval) and then later exclude it on the basis of recognizing it as being from the list, then the process dissociation procedure will overestimate how contaminated implicit tests are by intentional retrieval processes. Richardson-Klavehn et al. (1993) suggest that this is indeed the case. Whether these potential problems turn out to be serious or trivial must await future research.

**Method of triangulation and stochastic independence**

A final method we consider for separating explicit from implicit processes is Hayman and Tulving’s (1989a) method of triangulation. Like the process dissociation procedure, we cannot present the technique fully here, but will sketch in the general logic (which hinges in part on the retrieval intentionality criterion logic). One way of showing that two tests are measuring independent processes is by giving subjects two successive tests following study of a list (say an explicit recognition test and an implicit word fragment completion test) and examining the correlation between performance on the two tests across items. Are items that are well recognized on the first test those that are completed best on the second test? Tulving et al. (1982) showed that there was essentially a zero correlation between recognition of words and priming on a word fragment completion test, as long as the recognition test preceded the completion test.3 Others have reported similar findings with different explicit and implicit tests, although the usual explicit test is recognition and it must precede the implicit test for the reasons given in footnote 3 (e.g., Jacoby and Witherspoon, 1982).

A case can be made that such stochastic independence is a convincing criterion for incidental retrieval without conscious recollection in normal subjects: one and the same subject who shows no awareness of having seen some words from the study episode on the recognition test shows as much priming for those words as for the ones that are recognized. This simulates in normal subjects the results with amnesic patients that provided the initial impetus for implicit memory research: complete failure in recollection (even recognition) but intact priming on an implicit test.

Many researchers are unconvinced by such demonstrations of independence between tests because competing explanations exist for the lack of correlation between tests. Test ‘priming’ on the first test is alleged to influence performance on the second and to underlie the finding of independence (Shimamura, 1985); the priming measure may vary too little for reliable correlations to be observed (Ostergaard, 1992); or subject by item selection effects may

---

3 If the order is reversed, a positive correlation is induced. Items successfully completed are better recognized than those not completed. Presumably, successful completion acts as an additional study opportunity and makes recognition more likely.
operate in different directions to cancel out correlations that exist and give an illusion of independence between tests (Hintzman and Hartry, 1990). Each of these claims about stochastic independence is hotly debated (e.g. Flexser, 1991; Gardiner, 1991; Hintzman, 1991), but the method of triangulation would seem to overcome most of these complaints.

In the method of triangulation, subjects are given three tests (hence the name). Hayman and Tulving (1989a) had subjects study a list of words and then tested their recognition memory for the words. Following this test, two groups of subjects were tested with word fragment cues, but under different sets of instructions (as in experiments using the retrieval intentionality criterion). One group of subjects was told to use the fragments as retrieval cues and to recall words from the list that corresponded to these fragments (explicit test instructions). The other group of subjects was told to complete the fragments with the first words that came to mind (implicit test instructions). Note that priming from the recognition test, item selection, and other factors are better controlled here than in other studies because both groups receiving fragment cues have the same prior experimental history. The results showed a strong dependency between recognition and explicit cued recall with word fragments as cues, but little or no dependency between recognition and the implicit fragment completion test. As long as there is an adequate amount of priming under the conditions of the experiment so that constricted range problems will not produce artificially low correlations (Ostergaard, 1992), the method of triangulation should prove useful in separating tests into those requiring intentional retrieval and those reflecting incidental retrieval. Hayman and Tulving (1989b) and Challis, Chiu, Kerr et al. (1993) provide further evaluation of the method of triangulation.

Conclusion and recommendations

We have reviewed several methods that have been used for separating explicit from implicit memory. Three of these (the retrieval intentionality criterion, the process dissociation procedure, and the method of triangulation) are so new that not much use or critical analysis has been applied to them as of this writing. These are potentially the strongest procedures for drawing inferences, so no doubt these methods will receive careful scrutiny.

Given this state of affairs, we now make some suggestions as to how to conduct experiments that should produce relatively clean data from implicit memory tests using priming paradigms under normal circumstances. These recommendations are meant as guidelines, based both on prior experiments and logical considerations. First, subjects should be tested under conditions of both incidental learning (exposing material to them under some guise, such as norming materials for future research) and incidental (or unintentional) retrieval instructions. In our opinion, the evidence indicates that the test conditions are much more important than the study conditions (Roediger et al., 1992). Second, the instructions for responding during the test should encourage subjects to respond rapidly to each cue with the first item that comes to mind. We suspect subjects usually comply; it is often easier merely to respond quickly with whatever comes to mind than to engage in effortful retrieval, and the law of least effort will help insure more automatic responding, in Jacoby’s (1991) sense of automaticity. If the test cues are provided only briefly this will also make it more likely that a perceptual test will be based on incidental retrieval (Weldon, 1993). Third, use large sets of materials so that the possibility of using intentional strategies, even if tried, will be of little use (at least if subjects try to free recall items and match them to test cues). Under normal conditions, we would recommend list lengths in the range of at least 50 or so items and twice that number is not uncommon. (With short lists, say 10 – 20 items, subjects may be more likely to use explicit strategies). Fourth, use several filler tasks between the study list and the implicit test, so that the test seems just one more task in a long series. Some filler tasks may be made similar to the critical implicit task, but of course would include only nonstudied items. Fifth, begin the implicit memory test with 10 – 15 filler items that do not correspond to list items to en-
courage responding with the first item that comes to mind (and to discourage any attempt at intentional recollection). Sixth, keep the proportion of studied words in the test below 50%, for the same reason.4 Seventh, in most circumstances, include an explicit test in the experiment to show a dissociation between the implicit and explicit test. The strongest case would be to keep the overt test cues the same between the two tests and to vary only the instructions, thereby employing the logic of the retrieval intentionality criterion. This advice will also depend on the problem area under investigation, as effects on some explicit tests are so regular as not to require replication every time an implicit test is used (although at least one experiment might be necessary to show that the effect holds under the particular conditions employed).

Admittedly, some of these recommendations are not on secure empirical grounds (filler tasks, short lists, incidental learning, low proportion of overlap) and need evaluation. Still, they may be useful in quelling critics, even if they do not alter subjects' behavior. Some evidence exists that supposed contamination of implicit tests by intentional retrieval has been overblown (e.g., Richardson-Klavehn et al., 1993; Roediger et al., 1992) and may be receiving too much attention here and elsewhere.5

Having now considered the various criteria for establishing an implicit test as tapping incidental retrieval, we must admit that in the remainder of our review of the literature we largely dodge this issue and simply accept at face value claims in experimental reports. Most experiments do provide a dissociation between priming on an implicit test and performance on an explicit test, but this evidence is not particularly conclusive for reasons given above. The newer, more powerful methods of distinguishing intentional from incidental retrieval — the retrieval intentionality criterion, the process dissociation procedure, the method of triangulation — have not yet been widely enough applied to provide information about many variables.

Variables affecting priming on implicit memory tests

Although the study of implicit memory in normal subjects has been investigated only for about the past 10 years, the literature is already surprisingly large. In some cases findings are consistent, but in many others the effect of basic variables is unclear. There are occasional failures to replicate basic findings; different results are obtained using different measures and in different labs; and occasionally the worry exists (especially when some problematic result is obtained) that an implicit test may be contaminated by intentional retrieval strategies. Although this state of affairs is frustrating, it is probably not too surprising for a new area of study. It will take concentrated experimental effort to sort out many of the puzzles that already exist in the literature, and this process could take several decades, assuming researchers remain interested that long. We may hope that the new inferential techniques reviewed above will clarify the situation.

In the succeeding Sections, we review the main variables that have been investigated in experiments examining priming on various implicit memory tests. We include mostly variables under experimental control and exclude individual differences and other physiological variables such as the effects of drugs and the like. We include coverage of the 12 implicit memory tests in Table 1, but do not cover all the work on several others for reasons already mentioned. The main sections in the review involve

---

4 Actually, evidence on this advice is mixed, with Challis and Roediger (1993) reporting no effect of proportion of old items on a word fragment completion test (but see Jacoby, 1983 and Allen and Jacoby, 1990). However, this strategy of using a low proportion of studied items in the tests may be useful for satisfying reviewers (even if it does not affect subjects’ behavior), which is just as important to authors.

5 In commenting on an earlier version of this Chapter, Endel Tulving wrote that ‘‘your hand-wringing over how to tell a ‘pure’ implicit task, or implicit memory performance is a bit overdone … . The issue is not that important at this early stage of the game … . Putting it into the sharp focus of researchers’ attention may overly cramp their style. Instead of finding out about the real brain/mind, they — and editors and referees — may get fixed on the problem of how ‘pure’ are implicit and explicit tasks.’’

79
manipulations of (a) materials, (b) surface features of studied items, (c) presentation time and repetition, (d) instructional manipulations at study, (e) retention interval, and (f) some miscellaneous variables that do not fit comfortably into this scheme. This categorization works reasonably well, but some overlap occurs especially between variables in (a) and (b). In some cases we present data from the literature that, in our opinion, best represents the current state of knowledge in the field. More frequently, we present verbal summaries of the evidence.

**Materials**

We include in this section the effect on implicit tests of (a) presenting pictures and words, (b) presenting information in one or another language for bilingual subjects, (c) manipulating lexical variables (i.e., frequency of occurrence in the language, orthographic distinctiveness, and the number of graphemically similar words existing in the language) on verbal implicit tests, (d) indirectly priming a target word by presenting semantically- or phonologically-similar words, and (e) presenting nonsense materials (nonwords for verbal tests, nonobjects or relatively meaningless patterns for nonverbal tests).

**Picture/word manipulations**

On the standard, explicit tests of memory such as free recall, paired associate learning, and recognition, there exists a picture superiority effect: presentation of pictures during a study phase results in better retention than presentation of words, the names of the pictures (Madigan, 1983). This effect is generally attributed do pictures being encoded in two formats (pictorial and verbal) and words more likely being encoded in one format (verbal; see Paivio, 1986), although others attribute the effect to quicker, fuller encoding of pictures’ meaning (Nelson, 1979). Against this backdrop, the surprising finding from many studies of implicit memory using verbal, perceptual tests is that words produce more priming than to pictures. Winnick and Daniel (1970) showed more priming from words than from pictures on a word identification test (as did Kirsner, Milech and Stumpfel, 1986, and Weldon, 1991) and Durso and Johnson (1979) showed the same pattern on word naming. Word superiority has also been obtained in priming on several other verbal tests: word fragment completion (Roediger et al., 1992; Srinivas and Roediger, 1990; Weldon, 1991; Weldon and Roediger, 1987; Weldon, Roediger and Challis, 1989), word stem completion (Roediger et al., 1992; Weldon et al., 1989), anagram solution (Srinivas and Roediger, 1990), and the lexical decision task (Kroll and Potter, 1984).

Weldon (1991) directly compared priming on both word identification and word fragment completion from visually presented words, auditorily presented words, words generated from conceptual clues, and pictures. Pictures produced the least priming on both tests. This reversal of the typical picture superiority effect strongly implies that these tests are perceptual or data-driven; prior study of words leads to greater transfer on these tests because of the similarity in procedures required to read words and to decipher later their fragmented forms. There is some slight priming from pictures that is consistently observed in experiments employing verbal perceptual tests. This cross-form priming may arise from subjects covertly labeling the pictures as they are presented, but direct tests of this hypothesis have at best provided weak evidence for it (Weldon and Jackson, 1993), and other tests have yielded no evidence for it (Roediger and Weldon, 1987; Weldon and Roediger, 1987). The important point remains that on verbal, perceptual implicit tests, words produce more priming than do pictures, contrary to the usual picture superiority effect in standard explicit tests. (Again, on word recognition tests, prior study of pictures leads to better performance than prior study of words.)

The situation is quite different when nonverbal (pictorial) perceptual tests are considered. Now the match between the data-limited cues is greater for prior study of pictures than words. Most researchers have found greater priming from pictures than from words on tests of picture fragment naming (McDer-
mott and Roediger, 1992b; Srinivas 1993; Weldon and Roediger, 1987), identification of pictures from brief displays (Warren and Morton, 1982), and speed of naming intact pictures (Biederman, 1987; Durso and Johnson, 1979). This is exactly the pattern predicted by the transfer appropriate processing approach (Roediger et al., 1989), but the claim is that this type of picture superiority effect occurs for a different reason (overlap between perceptual processes of encoding and testing) than the standard picture superiority effect (overlap of conceptual elaboration at study with tests that draw on conceptual processes). It should be possible to test this claim (e.g., by dissociating the two picture superiority effects), but no one has tried yet.

The one exception to picture superiority on picture naming or picture fragment naming has been reported by Brown, Neblett, Jones and Mitchell (1991), who found that pictures and words produced equivalent priming when subjects had studied only one type of material before the test (i.e., in a between-subjects design), although they replicated the standard effect in within-subjects designs. It is unclear why they failed to replicate the picture superiority effect on picture naming, because others have reported finding greater priming from pictures than words on picture naming even in within-subjects designs (Biederman, 1987; Gabrieli, personal communication; Lachman and Lachman, 1980).

To our knowledge, no published reports analyze the effects of studying pictures and words on conceptual implicit memory tests such as category association or answering general knowledge questions. The prediction from the transfer appropriate processing approach is that the picture superiority effect should be present on this form of test (despite its verbal nature) because of the presumed conceptual nature of the standard picture superiority effect. These matters await future research. For the moment the situation with regard to this variable can be stated with relative clarity: on all verbal, perceptual implicit tests, words produce more priming than do pictures, whereas the converse is the case in almost all studies of nonverbal perceptual tests.

**Priming from alternate languages in bilinguals**

The comparison of the effects of pictures and their labels on memory permits (very roughly) the same concept to be presented in two different ways, one verbal and one pictorial. Presenting bilingual subjects with words in one or both of their two languages also permits interesting studies of the ‘same’ information represented in two symbol systems, but now within verbal systems of communication. If bilinguals are presented with a word in one language, will they show priming on implicit memory tests when the test is given in an alternate language? Thus far no studies have addressed this issue using conceptual implicit memory tests, but several have examined the impact of cross-language presentations on perceptual implicit memory tests. The most popular measure has been priming in the lexical decision task and studies of this measure have yielded fairly consistent effects. In the semantic priming type of experiment, where test items are presented in close temporal proximity, translation equivalents (e.g., *caballo* to *horse*) prime lexical decisions; however, when the delay between the first and second presentation is lengthened, there is little or no priming from translation equivalents, but repetition priming (*horse to horse*) remains intact (Kirsner, Brown, Abrol et al., 1980; Kirsner, Smith, Lockhart et al., 1984; Scarborough, Gerard and Cortese, 1984). Translation equivalents can produce long-term priming when subjects are told to translate the first term when it is presented, but this finding implies that subjects do not do so unless so instructed. If words in the two languages are morphologically similar, then cross-language priming will occur. Taken together, such results argue that language-specific lexical processes mediate priming on these perceptual implicit memory tests. Gerard and Scarborough (1989) report strong evidence that it is the word form (divorced from the word’s meaning) that produces priming in the bilingual situation on the lexical decision task.

Such specificity of priming does not only occur in the lexical decision task, but also has been obtained in word fragment completion (Durgunoğlu and Roediger, 1987) and word stem completion (Wat-
kins and Peynircioğlu, 1983). In both cases, cross-language priming was small and not significant, but same-language priming was robust.

One exception to these findings has been reported by Smith (1991), who argued that when words are read in the context of a sentence (relative to in a list), more conceptual processing of the word is engaged and subjects should show cross-language priming. Her results from two experiments supported this point of view: priming from translation equivalents occurred following presentation of words in sentences, but not following presentation of words in lists. Her results suggest that primed word fragment completion can be sensitive to semantic factors. Competing hypotheses are that intentional retrieval may have been involved in Smith’s (1991) tests (because the lists were rather short) or that covert translation was involved because of the difficulty of the materials. Future research will tell how susceptible bilingual priming is to influence by conceptual factors, but some other research argues that conceptual effects may be larger than the early studies would have led us to expect (e.g., Komatsu and Naito, 1992). However, most of the research reviewed previously reveals language specificity on verbal (perceptual) implicit memory tests even under conditions in which there can be no doubt that subjects processed the meaning of the materials during study (e.g., Durgunoğlu and Roediger, 1987, had subjects rate words’ pleasantness and found little cross-language priming).

**Lexical variables**

In this Section we include such factors that might affect verbal priming as word frequency, distinctive orthography, and the number of graphemically similar words that exist. Unlike most of the other research reviewed in this paper, studies of these variables are correlational in nature. Rather than the same concepts being used in various forms (as when such variables as modality are examined), researchers must select items that vary on the dimension of interest and then try to control other factors as well as possible.

Word frequency or familiarity provides a case in point. This variable has been of great interest in comparisons of recall and recognition, because low frequency words are better recognized than high frequency words, but high frequency words are better recalled (e.g., Balota and Neely, 1980; Gregg, 1976). At least, this is the commonly provided explanation, but Hintzman and Hartry (1990) point out that the story is not so simple: words of medium frequency are better recognized than those of high frequency, but rare words are poorly recognized relative to medium frequency words. The effects of word frequency are likely to be similarly complex on other tasks and this consideration may go some way to explaining the puzzling effects of word frequency reported thus far. For example, MacLeod (1989) found that low frequency words produced more priming than did high frequency words on word fragment completion, but Tenpenny and Shoben (1991) reported the converse. Roediger et al. (1992) found that low frequency words produced more priming on this task than did high frequency words (with length and concreteness held constant), but there was less effect of this variable on priming in word stem completion with the same target words. The size of the effects obtained by Roediger et al. (1992) was modest, even when they were significant. Finally, Hintzman and Hartry (1990) found no correlation (actually, +0.08) between word frequency and priming over the 192 words originally used by Tulving et al. (1982). In short, the effects of word frequency on primed word stem and word fragment completion are inconsistent.

The effects seem clearer on other tasks, but perhaps only because less research has been devoted to them. In word identification, Jacoby and Dallas (1981) found that low frequency words produced more priming than high frequency words; this outcome was replicated by Jacoby and Hayman (1987) and Kirsner et al. (1986). In the lexical decision task, Scarborough, Cortese and Scarborough (1977)

---

6 An additional complication in studies of word frequency, especially in tasks such as word stem completion, is that baseline differences are likely to exist for the two classes of items (see Roediger et al., 1992, Experiment 3 for elaboration).
found greater priming from low than from high frequency words. Kirsner et al. (1986) reported that word frequency affected priming on the picture naming test; pictures whose names appeared with low frequency produced more priming than those whose names were of higher frequency. Aside from this experiment, there has been little research on the effects of word frequency on nonverbal or conceptual implicit tests.

D.L. Nelson and his colleagues have conducted extensive research examining the effects of lexical set size and meaning set size on priming in fragment completion tasks (e.g., Nelson, Schreiber and McEvoy, 1992b). Lexical set size refers to the number of words subjects generate when given the fragment cue; meaning set size refers to the number of meaningfully associated words subjects produce to the target. (Both are estimated from production norms obtained from large groups of subjects.) Lexical set size has been shown to have large effects on implicit fragment completion tests, but meaning set size has no effect (Nelson, Canas, Bajo et al., 1987). On explicit tests with the same cues, both lexical set size and meaning set size affected recall. Nelson et al. (1987) concluded that implicit fragment completion tests depend on a lexical search through relevant sets of items sharing the same letters, but that explicit tests using these same cues rely on both lexical and conceptual searches. These conclusions were confirmed in later experiments reported by Nelson et al. (1989) and are broadly consistent with those of Weldon (1991) and Tulving and Schacter (1990) in emphasizing the role of lexical processes in verbal priming on perceptual tests.

On the other hand, results with word stem completion are more complicated than those for word fragment completion reviewed in the previous paragraph. With both beginning and ending stems (e.g., cl____ or __ud for cloud) lexical set size affects target recovery under both explicit and implicit test instructions, and the same is true of meaning set size (Nelson, Schreiber and Holley, 1992a). Meaning set size affects word stem completion even on implicit tests, unless the prime is strongly associated to the target. The contrasting effects of meaning set size on word fragment and word stem cues is puzzling in light of the fact that most evidence indicates that the two measures are similar across a variety of other manipulations (Rajaram and Roediger, 1993).

Hunt and Toth (1990) examined the variable of orthographic distinctiveness on priming in word fragment completion. Orthographically distinctive words are funny looking (sphinx, typhoon) because they violate usual expectations of English orthography. Relative to words of common orthography that are matched on several variables, words with distinctive orthographies produced more priming on a word fragment completion test than did those with more typical orthographies. However, the same effect did not occur in perceptual identification (although the mean difference, albeit nonsignificant, was 8% in one experiment). However, the Hunt and Toth (1990) research is based on samples of only 16 words, so further research would be welcomed to document this effect, with a wider sample of materials.

To return to the theme at the outset, lexical variables are correlational in nature and therefore attribution of an effect to the purported variable of interest requires more faith than usual. For example, it may be that orthographic distinctiveness underlies the finding that low frequency words often produce more priming than do high frequency words, under the reasonable (though unchecked) assumption that high frequency words are more regular orthographically. No one has yet equated orthographic regularity and manipulated word frequency or familiarity in a priming study. Nonetheless, within the limits of the research at hand, some lexical variables seem to affect performance on verbal perceptual tests, consonant with the idea that priming with such tests relies on lexical processes (Kirsner et al., 1989; Weldon, 1991), a lexical search (Nelson et al., 1989) or a visual word form memory system (Schacter, 1990; Tulving and Schacter, 1990).

Indirect priming
As described above, indirect priming refers to the effects of words semantically or associatively related
to the target on priming of the target. The boundaries between direct and indirect priming are fuzzy, and the work on priming between pictures and words and between words in bilinguals’ two languages could have been considered under this heading. Here we use the more traditional relations, such as associates, synonyms and the like. In studies of ‘semantic priming’ in which there is very brief separation between the prime and target, indirect priming occurs from many types of relation (see Neely, 1991). However, Shelton and Martin (1992) provide evidence that the term semantic priming is a misnomer because the priming is actually associative in nature. Words closely related in meaning but unassociated (according to norms) produce no priming even over short delays. The interest in this section is in indirect primes (whether bearing an associative or semantic relation) produce facilitation in the paradigms more typical of implicit memory research, with longer delays between study and test. Thus far, few experiments have addressed this issue.

Mandler, Graf and Kraft (1986) examined the effects of two types of indirect primes on word stem completion. Priming of phonological information was achieved by having subjects study words rhyming with the target, whereas semantic priming involved having subjects study category coordinates. Phonological priming occurred on the stem completion test at a ten minute retention interval, but semantic priming did not occur even at this delay, which is quite short relative to typical delays in implicit memory tests. In later experiments, Overton and Mandler (1987) and Mandler, Hamson and Dorfman (1990) obtained some evidence for semantic priming, but it was much weaker than direct or phonological priming and did not appear on some measures.

Roediger and Challis (1992) also failed to find indirect priming in three experiments in which delayed word fragment completion was the dependent measure. In Experiment 1 subjects studied synonyms (*tadpole*), associates (*pond*), category coordinates (*frog*), or visually similar words (*polygon*) before later completing a fragment for the target (po— for *polliwog*). None of these relations produced priming, under conditions in which there were large effects from direct priming. In addition, adding indirect primes to direct primes (putting both *polliwog* and *frog* in the list) did not enhance priming over the presentation of the target item itself. However, this manipulation had large effects on free recall for the target.

McDermott and Roediger (1992a) have replicated the cited findings by Roediger and Challis (1992) in word fragment completion and have extended them to an implicit category association test. That is, neither presentation of indirect primes nor adding indirect primes (*frog*) to direct primes (*polliwog*) affects priming when subjects generate words to category names, although adding indirect primes to direct primes did increase free recall of the targets. This finding is inconsistent with the transfer appropriate processing theory described above, because it predicts that indirect priming should occur on the category association test (Roediger, 1990a). The word fragment completion results are consistent with this viewpoint, but McDermott and Roediger’s category association results clearly are not. Further experiments are now being conducted with other semantic relations to examine indirect priming on conceptual tests.

**Priming of nonwords and nonobjects**

By one view of implicit memory, priming results from activation of preexisting structures (logogens for words, pictogens for objects or pictures) during study that are then reactivated with more facility later (e.g., Mandler, 1980; Morton, 1969; Warren and Morton, 1982). This leads to the prediction that priming should only occur for familiar words and objects and not for nonwords and nonobjects, because the latter have no pre-existing representation. Some researchers have indeed reported that priming for nonwords is at least much weaker than for words and may be absent altogether. Memory-impaired patients represent an interesting group to study because (unlike normals) priming of unfamiliar material is less likely to be due to explicit recollection. Indeed, Diamond and Rozin (1984)
originally reported no priming from nonwords in a stem completion test and Cermak, Talbot, Chandler et al. (1985) and Gabrieli, Cohen, Huff et al. (1984) produced similar findings using word identification and lexical decision measures, respectively.

These early results provided support for activation theories of priming. However, it is now clear that priming can be obtained for nonwords in both normal subjects and memory-impaired patients, and on a variety of tests. (Reasons for discrepancies across experiments are still a bit unclear). For example, Haist, Musen and Squire (1991) found priming for both words and nonwords on a word identification test; additionally, they found equivalent levels of priming for amnesic patients and normal controls. Other supportive evidence for the existence of nonword priming comes from Smith and Oscar-Berman (1990) using the lexical decision task in patient groups, although priming from nonwords was relatively weak. Numerous studies have found reliable nonword priming in normals in various perceptual priming tasks (e.g., Cermak et al., 1985; Feustal, Shiffrin and Salasoo, 1983; Jacoby, 1983a; Rueckl, 1990; Whitlow and Cebollero, 1989). These repeated findings provide difficulties for activation theories of priming, in that no permanent representations exist to be primed for nonwords.

Priming also occurs for unfamiliar visual objects in several paradigms (Gabrieli, Milberg, Keane et al., 1990; Kersteen-Tucker, 1991; Musen, 1991; Musen and Squire, 1991; Musen and Treisman, 1990; Schacter et al., 1990). Schacter et al. (1990) reported priming from objects that could be constructed in the real world (even if they were unfamiliar), but in numerous experiments he and his colleagues have found no priming from 'impossible' objects, Escher-like impossible figures that could not be realized in three dimensional space (see Schacter, Cooper, Delaney et al., 1991). The task that Schacter and his colleagues use is a decision task in which subjects decide if an object is possible or not. The failure of impossible objects to produce priming cannot be due to any failure of encoding because in recognition experiments under the same conditions performance is well above chance. It will be interesting to see if impossible objects produce no priming on other tasks that might be devised.

In short, with a few exceptions, priming on perceptual implicit memory tests does seem to occur both from nonwords and from unfamiliar objects. No one has yet examined the effect of nonwords on conceptual implicit tests, although doubtless relevant tasks could be constructed.

Summary of the section

Although the data are not perfectly consistent, the following conclusions seem warranted. Priming on perceptual implicit memory tests is affected by symbolic modality such that pictures produce more priming than do words on pictorial tests and words produce more priming than do pictures on verbal tests. For bilinguals, greater priming occurs when the study and test word appear in the same language than in different languages, although this effect may disappear when more meaning is imbued in processing the items at study (Smith, 1991). So far, indirect priming from semantically or associatively related materials (even synonyms) has not been found on perceptual implicit tests. Low frequency words produce more priming in most studies than do high frequency words, and orthographically distinctive words produce more perceptual priming, too. Finally, nonwords and unfamiliar objects can produce priming on perceptual tests. The evidence on the effects of these same variables on conceptual implicit tests is scant, pointing to some lacunae in the literature.

The bulk of the evidence from manipulation of these material variables on perceptual priming tests supports the idea that these priming effects are driven by processes used in word or object recognition. In general, the procedures used in recognizing words and objects in the study phase transfer to identifying their rapidly presented or fragmented forms in the test phase. The best guess is that these effects occur at intermediate stages in neural processing, in a lexical memory or word form system for words or a geon assembly layer or structural description system for objects (see Biederman, 1987; Schac-

**Manipulation of surface features**

In this Section we review experiments involving direct priming, but where there is an alteration in surface features between study and test presentations. In most cases, the study phase involved presentation of words and the test phase consisted of visually presented fragments or clues for the words (and other, nonstudied, words). Assuming that test stimuli were held constant, alterations in surface form may involve (a) changes of modality (auditory or visual presentation at study), (b) variation in the typography (font or case) of visually presented words, (c) alterations in physical characteristics (e.g., size) of objects, (d) variation in physical characteristics of auditorily presented stimuli, and (e) presentation in intact or in perceptually degraded forms. We consider these variables in turn. To presage this review, often manipulations of surface variables have large effects on perceptual (or data-driven) implicit memory tests and have no effect on explicit tests or on conceptual implicit tests. However, this pattern is not invariably observed and at least one variable shows the opposite outcome from that just described.

**Modality**

Probably the most studied variable in this set is modality. In general, auditory presentation produces about half as much priming as does visual presentation on the tests of word fragment completion (Blaxton, 1989; Challis and Sidhu, 1993; Craik, Moscovitch and McDowd, 1993; Donnelly, 1988; Roediger and Blaxton, 1987; Srinivas and Roediger, 1990; Weldon, 1991), word stem completion (Bassili et al., 1989; Craik et al., 1993; Graf, Shimamura and Squire, 1985; McClelland and Pring, 1991), partial word identification (Hashtrudi, Ferguson, Rappold et al., 1988), and lexical decision (Scarborough, Gerard and Cortese, 1979). The situation with respect to word identification appears much

![Diagram](image_url)

**Fig. 1.** Results from Rajaram and Roediger (1993). Direct comparison of priming in four implicit tests as a function of study condition. Priming scores represent relative priming (i.e., [completion rates of studied minus nonstudied item sets] / [1.0 – the nonstudied completion rate]) to correct for differences in baserate (nonstudied) differences across tests (see Snodgrass, 1989). Much the same picture appears in the raw priming scores.
the same. Jacoby and Dallas (1981) found virtually no priming from auditory presentations on this measure, but Kelley, Jacoby and Hollingshead (1989) found reliable priming from auditorily presented words on visual word identification, although it was significantly less than that from visually presented words. Additionally, experiments directly comparing word identification with the other measures have shown auditory priming significantly above baseline, but well below that arising from visual presentations (Rajaram and Roediger, 1993; Weldon, 1991). Schacter and Graf (1989) also found a modality effect in priming of new associations.

Two reviews of the literature on modality effects on data-driven tests exist (Donnelly, 1988; Kirnser et al., 1989), so we can keep this section short and refer readers to these reviews. We will discuss further only two recent results bearing on these issues. First, Rajaram and Roediger (1993) directly compared four implicit memory tests (and one explicit test) as a function of modality. Subjects either read words, heard words, or saw pictures of the same concepts (with high name agreement). Then they took one of four implicit tests: word identification; word fragment completion; word stem completion; or anagram solution. The basic results are shown in Fig. 1.

As can be seen, greatest priming occurred in each test from visual presentation of words, less from auditory presentation and least from pictures. The modality difference seems to be roughly equivalent for the four tasks, except for the anagram solution task in which it is a bit higher than the others. The conclusion about the modality effect across tasks from this experiment is more secure than the conclusions above from cross-experiment comparisons; comfortably, the conclusions agree.

All studies described thus far have employed visual tests. Several researchers have used auditory stem completion tests and varied modality of presentation. Bassili et al. (1989) showed that auditory presentation of words produced more priming than did visual presentation on this test. Jackson and Morton (1984) reported similar modality effects in identifying spoken words. Cross-modal priming (visual to auditory) occurred in these experiments, and McClelland and Pring (1991) reported evidence indicating that there was a phonological basis to this priming.

Turning to conceptual implicit memory tests, the data are sparse but the picture so far is clear: modality has no effect on priming in answering general knowledge questions (Blaxton, 1989; Challis and Sidhu, 1993) or in producing instances to category names (Srinivas and Roediger, 1990). Of course, there is as yet little systematic work on these tests, and modality has not been included as a variable in most studies; however, the preliminary evidence is consistent.

As this review indicates, the implicit memory tests used thus far overwhelmingly have been visual tests. A few experiments have been done on auditory priming, but to our knowledge none exist asking if priming occurs in (say) identification in smell, taste or touch. Technically, nothing would prevent such experiments, but they lie in the future. Extrapolating from current knowledge, we may expect sharp specificity effects in these other domains. Prior reading or listening to the word popcorn, or viewing a relevant picture, would likely cause modest priming on a ‘whiff identification test’ wherein subjects are given a whiff of the relevant smell and asked to name it, compared to the case in which the prior study phase involved a deep smell of the odor. Of course, smell and taste might show great transfer, because these senses seem to be linked. These speculations await future research, but clearly manipulations of modality in priming studies could be developed much further.

**Manipulations of typography**

Unlike manipulations of modality, where the literature is consistent, findings with regard to alterations in typography are quite messy. The typical experiment involves manipulation of some typographic feature at study and then testing in the same or a different format. Many experiments report small but significant effects of typography on perceptual implicit memory tests, but there are clear exceptions. To consider positive cases first, Kolers conducted
Ch. 2  H.L. Roediger and K.B. McDermott

There has been considerable research showing that radical manipulations of text (presenting it upside down or in various other rotations) conveys generally orderly transfer to rereading the same text later (see Kolers, 1979; Kolers and Roediger, 1984, for reviews). He interpreted these findings as showing that perceptual pattern analyzing operations used in reading transferred to later rereading. However, others have argued that there may be a conceptual basis to the transfer, too, and that intentional recollection may be involved (see Horton, 1985; Masson and Sala, 1978; Tardif and Craik, 1989).

Others have attempted to find effects of typographic specificity in the more standard priming tests of implicit memory. Jacoby and Hayman (1987) reported several experiments using primed word identification in which matching typographies at study and test enhanced priming relative to the case in which they were mismatched. Roediger and Blaxton (1987a) reported small effects from having subjects study words typed (in lower case letters) or handprinted (in upper case letters) and then tested on word fragments in the same or different typography. They generally found greater priming in the matching than the mismatching cases, but the greatest effect of matching typographies occurred in the handprinted conditions. They obtained greater effects in another experiment in the series, in which the word presented during study and the fragment given at test were either presented clearly or blurred so as to be just readable. Now greater priming occurred in the matching (blurred to blurred; clear to clear) conditions than in the mismatching (blurred to clear or vice versa) conditions, and the effect was symmetric. Blaxton (1989) reported effects of matching and mismatching typography in primed word fragment completion, but found no such effect on answering general knowledge questions (a conceptual test). Madigan, McDowd and Murphy (1991) also found effects of typography on primed fragment completion.

On the other hand, there have been many failures to obtain such effects under rather similar conditions. Clarke and Morton (1983) had subjects study words handwritten or typed and then tested them on word identification with typed words. There was equivalent transfer in the two cases. Rajaram and Roediger (1993) had subjects read words in one of two fonts before testing in one or the other font. The typography manipulation had no effect on any of four implicit memory tests. (Their results, in the black columns of Fig. 1 here, are collapsed across this variable). Similarly, Scarborough et al. (1977) found as much priming in the lexical decision task when the case of the words (upper or lower) mismatched as when it matched. Other similar results with various tasks can be found in Carr, Brown and Charalambous (1989), Feustal et al. (1983), Graf and Ryan (1990), Moscovitch and Umilta (1991), and Tardif and Craik (1989).

Why these discrepancies? There is no easy answer at the moment. One possibility is that priming on these verbal, visual implicit memory tests relies heavily on lexical processes (Kirsner et al., 1989; Weldon, 1991) or processing in a presemantic visual word form system (Tulving and Schacter, 1990). So long as the word is seen, priming will result; even radical changes in the surface features of the word create modest differences in priming (relative to changes of modality or symbolic form). Graf and Ryan (1990) suggested a second possibility. They argued that in order to obtain typographic specificity, one must employ an orienting task during the study phase that draws subjects' attention to perceptual aspects of the display. They report results consistent, in general, with this idea, but others have obtained specificity effects when using orienting tasks that did not overtly call attention to the perceptual character of the display. A third answer is provided by Marsolek, Kosslyn and Squire (1992), who investigated the contributions of left and right hemispheres to form-specific priming in word stem completion. Subjects studied words presented visually or auditorily and received a visual word fragment completion test in which words were presented to the left hemisphere (and thus went first to the right hemisphere) or the right hemisphere (and went first to the left hemisphere). They found that within-form priming was greater when the test stimulus was presented first to the right hemisphere.
than when presented first to the left hemisphere. They also found that manipulations of case (upper or lower) affected priming in the right but not left hemisphere. Thus form-specific priming appeared only when the right hemisphere was engaged in the task. Although important, these findings cannot be used in a principled way to clarify the successes and failures to find typographic specificity effects in the literature reviewed above, because the study and test conditions are dichoptic in those cases. Appeal to hemisphere differences in these cases cannot be justified by independent evidence (even though such differences conceivably could have played a role).

In sum, manipulations of typography on perceptual implicit memory tests are not well understood as yet. Often there is a small effect of typographic manipulations, but frequently not. A safe conclusion is that alterations in typography have much less effect than do changes in modality or symbolic form (pictures to words).

**Physical surface variations in object priming**

The previous two Sections have considered priming of written words, but other priming studies have been done with pictures, as we have already discussed above. As with words, we can also ask what variations in physical characteristics of objects affect priming. Indeed, in many ways the possibilities here are even more rich and varied. The literature on object priming (more accurately, priming of pictures of objects) has a long history and is intertwined with research in object recognition. What cues do we use to recognize objects and how do we recognize them from unusual viewpoints? Such issues have been of interest to psychologists studying perception for years, and the topic is actively researched, but in a different tradition from that of implicit memory (e.g., Biederman, 1987; Leeper, 1935; Rock, 1975). But if priming on many implicit tests is perceptual in nature, as we have been arguing, then this priming research in other domains is directly relevant to the topic of implicit memory (Roediger and Srinivas, 1993). Various tasks have been used to address these issues; researchers primarily have employed picture naming, object decision, and naming of picture fragments. We consider the effects of several variables on these tasks.

An interesting issue is whether priming transfers across different pictures of the same concept (study one picture of a bird, for example, and test on a different example). Warren and Morton (1982) tested subjects on a picture identification task (viewing brief exposures in a tachistoscope) when they had previously studied either an intact version of the same picture, a different picture of the same concept, or the name of the concept. Greatest transfer occurred in the identical condition, less transfer resulted from seeing the different exemplar, and no transfer occurred from reading the name. From a post hoc analysis, Warren and Morton (1982) concluded that perceptual similarity was not causing the priming effect for the different exemplars; they did not, however, directly manipulate perceptual similarity. In other studies using different paradigms, perceptual similarity does seem to play a determining role (e.g., Ellis et al., 1987; Goldinger, 1992).

Specificity in pictorial priming has also been reported when the study and test patterns differ in orientation along the picture plane. Jolicoeur (1985) showed that facilitation in naming line drawings of objects that were rotated from 0 to 120 degrees from the upright was quite specific to the particular objects seen. That is, subjects did not speed up because of some general mechanism for righting misoriented objects, but because of prior experience with particular objects. Specificity effects have also been discovered across objects rotated in depth so that they are seen from different viewpoints. Bartram (1974) reported experiments in which subjects named pictures across blocks of trials. Either the same picture was repeated, a different viewpoint of the same object was repeated, or a different exemplar of the object was repeated. Greatest transfer or priming occurred in naming the same pictures, less occurred from naming pictures of the same object seen in different views, and no priming occurred from naming different exemplars of the same object (unlike the results of Warren and Morton, 1982, cited above, who did find priming from different ex-
emplars). Srinivas (1993) has also reported greater priming when the same viewing angle of objects was used at study and test. Tarr and Pinker (1989) presented results from several experiments in which subjects repeatedly named pictures or classified them as normal or mirror reversed as they appeared in several orientations. The experiments were different from typical priming studies, but led to the conclusion that subjects represent shape-specific information about objects, in agreement with the other results here. This type of evidence fits well with theories of object recognition that maintain that objects are recognized through consultation of specific stored exemplars and their rotation (see Tarr, 1989). However, not all evidence indicates that object priming in different orientations produces specific effects; Biederman (1987) reports in passing several experiments in which he did not obtain such effects, but the details have never been published.

The bulk of the evidence reviewed above about priming across different exemplars and different viewpoints implicates, again, specificity of perceptual priming. There are two interesting exceptions to this claim, however. The dimensions of reflection (i.e., left-right orientation or mirror reversal) and size (small or large) have generally failed to affect priming. Biederman and E. Cooper (1992) found that priming in picture naming was not affected by left-right orientation; this variable also had little effect on priming on an object decision task, in which subjects decided if objects could possibly exist in 3-dimensional space (L. Cooper, Schacter, Ballesteros et al., 1992). Although this variable may warrant further examination (say, in fragment naming paradigms), the evidence thus far indicates that priming generalizes across left-right orientation.

The effects of size are most interesting. L. Cooper et al. (1992) showed that varying size of objects between study and test did not affect the amount of priming in the object decision task, although this variable did affect recognition memory in the expected way (matching the size of pictures at study and test resulted in better recognition than mismatching the sizes). Similarly, Biederman and Cooper (1992) showed that if subjects named small or large objects at study and test, priming in renaming them later was just as great when the size of the objects mismatched (small-large or the reverse) as when size matched (small-small or large-large). However, this variable had large effects on explicit recognition decisions. Kolers, Duchnicky and Sundstroem (1985) showed that size mattered in recognition memory for faces, but not for verbal materials, although they did not include implicit memory measures in their experiments.

The effects of size and orientation are interesting because the manipulation of a seemingly 'perceptual' variable has no effect on priming on a putative perceptual implicit memory test but has a large effect on explicit recognition. Schacter (1992) and Biederman and Cooper (1992) appeal to different neural mechanisms underlying the two forms of memory (object priming and recognition), one sensitive to size and one not. At some level, this suggestion must be correct; surely neural mechanisms underlie performance on these tasks. But their proposals go beyond this statement to draw on other findings in neuroscience that implicate specific neural mechanisms. However, the data still pose a problem for psychological theories of recognition that ascribe a large role to perceptual familiarity or perceptual fluency in making recognition judgments (Jacoby and Dallas, 1981; Mandler, 1980), because the experiments show that a perceptual variable can affect recognition without affecting perceptual priming.

To summarize this Section, several variables (manipulating pictorial examples, or their visual orientation in the picture plane or in depth) reveal effects on perceptual priming, but size and reflection are exceptions. It will be interesting to see if these conclusions generalize to other measures of priming, such as fragment naming tests. At the moment, the data present a challenge to the transfer appropriate processing viewpoint (but see Schacter, 1992, for a resolution) and to dual process theories of recognition.
Voice changes in auditory priming

Several studies have examined effects of changing surface variables — usually voice of the speaker — on auditory priming. Jackson and Morton (1984) had subjects listen to a list of words spoken by either a male or female; subjects later attempted to identify words through noise that were presented in the female voice. They found no greater priming from the prior experience of hearing the female voice than the male voice and concluded that priming in audition was indifferent to such surface changes. Schacter and Church (1992) replicated this finding in their first two experiments, in which subjects heard and then were tested on spoken word identification. However, in two other experiments they did report voice specificity in an auditory version of the word stem completion test. The difference between the two sets of experiments seems to lie in whether subjects had to identify spoken words in white noise (a procedure showing no perceptual specificity) or had to complete stems during the test (where perceptual specificity was obtained). Schacter and Church (1992) argued that this result may arise in a similar manner as the typographic effects of Marsalek et al. (1992) reviewed above. There is evidence that the right hemisphere processes prosodic information about a speaker’s voice and therefore might be responsible for voice specificity effects. Embedding the words in white noise may eliminate the contribution of the right hemisphere and hence the effect of voice specificity. Although speculative, Schacter and Church’s ideas have the advantage of tying together confusing literatures on typographical manipulations in visual priming with voice manipulations in auditory priming.

Goldinger (1992) has reported a finding different from one of those reported by Schacter and Church (1992). He had subjects listen to words presented by different speakers in noise both during a first (study) and a second (transfer) phase. He reported specificity of priming in that same voice presentations at study and test produced greater priming than did different voice presentations, unlike the results of Schacter and Church (1992). Goldinger (1992) noted several differences between the experiments, but the most likely is that noise was included during both phases of his research and during only the test phase by Schacter and Church (1992).

In general, the data from auditory priming experiments agree with findings from visual studies in that priming on perceptual implicit memory tests is enhanced by matching perceptual operations between study and test.

Repetition priming of degraded stimuli

In the foregoing review, we have considered direct priming on a variety of perceptual tests involving naming words or pictures from brief displays or from their fragmented forms, or both. Subjects studied a word or picture and were tested on its fragmented form, for example. Changes made in the surface form between study and test often affect amount of priming; the greatest priming occurs when perceptual features are held constant, presumably because perceptual operations used in recognizing words and pictures are recapitulated in this case. Carrying this transfer appropriate processing logic one step further, there should exist situations in which priming is increased even over that in which the study item matches the test item on features such as modality and typography. If the study phase involves subjects decoding a fragment, then one might expect priming on a fragment completion test to be enhanced above the level observed in the case of presenting the intact word or picture. In general, according to the transfer appropriate processing approach, the more the operations at test match those during study, the greater priming should be.

There are surprisingly few experiments employing this strategy, but most produce strikingly positive results. Gardiner and his associates have provided interesting demonstrations of how generating words from fragmented clues during study can produce positive effects on transfer, but only when the test fragment is identical to the fragment given at study (Gardiner, 1988, 1989; Gardiner, Dawson and Sutton, 1989). For example, Gardiner et al. (1989) had subjects generate words
from conceptual and fragment clues, such as single unmarried man — B—_— E—_—OR. In another condition, subjects read the intact version of BACHELOR at the end of the phrase. Relative to a nonstudied condition, all conditions produced priming for the target words; however, priming was 10–15% greater across experiments if subjects had generated the word from the same fragment that they later received at test. This advantage occurred relative to the condition in which subjects read the intact word, but also relative to the case in which subjects generated the word from a fragment that differed by one letter from the target fragment. That is, in order to obtain the enhanced positive transfer (relative to the Read condition), subjects had to receive exactly the same fragment at study and test. Such specificity of priming is often obtained, as we have seen, on perceptual implicit memory tests; in experiments that are in some ways similar to Gardiner’s, Hayman and Tulving (1986b) found evidence for ‘hyperspecificity’ of priming. The fact that Gardiner’s results were so specific to the match between the study and test cue indicates that the effect is unlikely to arise from explicit or controlled uses of memory, as would occur in more typical generation effects in explicit tests. Also, Gardiner (1989) found the effect under conditions in which subjects claimed (in a later questionnaire) to have been unaware that any test items had appeared in the study list.

Srinivas (1993) has obtained a result similar to Gardiner’s (1988, 1989) using pictorial materials. At study subjects saw intact pictures with their names, or fragmented forms of pictures with their names. At test subjects saw fragments that represented the pictures, and these fragments could be either the same as at study or different. Relative to a nonstudied condition, Srinivas obtained great amounts of priming (25% or greater) in all conditions. However, compared to the condition in which subjects saw intact pictures, seeing the same fragment at study and at test about doubled priming, from 25% to 50%. Yet studying one form of fragment and being tested on a different form produced priming equivalent to that from the intact picture. (In another experiment she showed that very little priming would accrue on picture fragment naming from prior presentation of only the word.) Srinivas’s results with pictures are similar to Gardiner’s with words and both point to the specificity of perceptual priming.

In other experiments using perceptually degraded pictures, Jacoby, Baker and Brooks (1989) produced an interesting demonstration of a dissociation between recall and priming. At study Jacoby et al.’s subjects either saw intact pictures or saw degraded pictures that gradually became clarified on the screen until they could be identified. Later they received either an explicit test (free recall) or an implicit test (subjects clarified pictures to identify them as in the study phase). The pictures studied in an intact form were recalled better than the degraded pictures, but when priming was measured through savings in the clarification procedure, pictures that had previously been clarified produced more priming than those that had been presented intact. Like the results of Jacoby (1983b), the finding of opposite effects on an explicit and an implicit test provides good evidence that the two types of test are indeed measuring different forms of mnemonic access. Jacoby et al. argued that the facilitation of picture clarification from prior clarification, relative to viewing the intact picture, represented a form of transfer appropriate processing and the basic finding is analogous to those of Gardiner (1988) and Srinivas (1992).

Snodgrass and Feenan (1990) showed a slightly different effect of degraded presentations that also produced more priming than intact presentations. Their materials involved seven versions of degraded pictures that increased in the amount of information given in levels from 1 (very degraded) to 7 (easy to identify). During an initial phase subjects were exposed to fragments of level 1, 4, or 7 and were given the names of pictures during study to aid their identification. During the test phase subjects tried to identify the target from fragments beginning with the level 1 fragments. Snodgrass and Feenan (1990) found that presentation of level 4 fragments produced greater priming than level 1 or level 7, both in terms of overall savings and in priming on the first,
most degraded, fragment subjects were given in the test (level 1). The finding that degraded fragments (level 4) produced greater priming than almost intact pictures (level 7) is consistent with prior research. However, they argued that the finding of greater priming from level 4 fragments than from level 1 fragments on a test involving the level 1 fragments was inconsistent with transfer appropriate processing ideas. This finding clearly deserves more study, because it is also inconsistent with other findings (e.g., Gardiner, 1988; Srinivas, 1993) showing that the studied forms of fragments produce the greatest priming on word fragment and picture fragment naming tests. However, Snodgrass and Feen’s (1990) level 1 fragments were much more degraded than those used in the other experiments.

Biederman and Cooper (1991) have reported one experiment in which different fragmented pictures primed each other as well as they primed themselves. The critical set of stimuli involved pictures in which alternate lines and vertices were deleted; with this manipulation, the objects could still be readily identified when presented leisurely (i.e., the object parts were preserved so that identification could occur). Test items were presented briefly and subjects tried to name them. As mentioned, on this special stimulus set, priming was as great from the same fragment as from the complementary fragment. However, with a different type of degradation (where stimuli were mutilated such that object parts could no longer be so easily recovered), Biederman and Cooper (1991) found the usual specificity of priming with these stimuli: greater priming occurred when the same fragment was used at study and test than when different fragments (of the same picture) occurred. Probably most fragments in other studies resemble those in Biederman and Cooper’s (1991) experiment that did not permit recovery of object parts.

Stimuli may be degraded by means other than fragmentation. Nairne (1988) had subjects study intact word or words that were presented very briefly and then masked. Of course, shortening presentation duration might be expected to harm memory — with less opportunity for processing, less retention might be expected (on explicit tests). However, using the same logic of Jacoby et al. (1989) in their clarification study, Nairne (1988) reasoned that a brief occluded glimpse of a stimulus might produce greater perceptual processing as the subject strives to accurately see the stimulus, relative to the case in which the word is presented intact. Nairne (1988) did not test perceptual priming, but rather examined recall and recognition, assuming that recognition (but not recall) might be affected by a perceptual component (Jacoby and Dallas, 1981; Mandler, 1980). Indeed, he obtained exactly this pattern: if items presented very fast could be identified during study, they were later recognized better than items presented intact. Nairne (1988) argued against the obvious item selection criticism (i.e., that subjects only identify items easy to remember at study) in several ways, one of which was the fact that the variable of presentation time produced no effect on free recall. However, in a follow-up experiment Hirshman and Mulligan (1991) found that briefly presented items were both recognized and recalled better than items presented intact (at least under conditions in which item selection was not a problem). More troubling for Nairne’s (1988) interpretation of his finding, Hirshman and Mulligan (1991) found no evidence that the two presentation modes produced an effect on priming in a word identification test. The effect of the presentation time variable over very brief intervals seems even more puzzling now, because it has been shown to affect two explicit tests but not an implicit test. The finding represents a severe violation of the total time law in explicit memory: masked presentation of a stimulus for 100 msec led to greater recall and recognition than presentation of a stimulus for 10 times as long (1 sec). We consider effects of presentation time per se in the next Section.

In general, the bulk of the evidence from presentation of degraded stimuli at study and/or test indicates greater priming from a degraded presentation of a stimulus (so long as it can eventually be identified) on a perceptual priming test. The exceptions come from part of Biederman and Cooper’s
reported that (1- versus 2- second) rate of presentation of words had no effect on primed word identification, but did affect recognition under the same conditions. Using a word fragment completion task, Neill, Beck, Bottalico et al. (1990) varied presentation time over the range of 1, 3, and 6.5 sec and found no effect on amount of priming. Again, the variable had a large effect on recognition. As we shall see below, highly similar studies of massed repetition lead to the same conclusion. To our knowledge, no experiments have addressed the issue of study time, per se, on conceptual implicit tests.

**Massed repetition**

In experiments manipulating study time, a stimulus is exposed for various lengths of time; in experiments investigating massed repetition, the stimulus is repeated with no extraneous intervening events. In both cases the stimulus is intended to engage the subject’s attention in a continuous manner, but in the case of massed repetition the item is continually re-presented, either by the experimenter or subject (who can be asked to rehearse it). The procedures for investigating study time and massed repetition are quite similar and, unsurprisingly, so are the ensuing results. Challis and Sidhu (1993) have recently reviewed the effects of massed repetition so we will simply state the conclusions of their analysis without reviewing the literature at length.

First, in every experiment in which effects of massed repetition of individual items have been examined on perceptual implicit tests, there has been little or no effect on priming. Jacoby and Dallas (1981), Perruchet (1989), and Greene (1990) found equivalent priming from one and two presentations of words on word identification procedures. Similarly, using word fragment completion tests, Roediger and Challis (1992) and Challis and Brodbeck (1992) found no difference in amount of priming from one and two massed presentations. In most of these cases, the number of massed presentations did significantly affect either recall or recognition in companion conditions.

More impressively, Greene (1986) had subjects repeat words aloud for either 2 sec or 10 sec before
taking either an implicit word stem completion or explicit word stem cued recall test. He found no differences on the implicit test, but a sizeable effect on the explicit test, thereby satisfying Schacter et al.'s (1989) retrieval intentionality criterion. Challis and Sidhu (1993) accomplished this feat with word fragment completion. They presented words 1, 4, or 16 times and tested subjects on explicit and implicit tests with word fragments as cues. Number of massed repetitions had no effect on the implicit test, but a reliable effect on the explicit test (as it did on free recall and recognition).

In the paradigm in which subjects must decide if an object is possible or impossible, Schacter et al. (1991) gave subjects one exposure or four exposures to the objects (each lasting 5 seconds) before the test. This manipulation produced equivalent priming for possible objects. In another experiment subjects received one or five 1-second exposures; in this case priming only occurred after the massed repetitions. However, as Schacter et al. point out, encoding these relatively complex objects may take several seconds; therefore the advantage of massed to single presentation in this case may simply be due to subjects being unable to accurately encode the objects in 1 second. The findings with regard to study time and massed repetition are relatively consistent: these variables have little or no effect on perceptual implicit tests (so long as the single presentation allows time for accurate encoding), but almost always have reliable effects on explicit tests.

Challis and Sidhu (1993) also examined effects of massed repetition on a conceptual implicit test (answering general knowledge questions) and found a reliable effect. Therefore, massed repetition dissociates perceptual and conceptual implicit tests just as do other variables (Blaxton, 1989). Further research will be needed to test the generalizability of this conclusion to other conceptual implicit tests.

Distributed repetition and spacing effects
Although massed repetition clearly seems to have little or no effect on perceptual implicit memory tests, the situation is murkier with regard to distributed practice and spacing effects. (We follow here the convention of referring to a distributed practice or spacing effect when two presentations occurring with other intervening items lead to better performance than two presentations that are massed. A lag effect refers to greater effects of spaced repetitions with increasing numbers of items between presentations.) Jacoby and Dallas (1981) found a small but not significant spacing effect on word identification performance in one experiment, but did find a significant effect in another experiment. Perruchet (1989) conducted four experiments on this issue using a perceptual clarification procedure, but the results were equivocal; he concluded that spacing has a 'slight and fluctuating effect' on priming (p. 113). Whitlow (1990), using word identification procedures, found a repetition effect for words, but only when subjects expected a recall test. Whitlow and Cebollero (1989) found a repetition effect for pseudowords and not words, under intentional conditions. The only researcher to find robust spacing effects in word identification procedures has been Greene (1990), who found spaced presentation better than massed presentation but found no lag effect (over lags of 1, 4, and 15).

The data from word fragment completion tests are also not clear cut. Parkin, Reid and Russo (1990) found no spacing effect in primed word fragment completion under conditions in which they did find an effect in recognition. Challis and Brodbeck (1992) obtained small and variable spacing effects in their experiments. Roediger and Challis (1992) found no reliable spacing effect in their Experiment 2 when comparing lags of 0 and 9; in Experiment 3, with more power, the effect was reliable with the same lags. In another experiment they obtained a lag effect across lags of 0, 10, 21, and 31. On the other hand, Greene (1990, Experiment 2) reported a spacing effect but no lag effect on the same kind of test.

In all of the above experiments, repetition, spacing, and lag effects have been examined by repeating items in a long list. Even the longest spacings under these conditions are, in some sense, relatively modest compared to other procedures that might be used. For example, Tulving et al. (1982) showed greater priming from studied items that were re-
exposed in a recognition test prior to being placed in a word fragment completion test, relative to words studied but not tested on the recognition test. Similarly, Musen and Treisman (1990) in a nonverbal priming test and Graf and Mandler (1984) and Chen and Squire (1990) in a word stem completion test showed that repeated presentations in lists or sets of items enhanced performance on the respective tests. Thus, the greater spacing of presentations afforded by these techniques may lead to robust effects on perceptual implicit tests.

Little work on repetition and spacing has been conducted using other implicit tests. Greene (1990) did report a reliable effect on a homophone spelling test. We are unaware of experiments examining spacing effects on conceptual implicit tests such as category association or answering general knowledge questions. However, because Challis and Sidhu (1993) reported positive effects with massed repetition (at least in answering general knowledge questions), it seems likely that spacing will produce positive effects, too, as it does on explicit (conceptual) tests.

Summary of the section
Massed repetition has no effect on implicit perceptual tests but a positive effect on at least one implicit conceptual test. However, beyond this, it seems fair to say that spacing and lag effects on perceptual implicit tests are not well understood. There is likely a small but real effect. As far as we can tell, nothing much hangs on the issue one way or the other in theoretical understanding of implicit memory, and therefore it is a bit surprising the variable has received so much attention. In all likelihood, spacing and lag effects on perceptual implicit tests occur in the same direction, but more weakly, than they do on standard explicit tests. No theory of test differences argues that implicit and explicit tests or perceptual and conceptual tests should differ as a function of every variable examined. Spacing and lag may represent variables that exert similar effects (of variable strength) on most or all tests.

Instructional manipulations at study

Probably the most studied aspects of performance on implicit memory tests are instructional manipulations that are intended to cause subjects to process materials in different ways, to determine the effects on implicit and explicit tests. In this Section we review research in which the following variables were manipulated: incidental versus intentional instructions; divided or focused attention; directed forgetting; levels of processing; imagery; effects of generating compared to reading; effects of varying context; and effects of organization.

Incidental/intentional learning
The distinction between incidental and intentional learning during study parallels the distinction between implicit and explicit instructions at test. In the former contrast, subjects are either forewarned that they will be tested on material to which they are exposed (intentional learning) or they are exposed to the material ostensibly for some other purpose and not warned of a test (incidental learning). Similarly, on explicit tests subjects are directed to recollect past experiences (intentional retrieval is involved), whereas in implicit tests the influence of past experience is revealed on transfer tests that index incidental retrieval. Of course, it is perfectly possible to examine the intentional/incidental study dimension on both types of test, and several such experiments have been reported with perceptual implicit tests. The general conclusion to emerge is that intention to learn has no effect on perceptual priming, but often affects performance on explicit tests.

Greene (1986) manipulated intention to learn using the digit-recall paradigm in which subjects were given digits to remember on every trial and then were given a word to repeat for various periods of time as a distractor task. Some subjects expected a memory test on the words, whereas others did not. Subjects were tested with word stems under either explicit or explicit test instructions. Greene found that intention to learn did not affect priming, but had a
sizeable effect on a word stem cued recall test. (As described previously, the same pattern occurred with rehearsal duration.) Therefore, both variables dissociated the explicit from implicit test and met Schacter et al.'s (1989) retrieval intentionality criterion.

Roediger et al. (1992) also manipulated intention to learn during study and found no effect on primed word stem and word fragment completion tests. However, they also failed to find any effect of intentionality on explicit tests with the same types of cues, unlike Greene (1986). The most likely source of the discrepancy is the differing natures of the orienting task in the two experiments. Roediger et al. did not use the digit word distractor paradigm, but used a standard levels-of-processing manipulation with a graphemic or semantic orienting task. Subjects were informed or not before study about the possible test. Nevertheless, both studies found no effect of intention to learn on implicit tests. Bowers and Schacter (1990) and Neill et al. (1990) also found no effect of intention to learn on tests of word stem completion and word fragment completion, respectively. Neill et al. found an effect of intention to learn on a recognition test.

One exception to the conclusion that intention to learn has no effect on perceptual priming appears in experiments by Whitlow (1989), who varied both the number of repetitions of words and pseudowords and intention to learn. He found that under incidental learning conditions there was no effect of repetition, but that number of repetitions did affect priming under intentional conditions. However, the intentionality variable had no effect on priming of pseudowords.

If we consider only cases in which a single exposure of material occurs, the results are clear-cut: intention to learn has no effect on perceptual priming from a single presentation, even under conditions in which the variable has large effects on explicit tests (e.g., Greene, 1986; Neill et al., 1990). Simply processing a word is sufficient to produce perceptual priming in full force; trying to remember it does not add to perceptual priming.

In an experiment related to this issue, Madigan et al. (1991) primed words by embedding them in the instructions subjects read before trying to complete word fragments. They reported a priming effect under these conditions, in which subjects were almost surely not engaging in intentional retrieval of words presented in the instructions. No list of words was presented prior to the test, so it is unlikely subjects were trying to intentionally recollect recent experiences during the word fragment completion test. However, not everyone obtains this finding (Oliprant, 1983) and the issue of when such incidental exposure will lead to priming and when it will not is complicated; we discuss it below under manipulations of context.

Divided and focused attention
Not many experiments have been reported in which subjects either focus attention on material during learning or divide attention between the learning material and some other task; however, the few available experiments agree that dividing attention during study has no effect on perceptual priming but large effects on explicit measures under the same study conditions. Parkin et al. (1990) had subjects perform a sentence verification task, but half the subjects also had to perform a tone monitoring task simultaneously. Retention was assessed a day later either by an explicit recognition test or an implicit word fragment completion test on items embedded in the sentences. Dividing attention at study affected recognition but had no effect on primed fragment completion. In another experiment on this issue, Parkin and Russo (1990) had subjects, during a first phase, identify pictures from successively presented fragmented forms, in which each successive fragment was more complete than the previous one (Snodgrass, 1989). Half the subjects simultaneously performed the tone monitoring task. A day later, subjects returned and received the same sort of picture identification task, which contained both previously studied items and new items so that savings (priming) could be calculated. Some subjects received a free recall test for the pictures prior to the picture fragment naming test. The results showed that dividing attention at study reliably harmed
recall, but had no effect on savings in the picture fragment identification test. Perceptual priming tests therefore seem to be measuring an automatic process that is unaffected by dividing attention, although surely other, more dramatic methods of dividing attention could affect performance.

Jacoby (1991) and Jacoby et al. (1993) have been especially interested in automatic influences on implicit tests and have developed the process dissociation procedure as a means of assessing a purely automatic component on such tests. The converging operation they have used to assess automatic performance as measured by this procedure is to divide attention at the time of test to reduce consciously controlled recollection. The strategy of dividing attention at test has provided impressive support for estimates of automatic processes obtained from the process dissociation procedure under full attention conditions, supplying further evidence that dividing attention is a useful strategy for studying incidental retrieval.

**Directed forgetting**

In experiments on directed forgetting, subjects study words or other materials and, for some items, are told that they can forget the material. Later subjects are told to try to remember both to-be-remembered (TBR) and to-be-forgotten (TBF) items. The general finding (see Bjork, 1972, 1989, for reviews) is that TBR material is better remembered on explicit tests than is TBF material, although the effect is larger on recall tests than on recognition tests (e.g., Elmes, Adams and Roediger, 1970), where it often disappears altogether (Block, 1971). The TBF and TBR items can be cued either on an item-by-item basis or on the basis of blocks of items (i.e., the instruction to remember or forget is given after each word or after blocks of words). The item-by-item method usually produces stronger effects (e.g., Basden, Basden and Gargano, 1993). Various theoretical mechanisms have been implicated in explaining directed forgetting, from rather prosaic differential rehearsal interpretations (more rehearsal for TBR than TBF words) to more mysterious retrieval inhibition mechanisms for TBF material (see Bjork, 1989).

From the foregoing review we might expect that directed forgetting would have little or no effect on implicit memory tests. If study time, massed presentations, intentionality of learning, and divided attention do not affect performance on perceptual implicit memory tests, then what chance would directed forgetting have, since the instruction is always given after the word has been fully encoded? In general, the literature bears out the supposition that the answer to this question is ‘very little’. However, in the first experiment reported on this issue, MacLeod (1989) did find that directed forgetting reduced priming on both a word fragment completion test (Experiment 1) as well as a lexical decision test (Experiment 2). Averaging across conditions, however, the directed forgetting effect was only 5% on primed word fragment completion and 12 msec in the lexical decision task. Although direct comparison to the effects of directed forgetting on explicit tests is obviously hazardous, the effects were quite substantial in the companion recognition and free recall conditions.

Later research has failed to produce even the slight effects that MacLeod (1989) obtained. Paller (1990) examined priming on the word stem completion test using the word-by-word cueing method and found no effect on word stem completion, although robust directed forgetting effects were obtained on word stem cued recall and free recall. Thus, Paller’s (1990) directed forgetting data satisfied Schacter et al.’s (1989) retrieval intentionality criterion; in addition, these conclusions from behavioral data were complemented by electrophysiological measures taken during study, which also led to the conclusion that performance on the explicit tests differed as a function of encoding processes, whereas performance on the implicit test did not.

Basden et al. (1993) examined directed forgetting on primed word fragment completion and on a word association test, thus providing both perceptual and conceptual implicit memory measures. In three experiments they found no effect of directed forget-
ting on either implicit test using both the item-by-item and the blocking method of directed forgetting. In most (but not all) conditions they did find directed forgetting effects in recall or in recognition, especially with the item-by-item method. In Experiment 4, Basden et al. did replicate MacLeod’s (1989) directed forgetting effect on word fragment completion with the word-by-word method.

Taken together, the few extant directed forgetting experiments reveal either little or no effect of this variable on perceptual priming under conditions in which large effects are obtained in free recall. Combined with the foregoing evidence on other variables, the evidence to date suggests that instructions to forget items after they have been studied have little influence on perceptual implicit memory tests, in line with the idea that these tests are little affected by subjects’ strategies engaged during study. The next Section is also concerned with this issue.

Levels of processing effects

Certainly the most popular variable examined on implicit memory tests has been levels of processing, in which subjects are exposed to material and directed to make ratings that focus attention on graphemic, phonemic, or semantic aspects of the material (Craik and Lockhart, 1972; Craik and Tulving, 1975; Hyde and Jenkins, 1969). This levels of processing variable has huge effects on standard recognition and recall tests, although explicit test conditions can be arranged that make the effect vanish (Fisher and Craik, 1977; McDaniel, Friedman and Bourne, 1978) or even reverse (Morris, Bransford and Franks, 1977). In general, these explicit tests show that no effect of level of processing occurs when the test is focused on lexical or phonemic properties of words rather than their meanings. Because perceptual implicit tests are also believed to access such lexical information (Nelson et al., 1987; Roediger et al., 1992; Tulving and Schacter, 1990; Weldon, 1991), we might expect that levels of processing would exert little or no effect on perceptual implicit memory tests, too.

The earliest experiments on levels of processing provided ringing support for this hypothesis. Jacoby and Dallas’ (1981) varied level of processing and showed no effect on priming in word identification under conditions that produced huge effects in recognition. In fact, this impressive demonstration was one that sparked the initial interest in studying implicit or indirect memory measures in normal subjects. Soon after Jacoby and Dallas’ (1981) demonstration, Graf, Mandler and Haden (1982) and Graf and Mandler (1984) showed that levels of processing similarly had little or no effect on word stem completion under conditions in which large effects were seen in recognition and word stem cued recall.

On the other hand, 10 years later it is clear that the assertion of the null hypothesis — levels of processing never has any effect on perceptual implicit tests — is wrong, at least as the variable is usually manipulated. Most experiments do report a slight (and sometimes significant) levels effect, although its magnitude is invariably modest, at best. However, the conclusion that levels of processing has absolutely no effect on performance is clearly wrong.

We base our summary here on two recent reviews of this literature by Challis and Brodbeck (1992) and by Brown and Mitchell (1992). These cover overlapping, but not identical, sets of experiments. Challis and Brodbeck (1992) reviewed some 35 experiments and added several of their own; Brown and Mitchell (1992) considered 114 comparisons from 30 experimental reports. (Unlike most experiments under consideration here, their reviews included data from memory-impaired patients.) Both sets of researchers concluded that levels of processing effects occur on implicit tests, but they differ somewhat in their interpretations of this state of affairs. Challis and Brodbeck (1992) argued that levels of processing effects are small but ubiquitous, yet are more likely to be found in certain types of experimental designs than others. In particular, their experiments revealed significant levels of processing effects in primed word fragment completion when the variable was manipulated between subjects (each group processed items with one orienting task) or within subjects when the variable was blocked (i.e., subjects processed a block of items with one ori-
ting task). When the orienting task changed randomly from item to item in their experiments, levels of processing effects were negligible. Although the results are consistent within their experiments, their meta-analysis of the prior literature did not bear out the importance of the between/within manipulation and neither did that of Brown and Mitchell (1992). Of course, in the literature at large, the between/within and blocked/random manipulations were not varied within one experiment, so confounding factors may have added noise and obscured a subtle effect.

Brown and Mitchell (1992) also emphasize the ubiquity of small levels of processing effects in implicit memory experiments but, unlike Challis and Brodbeck (1992), they considered experiments using both perceptual and conceptual implicit tests. (Challis and Brodbeck concentrated on the former.) Brown and Mitchell showed that the magnitude of levels of processing effects was smaller on perceptual than on conceptual implicit tests. Importantly, they also showed that the magnitude of the levels of processing effects on conceptual implicit tests correlated significantly with explicit measures of memory included in the same studies \( r = 0.70 \), but that the correlation between priming on perceptual implicit tests and companion explicit measures was negligible \( r = 0.14 \). (Of course, the accuracy of this conclusion hinges on there having been enough variability in the priming scores to permit reliable correlations, which is sometimes not the case; Ostergaard, 1992.) Therefore, levels of processing effects may occur to some small degree in perceptual implicit memory tests, but they seem uncorrelated with explicit measures of memory that have a conceptual base.

Levels of processing effects on explicit measures do correlate highly with priming on conceptual implicit tests (e.g., Hamann, 1990). In addition, Challis and Sidhu (1993) showed levels of processing effects in both explicit and implicit versions of the question answering task used by Blaxton (1989), although the effects were greater under explicit test conditions.

Why do the slight levels of processing effects occur on perceptual implicit tests? Perhaps there is a slight contamination from intentional retrieval processes that accounts for the effect and for the slight positive correlation. Schacter et al. (1989) built the logic of the retrieval intentionality criterion on dissociations of perceptual implicit tests from explicit tests with the same cues as a function of levels of processing (although other variables would do as well). Some studies have met this criterion reasonably well; for example, Roediger et al. (1992) found no effect of levels of processing on primed word fragment or word stem completion, but strong effects on explicit versions of both tests with the same cues (fragments or stems). It may be, then, that the levels of processing effect on perceptual implicit tests can be used as an index of how contaminated a test is by explicit retrieval processes. In addition, Toth, Reingold and Jacoby (1993) have shown that a small levels of processing effect on primed word stem completion vanishes when the process dissociation procedure is carried out and the automatic component of performance is calculated. However, if the small levels of processing effect sometimes found on perceptual implicit tests is due to contamination from explicit memory, then it is unclear why the correlation is not higher between the levels effect on explicit and perceptual implicit tests (although, as noted above, there may be a constricted range problem in priming scores; Ostergaard, 1992).

Another reason that levels of processing effects may occur on perceptual implicit tests is that the standard manipulation of the variable may affect both perceptual and conceptual aspects of processing (Challis and Brodbeck, 1992). The general assumption has been that the standard level of processing manipulation affects only conceptual elaboration and that perceptual (data-driven) encoding processes are the same across the manipulation. However, in the standard manipulation, subjects know before they see an item what task they will perform on it (Does it have an e? Does it fit in this sentence?, etc.). It may be that in graphemic or even phonemic processing tasks subjects truncate their perceptual analysis of the studied item, relative
to the semantic processing condition. That is, in checking for an e, or counting the number of syllables or the number of ascending and descending letters, subjects may be less likely to process the word as a lexical unit. If perceptual/lexical processing is short-circuited during study in the ‘shallow’ processing conditions relative to the ‘deep’ conditions, this would explain why the ‘shallow’ conditions show slightly less priming on the perceptual implicit tests. This hypothesis would also help explain Challis and Brodbeck’s (1992) finding that manipulating levels of processing in between-subjects (or in blocked, within-subjects) designs enhances the effect. In the shallow processing conditions of these designs, subjects can set themselves to (say) check for e’s and then ignore to a greater extent the lexical nature of the whole word; such a strategy would be less practicable in within-subjects designs when the processing task varies unpredictably from item to item (although of course subjects are still informed of the task before the items are presented).

Some evidence for this view comes from an experiment by Hayman and Jacoby (1989). The study phase consisted of a letter identification test in which subjects were cued as to what letters they should attempt to identify either before a word was presented or after the word was presented. (The word was always presented for 83 msec). The prediction was that if subjects knew the letters they were to seek before the word was displayed, they would truncate processing of the item as a lexical unit. Therefore, to the extent the manipulation was successful, they should not obtain priming on a later word identification test. On the other hand, when the letters were cued after the word was presented, it would be treated as a lexical unit and priming would occur. The results conformed nicely to these predictions, thereby supporting the conclusions that (a) lexical processing of the word is necessary for priming to occur, and (b) if such processing is truncated, priming will not occur. In the present context, these results make it more plausible that truncated processing of the word as a unit under shallow levels of processing (especially in blocked or between-subjects designs) could account for the small levels of processing effects sometimes observed in perceptual priming.

This truncated processing hypothesis of the slight levels of processing effects found in perceptual implicit tests can be tested in experiments in which subjects see target words and then get the instruction about the processing task after they have read the word. Under these conditions one can be more certain that initial perceptual processing is equated and that the levels of processing manipulation is affecting elaboration and not initial encoding. Manipulating levels of processing in this manner produces the usual effect on explicit tests (Moeser, 1983), but to our knowledge it has never been tried in implicit tests.

**Imagery**

It has long been known that forming an image of a word’s referent enhances later performance on standard, explicit memory tests (e.g., Paivio, 1986). Blaxton (1989, Experiment 3) tested the effects of imagery on two implicit tests: answering general knowledge questions and completing word fragments. She found that imagining referents of words enhanced performance on the conceptual test (answering general knowledge questions) but had no general effect on the perceptual test (word fragment completion). Durgunoğlu and Roediger (1987) also found no effect of imaging words’ referents on primed word fragment completion, although the manipulation did positively affect free recall and recognition.

Blaxton’s (1989) and Durgunoğlu and Roediger’s (1987) experiments, like most imagery studies, examined the effects of imagery as a conceptual variable; it is typically assumed, though, that the formation of images uses some of the same perceptual mechanisms as used in perceiving (Kosslyn, 1980). If so, can imagining relevant scenes be used to mimic the processes required to produce priming on perceptual implicit tests? In the prior experiments we would not expect to see a positive effect, because creation of images would rely on a very different form (perception of objects) than would be accessed by the verbal perceptual test (perception of
word forms). Could imagining a word form at study produce positive transfer on a verbal implicit perceptual test? Jacoby and Witherspoon (1982) presented evidence in support of this hypothesis. They had subjects spell auditorily presented words or simply listen to them. Spelling words produced more priming on a later word identification test than merely listening to words; indeed, spelling words produced almost as much priming as actually having read the words. Donnelly (1988) replicated this finding on a word fragment completion test.

Roediger and Blaxton (1987a) tested the imagery hypothesis directly, having subjects read, hear, or hear and image the word as it would appear printed or typed. On a word fragment completion test the image-the-word condition produced priming levels that were greater than those from regular auditory presentation, although less than from visual presentation. To further explore this effect, Roediger et al. (1992) gave subjects pictures and asked them either to form images of the names of the pictures (in order to count the number of ascending and descending letters) or to rate the pictures for pleasantness. In a later word fragment completion test they found greater priming from pictures whose names had been imaged. The same trend appeared on a word stem completion test, although it was not significant.

Similarly, Schacter and Graf (1989) reported that having subjects form visual images of auditorily presented words produced effects on word stem completion that were similar to those obtained from visual presentation of the word. Finally, Weldon (1991) showed that when the study material was anagrams (e.g., geldon), only those subjects who were instructed to mentally transpose the vowels to produce a word showed priming on word fragment completion and word identification tests when these tests gave clues to the target word (golden in this case).

All of the aforementioned experiments have involved examination of the effects of imagery on verbal implicit tests. The results have been positive in showing that imagining words from auditory or pictorial presentations enhances priming on verbal perceptual implicit tests. Thus far, little work has been done examining effects on nonverbal tests. Preliminary evidence seems to indicate that nonverbal priming from imagery is possible; McDermott and Roediger (1992b) have found that when subjects are presented with words and asked to image the referents, the obtained level of priming on a picture fragment naming test is approximately midway between priming of words given pleasantness ratings during study and that of studied pictures.

In summary, imagery seems to exert an effect on both conceptual and perceptual implicit memory tests, although for different reasons. The effects of imaging on tests such as answering general knowledge questions can be explained by drawing an analogy to the picture superiority effect; images may create dual codes (Paivio, 1986) or more meaningful elaborations (Nelson, 1979) and therefore transfer better to conceptual implicit (and explicit) tests. However, forming images can also have an effect on implicit perceptual tests if the test accesses perceptual operations used during encoding. Indeed, perceptual implicit tests may provide a new means of studying the effects of imagery.

Generating and reading
Winnick and Daniel (1970) reported an interesting experiment that, in retrospect, was 8 to 13 years ahead of its time. In their Experiment 2, subjects read words or generated them either from pictures (i.e., they named the pictures) or from definitions (i.e., they generated words when given their definitions). After this study phase, one group of subjects recalled the items in verbal form under free recall conditions, whereas a different group of subjects identified the words from tachistoscopic displays. In the explicit test, these researchers found what would later be called the generation effect (Sramecka and Graf, 1978): recall was better for items that had been generated from either pictures or definitions than for those that had been read. In the tachistoscopic identification test, however, Winnick and Daniel found greater priming from reading words than from generating them from either pictures or definitions; indeed priming on the word
identification test was negligible for both generation conditions. Years later, Jacoby (1983b) produced a similar negative generation effect in three experiments in which he had subjects produce antonyms from context words. That is, subjects either read words such as short or generated them from antonyms (tall -???). Generated words were recognized better than read words on the explicit test, but on an implicit word identification test, there was less priming from generated words.

Many claimed dissociations between explicit and implicit memory tests take the form of a variable affecting one test and having no effect on the other (e.g., levels of processing or intention to learn). The demonstrations of Winnick and Daniel (1970) and Jacoby (1983b) are especially impressive in that they showed opposite effects of a variable on the two types of test. Such interactions provide a more solid basis for concluding that different underlying factors are at work in the two tests (although see Dunn and Kirsner, 1988). Therefore, this generate/read manipulation has attracted considerable attention. Indeed, Blaxton (1989), Roediger and Blaxton (1987b) and Roediger et al. (1989b) suggested that the generate/read manipulation could be used to operationally define perceptual and conceptual implicit tests that were verbal in nature. (Several converging operations were also suggested.) In reading a word, the naming response is produced via bottom-up processes, with the data (the letters) driven up through the perceptual system. Generating occurs through top-down conceptual processes, because there is no perceptual resemblance between the cue and the word produced. Because generating from a definition, picture, or associative clue is a conceptual operation, a test showing an advantage of prior generating to reading during the study phase should be considered a conceptual test (or conceptually driven). On the other hand, reading a word out of context must be carried off by driving the information bottom-up through the perceptual system, and, relative to generating from conceptual clues, prior reading of words should transfer better to implicit tests that involve trying to name words from data-limited word cues (fragments, stems, brief flashes, etc.).

The idea that one can determine the relative perceptual and conceptual components in an implicit memory test from the generate/read contrast has produced a considerable amount of research. Most can be readily interpreted within this framework, but there are clearly problems, too. The approach today seems oversimplified, at best. There are several sources of difficulty, some logical and some empirical. Because no one has attempted to review this literature in detail, we provide a summary here.

Blaxton (1985, 1989; see also Roediger and Blaxton, 1987b) extended Jacoby's (1983b) method of using the Read/Generate contrast to dissociate perceptual from conceptual tests. She employed three study conditions: subjects read words out of context (eagle), read them in context (hawk-eagle) or generated them from conceptual clues and their first letters (hawk-e__). In each case subjects produced the word aloud during study, but the cognitive processes driving the production differed across conditions. When reading words out of context, subjects must derive the name from the perceptual clues on the page (thereby engaging in data-driven processing); when generating words from conceptual clues, they are engaging in conceptually driven processing (although the first letter provided a perceptual hint in Blaxton's experiment, but did not in Jacoby's or Winnick and Daniel's experiments). The read-in-context condition presumably involves a mix of data-driven and conceptually driven processing. Note that Blaxton (and Jacoby, 1983b) used two different reading conditions, in context and out of context. The former is the usual standard of comparison for the generation condition in the generation effect literature, because reading in context often produces greater performance than reading out of context on explicit tests and the interest in the standard case is on positive effects of generation. However, in contrast, for purposes of examining implicit tests, the contrast between reading out of context and generating is most
apt to provide the maximal contrast between data-driven and conceptually driven processing during study.

Blaxton (1989) employed five different tests, three explicit and two implicit, as seen in Table 5. Two tests were hypothesized to be perceptual and three to be conceptual; the results obtained fully backed this categorization. Prior generation of words led to better performance than reading words on conceptual tests, whereas the reverse occurred on perceptual tests. Note that in Table 5 explicit tests were dissociated from one another (free recall and semantic cued recall differed from graphemic cued recall) and so were implicit tests. (Primings in answering general knowledge questions differed from primed word fragment completion). The tests divided themselves as a function of the match between processing requirements between study and test; study conditions involving reading transferred better to perceptual tests (whether explicit or implicit) whereas those involving generating produced better performance on conceptual tests (again, whether explicit or implicit). Blaxton’s (1989) results have generally replicated well in normal subjects (e.g., Blaxton, 1992; Srinivas and Roediger, 1990; Tajika and Neumann, 1992). We turn next to consider how well Blaxton’s findings have held up for the two classes of implicit test (conceptual and perceptual) across the literature.

In conceptual implicit tests, the results generally replicate Blaxton (1989), but there are relatively few studies. Relevant data are provided in Table 6. Blaxton (1992, Experiments 1 and 2) replicated twice the finding that generating produces greater priming than does reading or answering general knowledge questions, at least in normal subjects. (Memory-impaired patients who were epileptic and had left temporal lobe foci for their seizures failed to show this generation effect). Tajika and Neumann (1992, Experiment 1) replicated Blaxton’s three study conditions and showed the same pattern of outcome in answering general knowledge questions: reading out of context produced least priming, reading in context led to intermediate priming, and generating produced greatest priming. However, in Experiment 2, using only the No Context and Generate conditions, Tajika and Neumann failed to find a generation effect; reading and generating led to equal priming. Tajika and Neumann offer no explanation for the failure to replicate the effect in their Experiment 2.

Turning to other conceptual implicit tests, Srinivas and Roediger (1990, Experiment 1) showed that generating produced more priming than reading in or out of context on the test of category instance production. In another paradigm, Smith and Branscombe (1988) showed greater priming from generating than from reading out of context on a test of category accessibility (providing a trait adjective to describe behaviors of a person; see Hig-
The evidence seems relatively consistent, then, that generating produces greater priming than does reading on these three conceptual implicit tests.

Moving on to perceptual implicit tests, the Read/Generate comparison has been made numerous times since the original Winnick and Daniel (1970), Jacoby (1983b), and Blaxton (1989) results (all of which showed greater priming from Read than from Generate conditions on word identification or word fragment completion tests). Most experiments have used word fragment completion, word stem completion, or word identification as the implicit measures. We consider each of these in turn, with results summarized in Tables 7 and 8. We include only experiments in which (a) words were read (either in or out of context), (b) words were generated from a conceptual clue with zero or one letter accompanying the clue, and (c) the test item was presented visually without a meaningful context. These are the conditions first used by Winnick and Daniel (1970) and Jacoby (1983) and define the phenomenon of interest.

### Table 6

<table>
<thead>
<tr>
<th>Reference</th>
<th>Test</th>
<th>Nonstudied</th>
<th>Read No context</th>
<th>Read Context</th>
<th>Generate</th>
<th>Generate Difference (G – NC)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blaxton (1989, Expt. 1)</td>
<td>Answering general knowledge questions</td>
<td>0.25</td>
<td>0.08</td>
<td>0.13</td>
<td>0.25</td>
<td>+0.17</td>
</tr>
<tr>
<td>Blaxton (1992)*</td>
<td>Answering general knowledge questions</td>
<td>0.57</td>
<td>0.13</td>
<td>0.29</td>
<td>+0.16</td>
<td></td>
</tr>
<tr>
<td>Expt. 1</td>
<td></td>
<td>0.40</td>
<td>0.15</td>
<td>0.30</td>
<td>+0.15</td>
<td></td>
</tr>
<tr>
<td>Expt. 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tajika and Neuman (1992)</td>
<td>Answering general knowledge questions</td>
<td>0.40</td>
<td>0.05</td>
<td>0.25</td>
<td>0.31</td>
<td>+0.26</td>
</tr>
<tr>
<td>Expt. 1</td>
<td></td>
<td>0.41</td>
<td>0.19</td>
<td>0.20</td>
<td>+0.01 n.s.</td>
<td></td>
</tr>
<tr>
<td>Expt. 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Srinivas and Roediger</td>
<td>Category instance production</td>
<td>0.16</td>
<td>0.07</td>
<td>0.09</td>
<td>0.16</td>
<td>+0.11</td>
</tr>
<tr>
<td>(1990, Expt. 1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Smith and Branscombe</td>
<td>Category accessibility</td>
<td>0.34</td>
<td>0.09</td>
<td>0.18</td>
<td>+0.09</td>
<td></td>
</tr>
<tr>
<td>(1988)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*a Data estimated from figures. n.s., not significant.*

The relevant word fragment and word stem completion results are shown in Table 7, where the experiment, nonstudied baserate, priming scores for Read and Generate conditions and their difference (Generate – No Context Read) are listed. As can be seen, all 19 relevant comparisons using these two measures show priming to be greater following Read than Generate conditions, although the effect was not statistically significant in all cases. In particular, Komatsu and Naito (1992) found only 0.07 and 0.02 advantages of Read to Generate in their experiments using Japanese word fragments. Schwartz (1989) similarly found only slight advantages of the Read condition in her two experiments, where the word stem was four letters of five-letter words (i.e., only one letter needed to be added). It is unclear why these two sets of researchers reported results discrepant from the rest, but each involved either materials (Japanese words in different scripts) or procedures (stem completion with 80% of the letters in the stem) less typical of implicit memory research using these paradigms.

In addition to the results reported in Table 7,
TABLE 7
Primin in word fragment and word stem completion following a Read/Generate manipulation

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Test</th>
<th>Nonstudied</th>
<th>Read</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>No context</td>
<td>Context</td>
<td>Generate</td>
</tr>
<tr>
<td>Bassili, Smith and MacLeod (1989)</td>
<td>WSC</td>
<td>0.27</td>
<td>0.27</td>
<td>0.27</td>
<td>0.13</td>
<td>-0.14</td>
</tr>
<tr>
<td>Blaxton (1989, Expt. 1)</td>
<td>WFC</td>
<td>0.27</td>
<td>0.48</td>
<td>0.35</td>
<td>0.19</td>
<td>-0.29</td>
</tr>
<tr>
<td>Blaxton (1992)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exp. 1</td>
<td>WFC</td>
<td>0.47</td>
<td>0.39</td>
<td>0.24</td>
<td>-0.15</td>
<td></td>
</tr>
<tr>
<td>Exp. 2</td>
<td>WFC</td>
<td>0.45</td>
<td>0.41</td>
<td>0.28</td>
<td>-0.13</td>
<td></td>
</tr>
<tr>
<td>Java (1993)</td>
<td>WSC</td>
<td>0.14/0.15</td>
<td>0.33</td>
<td>0.19</td>
<td>-0.14</td>
<td></td>
</tr>
<tr>
<td>Java and Gardiner (1991)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>young</td>
<td>WSC</td>
<td>0.18</td>
<td>0.30</td>
<td>0.23</td>
<td>-0.07</td>
<td></td>
</tr>
<tr>
<td>old</td>
<td>WSC</td>
<td>0.16</td>
<td>0.27</td>
<td>0.15</td>
<td>-0.12</td>
<td></td>
</tr>
<tr>
<td>Komatsu and Naito (1992)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Expt. 1</td>
<td>WFC</td>
<td>0.44</td>
<td>0.23</td>
<td>0.16</td>
<td>-0.07 n.s.</td>
<td></td>
</tr>
<tr>
<td>Expt. 2</td>
<td>WFC</td>
<td>0.28</td>
<td>0.20</td>
<td>0.18</td>
<td>-0.02 n.s.</td>
<td></td>
</tr>
<tr>
<td>McClelland and Pring (1991)</td>
<td>WSC</td>
<td>0.06</td>
<td>0.42</td>
<td>0.21</td>
<td>-0.21</td>
<td></td>
</tr>
<tr>
<td>Roediger, Guynn and Jones (in press)</td>
<td>WFC</td>
<td>0.47</td>
<td>0.16</td>
<td>0.06</td>
<td>-0.10</td>
<td></td>
</tr>
<tr>
<td>Schwartz (1989, Expt. 1)</td>
<td>WSC</td>
<td>0.47</td>
<td>0.18</td>
<td>0.16</td>
<td>-0.02 n.s.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.46</td>
<td>0.18</td>
<td>0.15</td>
<td>-0.03 n.s.</td>
<td></td>
</tr>
<tr>
<td>Smith and Branscombe (1988)</td>
<td>WFC</td>
<td>0.41</td>
<td>0.21</td>
<td>0.02</td>
<td>-0.19</td>
<td></td>
</tr>
<tr>
<td>Srinivas and Roediger (1990, Expt. 1)</td>
<td>WFC</td>
<td>0.21</td>
<td>0.24</td>
<td>0.20</td>
<td>0.14</td>
<td>-0.10</td>
</tr>
<tr>
<td>Tajika and Neuman (1992)</td>
<td>WSC</td>
<td>0.15</td>
<td>0.22</td>
<td>0.07</td>
<td>0.12</td>
<td>-0.10</td>
</tr>
<tr>
<td>Expt. 1</td>
<td>WSC</td>
<td>0.11</td>
<td>0.46</td>
<td>0.29</td>
<td>-0.17</td>
<td></td>
</tr>
<tr>
<td>Expt. 2</td>
<td>WSC</td>
<td>0.30</td>
<td>0.24</td>
<td>0.14</td>
<td>-0.10 p</td>
<td></td>
</tr>
<tr>
<td>Toth, Reingold and Jacoby (1993)</td>
<td>WSC</td>
<td>0.30</td>
<td>0.24</td>
<td>0.14</td>
<td>-0.10 p</td>
<td></td>
</tr>
</tbody>
</table>
TABLE 7  Continued

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Test</th>
<th>Nonstudied</th>
<th>Read</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>No context</td>
</tr>
<tr>
<td>Weldon (1991, Expt. 1)</td>
<td>WFC</td>
<td>0.29</td>
<td>0.25</td>
</tr>
</tbody>
</table>

*a Estimated from figures;  
*b Generate — Read-in-context; n.s., not significant.

Blaxton (1992) reported three other relevant contrasts with memory-impaired patients that also showed significant benefits of Read to Generate conditions in primed word fragment completion (with advantages of 0.14, 0.13, and 0.22 in three different patient groups). In general, the great bulk of research with word fragment and word stem completion shows greater priming from prior reading of words than from generating words from conceptual clues.7

Turning to word identification (or perceptual identification), matters grow more complex. In these paradigms the test involves brief flashes of intact words and the measure is simply probability of naming the word, or the reduced threshold for doing so, following various study conditions. This task is the one in which the superiority of Read to Generate conditions was discovered by Winnick and Daniel (1970) in a threshold procedure. Clarke and Morton (1983) replicated these findings using a similar procedure. Jacoby (1983b) changed the procedure from a threshold measure of priming to giving subjects one brief glimpse of the stimulus during the test and measuring the probability correct, but this test, too, revealed more priming from a Read than from a Generate condition in three experiments.

Table 8 summarizes the results of which we are aware in which a word identification procedure was used. (Unless otherwise noted, the differences reported in the table are reliable). In 13 of the 18 nontried comparisons in the table, the Read condition produced more priming than did the Generate condition. All the statistically significant results show this pattern.

The primary exception to the usual pattern of Read producing greater priming than Generate in word identification appears in several experiments of Masson and MacLeod (1992). In these cases, they found priming after a Generate condition equal to that found after a Read condition. They replicated Jacoby’s (1983b) finding of Read superiority when the generation task and materials involved antonyms (Experiment 2) or idioms (e.g., kick the b———, Experiment 6). However, in their other experiments in which subjects generated words from definitions, synonyms, or sentences, there was no difference between Read and Generate conditions despite large differences (favoring the Generate condition) in recognition.

It is unclear why Masson and MacLeod’s results differ from those described above. For example, in their Experiment 1 (in which subjects generated from a definition and first letter) they found that a Generate condition produced slightly greater performance than the Read condition, although the difference was not significant; however, neither Win-
### Table 8

**Priming in word identification experiments following a Read/Generate manipulation**

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Nonstudied</th>
<th>No context</th>
<th>Context</th>
<th>Generate</th>
<th>Difference (G – R)</th>
<th>Generation cue</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clarke and Morton</td>
<td>35.6 ms</td>
<td>6.7</td>
<td>0.1</td>
<td>– 6.6</td>
<td>Definition</td>
<td></td>
</tr>
<tr>
<td>(1983)a</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jacoby (1983) Expt. 1</td>
<td>0.47</td>
<td>0.15</td>
<td>0.08</td>
<td>0.00</td>
<td>– 0.15</td>
<td>Antonym</td>
</tr>
<tr>
<td>Expt. 2</td>
<td>0.60</td>
<td>0.22</td>
<td>0.16</td>
<td>0.07</td>
<td>– 0.15</td>
<td>Antonym</td>
</tr>
<tr>
<td>Expt. 3</td>
<td>0.47</td>
<td>0.18</td>
<td>0.11</td>
<td>0.05</td>
<td>– 0.13</td>
<td>Antonym</td>
</tr>
<tr>
<td>Masson and MacLeod</td>
<td>0.47</td>
<td>0.12</td>
<td>0.15</td>
<td>+0.03 n.s.</td>
<td>Definition and first letter</td>
<td></td>
</tr>
<tr>
<td>(1992) Expt. 1</td>
<td>0.37</td>
<td>0.13</td>
<td>–0.01</td>
<td>– 0.14</td>
<td>Antonym</td>
<td></td>
</tr>
<tr>
<td>Expt. 2</td>
<td>0.58</td>
<td>0.07</td>
<td>0.09</td>
<td>+0.02, n.s.</td>
<td>Synonym/Associate</td>
<td></td>
</tr>
<tr>
<td>Expt. 4</td>
<td>0.60</td>
<td>0.12</td>
<td>0.12</td>
<td>0.00, n.s.</td>
<td>First name</td>
<td></td>
</tr>
<tr>
<td>Expt. 5</td>
<td>0.60</td>
<td>0.10</td>
<td>0.11</td>
<td>+0.01, n.s.</td>
<td>Phrase</td>
<td></td>
</tr>
<tr>
<td>Expt. 6</td>
<td>0.50</td>
<td>0.19</td>
<td>0.10</td>
<td>– 0.09</td>
<td>Idiom</td>
<td></td>
</tr>
<tr>
<td>Expt. 7</td>
<td>0.52</td>
<td>0.20</td>
<td>0.20</td>
<td>0.00, n.s.</td>
<td>Sentence</td>
<td></td>
</tr>
<tr>
<td>Expt. 11</td>
<td>0.68</td>
<td>0.07</td>
<td>0.16</td>
<td>+0.09, n.s.</td>
<td>Synonym</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.61</td>
<td>0.08</td>
<td>0.11</td>
<td>+0.03, n.s.</td>
<td>Antonym</td>
<td></td>
</tr>
<tr>
<td>Schwartz (1989) Expt. 2a</td>
<td>0.42</td>
<td>0.19</td>
<td>0.12</td>
<td>– 0.07</td>
<td>Definition and first letter</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.54</td>
<td>0.17</td>
<td>0.12</td>
<td>– 0.05</td>
<td>Definition</td>
<td></td>
</tr>
<tr>
<td>Weldon (1991) Expt. 1</td>
<td>0.57</td>
<td>0.17</td>
<td>0.08</td>
<td>– 0.09</td>
<td>Conceptual clue and first letter</td>
<td></td>
</tr>
<tr>
<td>Winnick and Daniel</td>
<td>61.9 ms</td>
<td>8.5</td>
<td>3.4</td>
<td>– 5.1</td>
<td>Definition</td>
<td></td>
</tr>
<tr>
<td>(1970)a</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Expt. 2</td>
<td>57.2 ms</td>
<td>12.5</td>
<td>0.0</td>
<td>– 12.5</td>
<td>Definition</td>
<td></td>
</tr>
<tr>
<td></td>
<td>68.7 ms</td>
<td>16.2</td>
<td>3.1</td>
<td>– 13.1</td>
<td>Definition</td>
<td></td>
</tr>
<tr>
<td></td>
<td>59.7 ms</td>
<td>13.8</td>
<td>3.8</td>
<td>– 10.0</td>
<td>Definition</td>
<td></td>
</tr>
</tbody>
</table>

*a Visual duration thresholds were the measures taken. n.s., not significant.*

Nico and Daniel (1970) nor Clarke and Morton (1983) obtained any priming relative to the baseline in their conditions in which subjects generated from definitions. Masson and MacLeod (1992) do not comment on the discrepancy, but one possible explanation for the varying outcomes is slight differences in the identification procedures used. On the other hand, given that the Read to Generate superiority has been found by so many other researchers using both word identification and other perceptual priming measures, and given that Masson and MacLeod do obtain the effect under certain conditions, it seems unlikely that any simple procedural difference provides an answer to this puzzle.

Masson and MacLeod (1992) offer two primary reasons that priming after Read conditions equaled priming after Generate conditions in some of their experiments: (1) integrating words into a context produces less perceptual enhancement, and this occurs under some generation conditions; and (2) subjects interpret words differently — assign different meanings — in Generate and Read conditions, with greater priming occurring in the Read condition.
because subjects are more likely to assign the same meaning at study and test. These are interesting ideas, but they seem at variance with a considerable amount of the literature, because in other experiments subjects show an advantage of reading to generating under conditions that Masson and MacLeod claim are nonintegrative (e.g., generating from definitions or sentences). More importantly, evidence to refute both hypotheses exists in these authors’ own data. In Experiment 4, they used strongly unitized materials by having subjects generate first names from last names and initials (e.g., M. Monroe) or by reading the name in isolation (Marilyn). They found Read equal to Generate even under these conditions, in which items formed strong integrated units. Similarly, the authors also directly tested the change-of-meaning idea by using homographs and using cues that would lead subjects to generate the primary or secondary meaning of the homograph. However, priming was equal in the two conditions and did not differ from the reading condition (Experiment 5), contradicting the change-of-meaning hypothesis under consideration.

At this writing, it is impossible to provide an account for Masson and MacLeod’s results showing priming in word identification to be the same following Read and Generate conditions and simultaneously to explain the numerous findings in which Read conditions produced greater priming than Generate conditions. It will be interesting to see if others can replicate and confirm the findings of Masson and MacLeod (1992), because most experiments using perceptual implicit tests have shown priming from Read conditions to be greater than that from Generate conditions. However, we suspect that finding is replicable, because M. Khan (working in the senior author’s laboratory) has conducted several experiments in which priming from generating words equaled that of reading words on word fragment completion tests. Therefore, it remains a puzzle as to what factors determine whether reading produces greater priming than does generating.

The foregoing paragraphs have concentrated on one type of Read/Generate comparison, viz., when subjects read a word (usually out of context) or generate the word from a conceptual clue (with at most only one letter presented). There are many other sorts of generation procedures that researchers have used. We have previously discussed the work by Gardiner (1988, 1989; Gardiner et al., 1989) that showed generation of information from the same fragment as used on the test enhanced priming relative to reading the word. When the fragment was changed between study and test, prior reading of the word produced the same amount of priming as generating it (see also Toth and Hunt, 1990, Experiment 1).

Other researchers have compared conditions in which subjects read a word to those in which subjects generate the word from an anagram or some other perceptually altered display. This type of generation seems quite different from other examples; although rearranging letters to form a word surely involves knowledge of lexical rules, the generation involved is likely not the same as generating from a meaningful clue. Presented in Table 9 are results from experiments in which researchers have made such comparisons (Schwartz, 1989; Weldon, 1991; Wippich, 1992). In each case, subjects read words or generated them from spatial rearrangements of their letters (anagrams or scattered words) that presumably required more extensive lexical processing than does reading. Subjects were tested on three different perceptual implicit tests (word identification, word stem completion, or word fragment completion) in these experiments. The results in Table 9 reveal an interesting pattern: priming in word identification was greater following reading than generating words, but this outcome disappeared on word stem and word fragment tests. This interaction is not the result of comparing across experiments, because both Schwartz (1989) and Weldon (1991) obtained this pattern within their experiments. They both concluded that word identification is a more perceptual or data-driven test than is stem completion.

Viewing all the literature, we see that on perceptual implicit memory tests, reading produces greater priming than does generating in most cases (al-
though this depends on the rule for generating and the test; see Tables 7 and 8). Interestingly, most of these outcomes occur under conditions in which generating produces much better performance than does reading in recognition or recall. The only occurrence of a positive generation effect in perceptual implicit memory tests of which we are aware occurs in Gardiner’s (1988) work in which generation is accomplished from the same fragment on which subjects are tested, a result predicted from the transfer appropriate processing framework. However, there are clearly exceptions to the finding that reading is better than generating, both in the case of generation from conceptual clues (Masson and MacLeod, 1992; Khan’s unpublished experiments) and generation from letter displays (Schwartz, 1989; Weldon, 1991). There are several approaches to understanding these occasional discrepancies. One is to appeal to test differences, which seems to work in the case of generation from letter cues, in which superiority of reading to generating occurred on word identification tests but not fragment or stem completion tests. However, this tactic seems unlikely to help in understanding the exceptions produced by Masson and MacLeod (1992), because they used word identification (by other accounts the most data-driven or perceptual test).

Another approach to sorting out these puzzles is to assume that these perceptual implicit memory tests are not purely incidental (or automatic) and may be affected by intentional retrieval processes. Although implicit tests are assumed to measure a relatively incidental (or automatic) form of learning, under some conditions subjects may base responding in part on recollection that is under conscious control (Jacoby et al., 1989, 1993a,b). So, perhaps when priming is equivalent after Generate and Read conditions, performance in the Generate condition has been augmented by conscious recollection. This interpretation cannot be entertained post hoc in these cases without evidence, of course. Jacoby et al. (1993a,b) have conducted several experiments examining performance following Read and Generate conditions where the generation occurred either from anagrams or from conceptual clues. We presented some of these results earlier in this Chapter to illustrate the logic of the process dissociation procedure (see Table 4). On the implicit
test of word stem completion, Read and Generate conditions produced about equivalent priming, in the data shown there. However, when the process dissociation procedure was used to estimate the contributions of the recollective and automatic influences, this apparent null result is shown to be the outcome of two opposing processes that affected performance. Generating produced better performance than did reading in recollection, but reading led to a stronger automatic influence than did generating. Indeed, using responding to New items as the appropriate baseline, we can see that there was no effect of generating words on the automatic component (as, indeed, there is sometimes no effect on perceptual priming following generating; Jacoby, 1983b, Experiment 3; Smith and Branscombe, 1988). This outcome helps us understand the results shown in Table 9, where priming on word stem and word fragment completion were the same following Read and Generate conditions, and why the difference existed between these measures and word identification. Perhaps if the process dissociation procedure were applied to Masson and MacLeod’s (1992) puzzling findings, then they too would be seen to be the result of these two offsetting processes that, for reasons not yet understood, seem to contribute differentially across various implicit tests.

Context effects
The word context carries many connotations in psychology (Davies and Thomson, 1989). We confine ourselves here to manipulations of verbal context in which target items were embedded, with an implicit memory test given later. Many of the experiments directed at the effect of reading or generating information, just reviewed, included a manipulation of context wherein subjects read the generation clue as well as the target. Therefore this section follows naturally from the foregoing one.

Examination of Tables 7 and 8 shows that, for most of the experiments that provided the relevant comparison, reading an isolated word aloud without context produced more priming than reading it embedded within a sentence context. In most cases the difference was statistically significant. In addition, Masson and MacLeod (1992) reported three experiments (8A, 8B, and 9) in which they showed reading a word in isolation to produce more priming on word identification than reading it in a meaningful context. This outcome helps clarify some puzzling research reported by Oliphant (1983).

Oliphant (1983) reported experiments in which subjects were exposed to words in a first phase prior to performing a lexical decision task. If subjects made a lexical decision on the word during the first phase, considerable priming was found on the later test. However, if subjects simply read the word in the instructions during the first part of the experiment, then no priming was found. The context of the original presentation of the word determined whether later priming would be obtained. MacLeod (1989) addressed the same issue by comparing priming on the word fragment completion test following study of words in a list compared to their presentation in a prose passage. He found greater priming from the list presentation, although (unlike Oliphant) he did obtain some priming when the words appeared in the passage, at least in two experiments. In addition, MacLeod obtained greater priming when the words embedded in the prose passage did not make sense and had to be crossed out than when they fit sensibly into the text. MacLeod concluded that ‘. . . context plays a critical role in priming: As a word moves from being contextually bound in meaningful discourse to being isolated in a list, its probability of priming increases’ (1989, p. 398).

Levy and Kirsner (1989) reported interesting results also bearing on the issue of context effects in implicit tests. Like Oliphant (1983), they showed no priming from prior reading of words in text on a word identification test, although presentation of words in a list did cause priming. However, in other experiments they showed that when the transfer task was rereading texts, prior reading of text produced quite reliable transfer that also held some trademarks of data-driven or perceptual processing (i.e., prior presentation of the passage auditorily did not lead to as much benefit in rereading speed as did prior visual presentation). Levy and Kirsner suggested that it is difficult to interpret transfer or prim-
ing measures 'when the original and reprocessing tasks are at different linguistic levels' (1989, p. 407). Reading words in passages does produce priming (or positive transfer), unlike the original claim by Oliphant (1983), but it is necessary for the test phase to permit appropriate transfer by being a similar sort of experience.

As mentioned previously, Madigan et al. (1991) did find priming from words embedded in instructions given just prior to completing fragmented words. However, in this case the retention interval between exposure and test of the words was much shorter than in Oliphant's (1983) experiment (and the dependent measure was different, too). These may be reasons why the outcomes differ.

Organization
A robust finding in free recall is that a list of items belonging to common categories (some types of furniture, some animals, etc.) is better recalled if the items are blocked by category during study than if they are randomly arranged (Cofer, Bruce, and Reicher, 1966). There is some controversy as to whether this effect of organization of materials, called the blocked/random effect, also occurs in recognition (e.g., Kintsch, 1968; Neely and Balota, 1981). Because organization is a quintessential conceptual variable, one might expect it to affect implicit conceptual tests but not perceptual tests. The evidence is sparse, but thus far is consistent with this prediction. Rappold and Hashtroudi (1991) presented a categorized list with words either blocked according to category or randomly arranged and showed that this variable had no effect on priming in a word identification test but did affect priming in a category production test (as it did in free recall and cued recall with category names). Blaxton (1992) replicated the blocked/random effect in primed category production (and in cued recall).

In another experiment Rappold and Hashtroudi (1991) gave subjects a randomized list of categorized items either with instructions to organize them or not. The group instructed to organize the words showed greater priming in the category production test as well as enhanced free recall. Thus the data reported so far show that organization affects category association tests, but not word identification. However, this variable has not been examined in many experiments or with many implicit tests, so further work is clearly warranted.

Summary of the section
To summarize the evidence from this section, on perceptual implicit tests the following instructional encoding variables have little or no effect: intention to learn; divided attention; directed forgetting; levels of processing; and organization. Directed forgetting also seems to have little effect on conceptual implicit tests (although little evidence yet exists on this issue), but levels of processing and organization of materials do positively affect conceptual tests. The variables of reading and generating words, or of reading them in various contexts, have produced less consistent findings. Generating words generally has positive effects on priming on conceptual tests relative to reading them out of context, just as occurs on explicit conceptual tests. Conversely, in the great majority of experiments, reading words produces more priming on perceptual implicit tests than does generating words. However, there are clear exceptions where reading and generating produce equivalent priming. Context also has variable effects. In general, it seems that when a target becomes submerged in a larger context, priming of it will be reduced if the test involves single, degraded items.

Retention interval
In one of the first experiments conducted with normal subjects on an implicit memory test, Jacoby and Dallas (1981, Experiment 5) examined the effects of retention interval on priming in a word identification paradigm and obtained no reliable decrease from an immediate test to tests delayed 15 minutes or 24 hours (although proportion of completions dropped from 0.73 to 0.72 to 0.67 over these delays, compared to a nonstudied baseline of 0.50). Recognition memory did drop reliably over these retention intervals, so this variable seemed to
dissociate recognition and primed word identification, at least over 24 h. This finding helped to excite interest in the study of implicit memory, especially when Tulving et al. (1982) seemed to replicate and extend it the next year.

Tulving et al. (1982) reported that primed word fragment completion was especially resistant to loss over time. They presented subjects long lists of words and then tested recognition and word fragment completion 1 hour and 1 week after study. Recognition memory scores showed a sharp decline over the week, but priming on the fragment completion test dropped only 3%, from 17% to 14%. Thus, retention interval seemed to dissociate explicit and implicit tests, having huge effects on the former and negligible effects on the latter.

Probably because of these experiments by Jacoby and Dallas (1981) and Tulving et al. (1982), many researchers have included retention interval in their experiments on implicit memory tests. The literature is large and the evidence is somewhat contradictory. However, it seems safe to say that any claim of implicit memory measures being especially resistant to forgetting has been conclusively disconfirmed. Only one published report of many has replicated the complete resistance of primed word fragment completion to forgetting. Komatsu and Ohta (1984, Experiment 2) reported that priming from presentation of a word dropped only 3% (from 0.20 to 0.17) from 1 minute to 1 week after study. However, virtually every other experiment shows significant losses over a week. Without going into exhaustive detail, in a direct replication of the original Tulving et al. (1982) experiment (reported in Sloman, Hayman, Ohta et al., 1988), priming dropped from 0.22 to 0.16 (more than a 25% proportional loss from the original level, i.e., \(0.22 - 0.16\)/0.22 = 0.27). Using this same measure of proportional loss (to take different baseline and priming rates into account), Light, Singh and Capps (1986) found losses of 41% and 47% for young and old subjects, respectively, over a one week interval; Chandler (1983; reported in Sloman et al., 1988) found losses of 68% and 50% for two different list lengths; and Roediger and Blaxton (1987a) reported a 52% drop. Analyzed this way, Tulving et al. (1982) originally found a 17% loss and Komatsu and Ohta (1984) a 15% drop, so it is clear that this proportional loss score dramatizes the forgetting rate; however, priming is often rather low to begin with in some studies, and therefore little loss over time can be expected. The proper measure of forgetting is still debated (e.g., Loftus, 1985a,b; Slamecka, 1985), but whether one examines absolute or proportional losses, most experiments show loss of priming over time in word fragment completion.

The most complete attempt to come to grips with losses of priming over time in the word fragment completion task was published by Sloman et al. (1988), who reported six new experiments and briefly presented results of several unpublished studies. They showed very rapid losses of priming over short delays (less than 5 min since study of a word) and then more gradual decay thereafter. However, priming lasted for months in some of their experiments and over a year in one. Most researchers who have looked have found reliable priming after at least a week. Hashtroudi et al. (1988) reported similar long-lasting priming with the measure of partial word identification. We should stop for a moment to marvel at the fact that a single presentation of a word can lead to priming that certainly lasts a week and may last months.

Researchers employing word stem completion as a measure of implicit memory have also examined retention interval and many report a very different pattern. In this task, complete loss of priming is often found after two hours or less (e.g., Chen and Squire, 1990; Diamond and Rozin, 1984; Graf and Mandler, 1984; Mandler et al., 1986). Coupled with the long-lasting priming in word fragment completion, this fact led Graf and Mandler (1984; see also Squire, Shimamura and Graf, 1987) to the conclusion that word stem completion measures automatic activation or priming more accurately than does word fragment completion and that the apparent long-lasting 'priming' in fragment completion was really due to intentional recall. However, the experiments measuring priming on the two types of tests differed on many other variables, such as fre-
quency or familiarity of the words used. Researchers using word stem completion often use short, familiar words (so as to have ten completions from the stem), whereas those employing word fragment completion often use long, less familiar words (to obtain unique fragments). Roediger et al. (1992) directly compared loss of priming on word fragment and word stem completion in the same experiment with the same materials and found little difference between the measures. They obtained gradual declines in priming from an immediate test to tests coming 1.5, 2, 48 and 168 hours after study, both on word stem and word fragment completion. Priming on both tests existed one week after study. Craik et al. (1993) also directly compared word fragment and word stem completion relatively soon after study or a day later and found reliable priming on both tests at both intervals. Interestingly, priming on neither test dropped over the 24-hour retention interval.

There are miscellaneous reports of losses of priming on other implicit memory tasks, but only word fragment and word stem completion have been studied extensively enough to warrant careful review here. However, for the sake of completeness we list a few other findings with various tests. Mitchell and Brown (1988) showed stable repetition priming in picture naming between 1 and 6 weeks, whereas recognition memory for the pictures declined over this interval. Durso and Johnson (1979) also found stable repetition priming in picture naming, but over minutes rather than weeks. However, Lachman and Lachman (1980) appeared to find a decline in primed picture naming from 1 to 7 to 14 days, although no inferential statistics were reported. Attempts to replicate Mitchell and Brown (1988) would be worthwhile, because their data seemed to show a decline between 1 and 4 weeks, but not 1 and 6 weeks (i.e., priming was greater after 6 weeks than after 4 weeks). On a different task, McAndrews and Moscovitch (1990) found priming in anagram solution after 2 weeks and Roediger et al. (1992) obtained reliable priming on this test after 2 and 28 hours. Using the lexical decision task, researchers have been interested in the durability of repetition priming as a function of lag (the number of intervening words and nonwords) within a session. Repetition priming has proved remarkably durable over lags of 16 (Dannenbring and Briand, 1982), 29 (Moscovitch, 1985), 31 (Scarborough et al., 1977, in most of their experiments) and even 84 intervening items (Forbach, Stanners and Hochhaus, 1974). Curiously, we can find only two experiments in which repetition priming in the lexical decision task was measured over intervals longer than a few minutes. Scarborough et al. (1977, Experiment 4) found reliable repetition priming for words (but not nonwords) over a 2-day interval, and Connor, Balota and Neely (1992, Experiment 3) showed a similar effect for very low frequency words after a week.

Turning to conceptual implicit tests, Hamann (1990) showed that priming on the category association test declined rather rapidly over 90 min and that the decline was slower following semantic processing of the targets at study than following physical processing. However, Rappold and Hashtroudi (1991) did find reliable priming on a category production test after 24 hours and (unlike Hamann) they did not find that a variable (blocked presentation of category exemplars at study) that affected immediate priming levels also enhanced priming a day later.

Indirect (or semantic) priming from related words in tasks such as lexical decision or word naming is evanescent, disappearing if only one or two items are presented between the priming and test events (e.g., Joordens and Besner, 1992). The results in the foregoing Section make it clear that direct priming and repetition priming have a different basis from indirect (semantic) priming, because of their longevity. Direct priming is often obtained over hours, days, weeks, and sometimes even months in virtually all implicit memory tests. From this observation, it is clear that the notion that priming is due to temporary activation of abstract representations such as logogens (Morton, 1969) or pictogens (Warren and Morton, 1982) simply does not apply to priming in these implicit memory tasks. Direct priming in these tests may not be as completely resistant to forgetting as was once believed (Jacoby and
Dallas, 1981; Tulving et al., 1982), but it is still remarkably persistent over long intervals, relative to semantic priming.

Miscellaneous variables

We could not squeeze three variables into the organization of our review, yet each deserves brief attention. The variables are serial position, interference, and the proportion overlap of studied items on implicit memory tests (i.e., the ratio of studied items to the total number tested). The first two are classic memory variables that have as yet received little attention in the implicit memory literature.

Serial position

Procedures in implicit memory tests often include study of long lists and a lengthy interval between study and test. Perhaps it is these features that have encouraged researchers from examining serial position curves even in the more standard implicit memory tests. The existence of a recency effect in primed word fragment completion is strongly implied by Sloman et al.’s (1988, Experiment 1) findings, described earlier, in which they showed rapid loss of priming in word fragment completion over the initial few minutes after study. After all, one way of thinking about a recency effect in a serial position curve is as a forgetting function (albeit plotted in the opposite direction from the usual way) — performance declines with delay from study, with the most recent item the best produced. In the one experiment to examine recency per se, McKenzie and Humphreys (1991) did indeed find recency effects on implicit tests with part word cues (beginning and ending word stems). Recency also occurred in explicit test conditions with the same types of cues. They also examined recency with conceptual cues (associates of the targets) on both implicit and explicit tests, but found little evidence for it on the implicit word association test.

As far as we can tell, the voluminous implicit memory literature is almost mute on the issue as to whether a primacy effect exists on any implicit memory test. The only exception to this generalization is represented by Sloman et al.’s (1988) finding that a single item primacy effect appeared in three replications of their Experiment 1. Although examination of serial position effects has guided research about other issues in the study of memory, work on implicit memory is a clear exception at this point in time.

Interference

Both proactive and retroactive interference exert powerful effects on retention in many explicit memory tests (Postman, 1976; Underwood, 1957). Therefore, heightened interest should be directed to claims that interference has little or no effect on priming in implicit memory tests. For example, Jacoby (1983a) presented subjects with a different word list on each of five successive days and tested for word identification. Although he found healthy priming effects, he obtained no evidence of either proactive or retroactive interference.

In a related vein, Sloman et al. (1988, Experiment 5) had two groups of subjects study 25 words and then perform either a verbal or nonverbal interpolated task for 20 min before receiving a word fragment completion test. Subjects in the nonverbal distractor condition played a video game for 20 min, whereas those in the verbal condition studied and recalled 12 lists of 12 words. The verbal interpolated condition might be expected to yield greater interference in retention of the original word list, but no reliable difference was found (although priming was 4% greater in the nonverbal than the verbal condition). However, in an explicit recognition test (two alternative forced choice), there was also little evidence for retroactive interference — a 5% difference in hit rates favoring the nonverbal condition. Of course, many studies show little RI in recognition relative to recall (Underwood and Postman, 1973). Therefore, lack of interference on implicit tests may not indicate any fundamental difference between explicit and implicit memory tests, but rather that interference is greater on tests without strong cues. Sloman et al. (1988, p. 238) left ‘open the possibility that retroactive interference ef-
Effects may occur in primed fragment completion, but they may be revealed only in experiments that are much more sensitive than ours.

Nelson et al. (1989, Experiments 2 and 3) did indeed report retroactive interference in a word completion task the next year, but the task involved presenting subjects word endings and asking them to produce a response under either explicit or implicit test instructions. Nelson et al. (1989) found RI when the interfering words were lexically related to the target list (i.e., had similar letter patterns and/or rhymed with the target set), but not when they were similar in meaning. This pattern fits with Nelson’s view that a critical component in word fragment tasks is lexical access or a search of lexical memory. Interestingly, the lexical interference effects appeared on both explicit and implicit versions of the test; again, no evidence was produced that interference occurs in explicit but not implicit tests. However, we note in passing that the Nelson et al. (1989) results are different from results of a pilot experiment by Sloman et al. (1988, pp. 232–233). They reported no difference in RI on a word fragment completion test from interpolated lists that were orthographically similar to the target list relative to interfering lists that were semantically similar. (Sloman et al. did not include explicit tests to see if interference would have occurred on such tests under these conditions).

Graf and Schacter (1987) did report a dissociation between an explicit and implicit version of a test (priming of new associations) as a function of interference. In the priming of new associations, subjects study paired associates (shirt-window) and then later receive a test in which they are told to complete a word stem for the right hand member of the pair (win-______) in the context of another word. The other word can be the original context word from the study list (shirt-win______) or a word studied in another pair (e.g., bottle-win______). Under implicit test instructions subjects are more likely to complete the stem with the target word if the original context word is given at test than if another word is given. In the Graf and Schacter (1987) experiments subjects were placed in A-B, A-D interference conditions in which the same stimulus was paired with different responses across lists. Interestingly, this manipulation produced no effect on implicit versions of the test for the A-B list, but did affect the test when it was given with explicit memory instructions. The RI manipulation also affected performance on other explicit tests, such as pair matching.

Despite speculation that interference effects may differ on explicit and implicit tests, Graf and Schacter’s (1987) demonstration is the only piece of convincing evidence available in the literature. However, they showed this in the implicit tests of memory for new associations, which seem to differ from other tests in various ways, such as not showing intact priming in memory-impaired subjects (Shimamura and Squire, 1989) and requiring elaborative processing at study to produce any priming at all (Graf and Schacter, 1986). It remains to be seen if absence of interference can be demonstrated on other implicit tests under conditions that satisfy the retrieval intentionality criterion (as did Graf and Schacter’s (1987) experiments).

An issue indirectly related to the issue of interference is that of list length. In explicit memory tests, longer lists are associated with worse recall or recognition of particular items appearing in the list (e.g., Murdock, 1960). The same holds true in primed word fragment completion, too, as Chandler (1983) and Sloman et al. (1988, Experiment 5) have shown. In Chandler’s (1983) experiments, subjects studied lists of 12 or 48 words and took tests at two retention intervals. Collapsing over retention interval, she found priming effects of 0.30 in the short list and 0.08 in the long list, which is rather unnerving in that many researchers have used long lists and we advocate the practice in our recommendations for implicit memory test procedures given above. At any rate, for present purposes the point is that list length effects are often attributed to interference, although other variables (such as delay between study and test) are mixed in. If list length effects are obtained with other measures of performance, then experiments will be needed to tease apart effects of interfering items (input and output
interference) from delay (e.g., Tulving and Ar- 
buckle, 1963).

Proportion overlap
Jacoby (1983a, Experiment 1) performed a word identification experiment in which the proportion of studied words appearing on the test was varied. In one condition, only 10% of studied words appeared on the test (i.e., the remainder were nonstudied) whereas in another condition 90% of the test items had been studied. The priming effect was larger when the test contained a large proportion of studied items. Allen and Jacoby (1990) replicated this finding and provided evidence that it was not mediated by intentional uses of memory. (However, the same sort of effect also occurs in recognition memory experiments; see Jacoby, 1972, and Todres and Watkins, 1981).

Challis and Roediger (1993) attempted to extend the finding of a proportion overlap effect to primed word fragment completion. Subjects studied 80 words and then took a word fragment completion test in which the proportion of items overlapping between study and test varied (for different groups) across the range of 0 to 100% in 25% increments. After subjects took an initial test (80 fragments) they received exactly the same test again. Challis and Roediger found no effect of proportion overlap on either the first or second test. Although performance on the second test was better than on the first, priming was no greater; that is, subjects completed more fragments on the second test than on the first, but equally so for studied and nonstudied items.

Why did Jacoby (1983a) and Allen and Jacoby (1990) find robust proportion overlap effects, but Challis and Roediger (1993) find none? Without further work there is no sure answer to this question. However, one difference in the procedures (besides the difference in the type of test) was that subjects in Jacoby's experiments were told, in the high proportion overlap condition, that many studied words would indeed appear on the test, whereas subjects in the low overlap condition were not informed about any relation between study and test. However, Challis and Roediger (1993) made no mention of the independent variable to their subjects in any of the overlap conditions. Only further work can say whether this difference is critical; it may well not be, because Allen and Jacoby (1990) provided evidence against contamination by intentional uses of memory as causing their proportion overlap effect.

Theoretical approaches for explaining test dissociations
Prior reviews of the literature have identified three primary approaches towards explaining implicit memory phenomena and their frequent dissociation from explicit measures of retention (Schacter, 1987): the transfer appropriate processing approach; theories proposing memory systems; and dual process theories postulating activation and elaboration as critical constructs. We cannot present the theories in their full form here, but we sketch in a few observations about these three approaches, their strengths and weaknesses, and point out how the theories may be converging into one general theory of this field. One newer approach is considered briefly, but we cannot take the space to consider a variety of other new ideas (bearing various shades of resemblance to those above) in this review.

Processing theories
This entire review has been oriented around one version of the processing approach to explaining dissociations between explicit and implicit memory tests (Roediger, 1990a). The theory obviously can account for many of the test differences appearing in the literature. In particular, the theory predicts many of the findings from manipulation of surface features or symbolic forms (variations of modality, picture/word processing, and language in bilinguals) on tests. These variables have large effects on perceptual implicit tests and either no effect or opposite effects on conceptual tests, whether explicit or implicit. Similarly, the bulk of the findings from manipulations of processing strategies (levels of processing, etc.) are accounted for quite
reasonably. We need not rehearse here the supportive findings.

Some of the advantages of the transfer appropriate processing approach are considered here. First, the approach has been fruitful in providing a program of research and identifying relevant variables to examine. The experiments conducted within this framework have provided a wealth of information about the relations among various tests. Second, and related, this approach has urged researchers to use multiple implicit and explicit tests to assess explicit and implicit memory, rather than relying on a single test of each type (although studies using only two tests still dominate the literature). Third, the approach has provided a set of converging operations to identify the primary theoretical constructs (perceptual and conceptual tests). Although some have argued that the approach is circular (Tenpenny and Shoben, 1992), we disagree with this criticism because the set of converging operations for identifying tests has been specified. (This has rarely been the case with the other theoretical approaches.) Fourth, because of the second and third features here, the approach makes specific predictions about performance on many different tests (not all of which are satisfied, of course). Fifth, the approach emphasizes interaction of observable variables — the relation between study conditions and tests, primarily — in a direct way without being caught up in labyrinthine mediationistic accounts of performance (Watkins, 1990).

Drawbacks of the approach include the following. First, the theory emphasizes differences in types of processing, but does not explain well pure differences in intentional versus incidental retrieval. These occur in amnesic patients (who are deficient at intentional retrieval and intact on incidental retrieval) and in normal subjects (when dissociations occur in tests when only test instructions are manipulated; Graf and Mandler, 1984; Roediger et al., 1992). Some steps have been taken to overcome this deficiency, by suggesting that conceptual processes are more engaged in explicit than implicit tests with data-limited cues (Roediger et al., 1989); by extending the transfer appropriate processing approach to account for patterns of performance in memory-impaired patients (Blaxton, 1992); and by suggesting that Jacoby, Kelley and Dywan's (1989) memory attributions approach can be melded with the processing ideas to provide an account of intentional remembering and its dissociations from incidental retrieval (Roediger, Rajaram and Srinivas, 1990). However, it remains the case that the transfer appropriate processing approach has had virtually nothing to say about the important issue of consciousness in intentional and incidental retrieval (Gardiner and Java, 1993).

Second, the approach obviously has difficulty explaining some findings that we have reviewed, particularly those of Masson and MacLeod (1992), Smith (1991), Hunt and Toth (1990), and Tenpenny and Shoben (1992). The basic problem pointed out in the last two papers is that parallel effects and dissociations that are not predicted by the theory can exist between tests. Hunt and Toth (1992) found that a perceptual variable (orthographic distinctiveness) could affect both priming on a perceptual implicit test and performance on a conceptual explicit test (free recall). Tenpenny and Shoben (1992) found dissociations between two conceptual tests. How can we reconcile these findings with the large body of evidence consistent with the processing approach? We think the answer lies in the fact that a distinction between perceptual and conceptual processing is likely too rough and that more fine-grained distinctions are needed. That is, just as we can distinguish different sorts of perceptual processes and use them to make principled predictions on implicit perceptual tests (as in Weldon and Roediger, 1987, Experiment 4), we need to make finer grained distinctions in the conceptual domain, too. We need to distinguish further between types of perceptual processing and types of conceptual processing. We explore this idea below when we consider the componential approach.

Memory systems theories

Some of the early observations that led to the study of implicit memory tests arose from neuropsycho-
logical studies of amnesic patients. Brain damage from various causes rendered these people incapable (or at least much less capable than normal subjects) of remembering events from their lives, but they generally showed intact priming on a number of implicit memory tests (Moscovitch et al., this Volume; Schacter, 1987; Shimamura, 1986; Warrington and Weiskrantz, 1970). The natural interpretation of this state of affairs is that there are distinct memory systems in the brain, that these systems provide the neural underpinnings for performance on various memory tests, and that the systems can be selectively impaired. This general view is often impressively backed by experimental work with animals that reveals different neural systems involved in different memory tasks (e.g., Mishin and Appenzeller, 1987).

Cohen and Squire (1980) proposed a distinction between declarative and procedural memory, the former underlying both remembrance of events (episodic memory) and knowledge of facts (semantic memory) and the latter responsible for carrying out cognitive and motor procedures. In the original proposal (see also Squire, 1987), the declarative memory system was considered impaired in amnesic patients, whereas the procedural system, responsible for priming and other forms of learning, was intact.

Tulving (1983; see also Tulving, 1972) distinguished between episodic and semantic memory (memory for personal episodes and for general knowledge, respectively); brain damage that affected recall and recognition was attributed to damage to the episodic system; preserved priming was attributed to an intact semantic memory system. Tulving (1985) later proposed a mono-hierarchical arrangement of systems that included procedural memory as the most basic system, underlying performance on the other two (i.e., semantic and episodic memory). Semantic memory depends on the procedural system. The episodic system is a yet more developed system for remembering personal experiences, which is dependent on the other two systems. The primary differences between Tulving's (1985a) proposal and that of Squire (1987) are that (a) Squire's model proposes systems working in parallel, whereas in Tulving's view they are dependent on one another, and (b) semantic memory is impaired in amnesic patients according to Squire (who attributes semantic priming to a subsystem of the procedural or nondeclarative memory system), whereas Tulving sees semantic memory as generally intact in such patients. Semantic learning is not normal in amnesics, according to Tulving, because episodic memory is impaired and must be used to acquire new facts (see Tulving, Hayman and Macdonald, 1991). Several other proposals in the 1970s and 1980s proposed two different memory systems to account for dissociations being found both in humans and in other animals (see Schacter, 1987).

These original proposals emphasized that there were a small number of memory systems — the above authors and others started out with two — and that criteria for proposing systems were fairly tight (see Sherry and Schacter, 1987; Tulving, 1983). Since the early 1980s there has been an explosion in the number of memory systems (or subsystems) postulated, because increasing numbers of dissociations between tests have been shown in both normal and pathological populations (e.g., Blaxton, 1989). The latest versions of these theories provide at least a half dozen systems (e.g., Squire et al., 1992; Tulving and Schacter, 1992; Weiskrantz, 1985) and by using the same sort of criteria, the count can plausibly be raised to 25 or so (Roediger, 1990b). Not surprisingly, with enough memory systems, any pattern of performance appearing in data of normal or brain-damaged subjects can be explained. The rather strict criteria for postulating memory systems recommended by Sherry and Schacter (1987) have never taken hold. The current state of affairs leaves the field in ferment, which is healthy, but makes testability of the theories rather beside the point. Patterns of results interpreted as contradictory to one proposal with a small number of memory systems, such as Blaxton's (1989) experiments, can always be accommodated by proposing new systems.
(although in this case evidence from patients also converges on the newly proposed systems; see Schacter, 1990).

There has also been a tendency to shift responsibility for tasks among the various proposed memory systems. The word fragment completion task first popularized by Tulving et al. (1982) was then, in 1982, ‘clearly a semantic memory task’ (p. 341). This was because semantic memory tasks were then defined as tasks for which episodic memory was not needed and ‘for which general pre-experimental knowledge of the world is sufficient’ (Tulving, 1983b, p. 77). But even in 1982, Tulving et al. noted that ‘it is not clear that [priming effects in this task] can be regarded as a phenomenon of semantic memory’. Instead, they ‘were tempted to think that [the priming effects] reflect the operation of some other, as yet little understood, memory system’ (quotes from p. 341). Eight years later, priming on this same word fragment completion task was now taken (by two of the original authors) to reflect output from one of several perceptual representation systems, the visual word form system (Tulving and Schacter, 1990). Representations in this system are said to be ‘presemantic’, and the test is not thought to reflect output from semantic memory at all.

All scientific ideas undergo change, and in this area the changes have occurred at a fast rate. These changes have made the idea of there being only two (or even a handful of) memory systems obsolete. In our opinion, the changes in these approaches to postulate increasing numbers of memory systems (see Schacter, 1992; Squire et al., 1992; Tulving and Schacter, 1990; Weiskrantz, 1985) have made these ideas more similar to those emphasizing mental procedures or processes. The procedural views showed that dissociations could be found on tests on many bases besides those included in the original memory systems theories, with only two or three systems (Blaxton, 1989; Kolers and Roediger, 1984). These processing ideas were originally seen as competitors to the systems views, but can now be viewed as complementary (Schacter, 1990, 1992).

One advantage of memory systems theories is that they relate performance on behavioral tasks to underlying neural structures and processes. They thus provide more natural explanations of performance in memory-impaired patients and also provide a link to work with similar preparations in other animals. In addition, they frequently provide a plausible account of performance in normal subjects, too, when dissociations occur between tests, especially when the tests are equated on other dimensions (Graf and Mandler, 1984; Roediger et al., 1992; Schacter et al., 1989). Another advantage of this approach is that some theories provide specific statements about the relation between states of consciousness and performance based on the different systems (see Tulving, 1985b).

Drawbacks of the approach include the complaint that much of the evidence depends on contrasting performance on two measures of performance (an explicit test thought to reflect one system and an implicit test thought to reflect another) when these tasks often differ in many other ways (see Neely, 1989; Roediger, 1984). As previously remarked, most of the literature on test differences is, unfortunately, marked by contrasts between only two tests. Converging tests are needed for each of the systems postulated and attempts to provide such converging operations for memory systems views have not been particularly successful in some cases (Blaxton, 1989). A second problem, as already mentioned, is that the increasing number of memory systems postulated makes testability difficult if not altogether beside the point. With the increasing number of systems, this approach resembles the processing approach in emphasizing the match between the type of information provided to subjects at study and the information processing requirements (and underlying neural structures) engaged by the test. Roediger and Srinivas (1993) have noted that if one defines a memory system as ‘whatever neural structures underlie performance on a particular test’, then there is no distinction to be made between processing and systems theories. Whenever a dissociation occurs between two behavioral measures of learning, there must be concomitant neural changes causing the behavioral effects. Therefore, in this limited sense, two different
neural systems must be active. However, the implication from the numerous dissociations now available in the literature is that the number of memory systems in this sense is quite large (Roediger, 1990b).

Roediger (1993) has noted that memory is defined quite broadly in the study of implicit memory tests, so that virtually all aftereffects of experience that appear in later behavior are considered manifestations of memory. If this is so, then most biological systems have memories associated with them. The muscular system, the respiratory system, and the immune system all show effects of past experience, or priming. Roediger (1993) even presented evidence that the female reproductive system is a memory system, because it learns to produce babies faster when women repeatedly go into labor (i.e., there is priming caused by the first experience of labor that speeds subsequent labors). Given these considerations, priming on perceptual tests of implicit memory may then be seen as priming in another set of biological systems, the perceptual systems. In short, there can be no quarrel at some level that dissociations between performance on various tests and indices of behavior reflect different memory systems; however, starting with this assumption indicates that there will be a very large number of such systems, and not just some 2 – 5 as current theories endorse.

Activation and elaboration

A third primary approach identified in Schacter’s (1987) review is one advocated by Mandler and his colleagues (e.g., Mandler, 1980, 1989; see also Graf and Mandler, 1984), which distinguishes between processes of activation and elaboration. Briefly, the idea is that studied events (such as the presentation of words in lists) activate their preexisting representations for a relatively short period. This activation is responsible for priming, and in many theories the activation is thought of as a general (amodal) sort, of a logogen (Morton, 1969), or some other form of knowledge structure. Activation occurs automatically, but in addition subjects may elaborate on information (form associations, process it meaningfully, form images, etc.). The distinction between activation and elaboration can account for many dissociations in the literature, such as the fact that level of processing generally has little effect on perceptual priming, but does affect standard explicit tests. Activation of knowledge occurs under all levels of processing and thereby supports priming, whereas elaboration is responsible for coding events in such a way that they can be recognized and recalled later.

The notion that amodal activation is responsible for priming has foundered on many empirical observations. First, numerous studies reviewed above show that surface features of materials affect priming (modality, typography, and so on), and therefore any activated representation is clearly not amodal. In addition, priming in all tasks lasts much longer than can be reasonably explained by a temporary activation of knowledge structures, unless one stretches the notion of ‘temporary activation’ to great lengths. However, in later versions of this view, the notion of amodal activated representations has been abandoned and, in general, the distinction has come to resemble the transfer appropriate processing approach by emphasizing modality-specific representations that can then be elaborated (Mandler, 1989). This trend can be most clearly seen in Graf and Ryan (1990), who even refer to their account as a transfer appropriate processing view. Once again, we see that the approaches to explaining dissociations between explicit and implicit tests seem to be converging. If we were to add to this Section, our comments would be quite similar to those for the transfer appropriate processing approach described above, with similar advantages and drawbacks.

Components of processing

A new approach that has been advocated by Moscovitch et al. (Chapter 3, this Volume; see also Moscovitch, 1992) endorses the notion of the cognitive systems being composed of numerous components that may be assembled for use in a given
task. Rather than thinking in terms of independent memory systems or one or two different types of processing (perceptual and conceptual), the components of processing approach represents a middle ground. As Witherspoon and Moscovitch (1989, p. 29) express the idea, this approach is '... based on the assumption that performance on each task requires the operation of many components, some of which are common to tasks and some of which are not. Performances on each task may be independent from each other to the extent that their components differ (or the information they use is different), leaving open the possibility that some components (or types of information) may be more critical in this regard than others. Witherspoon and Moscovitch (1990) observed stochastic independence between priming on two perceptual implicit memory tests and argued that the two must have some component that is not in common (see also McAndrews and Moscovitch, 1990). Tenpenny and Shoben (1992) endorsed the same idea to explain why they found dissociations between two conceptual tests — they must have some component process or processes that are not shared. Hintzman (1990) has made a similar point.

The advantages of the processing components view are, in some ways, similar to those of the transfer appropriate processing view of Roediger (1990) and others, which it incorporates. Dissociations are explained by the match between processing components used at study and at test. Moscovitch et al. (Chapter 3, this Volume) argue that the distinction between perceptual and conceptual components represents two major components of processing, whereas Tenpenny and Shoben (1992) find little use for the distinction (although they endorse the componential approach). Not surprisingly, we side with Moscovitch et al. on this issue and think it would be difficult to explain the great body of evidence we have reviewed here without introducing some notion of a perceptual/conceptual contrast, both to explain performance of normal subjects and that of memory-impaired patients (Blaxton, 1992; Schacter, 1990; Tulving and Schacter, 1990). We readily concede, however, that this distinction is too gross to capture all dissociations that have been observed, as we noted above.

Another advantage of the processing components approach is that it provides a natural way to explain behavior of brain-damaged patients with specific deficits. Processing components are similar to the processing modules endorsed in most of cognitive neuropsychology; they can be thought of as mini (many) memory systems, with the proviso that there are not a few such systems, but scads of them.

These considerations point up the main difficulty of this new approach: its vagueness. Unless one can identify what the components are, how many and in what configuration they occur in a given task, and their importance to the task, the approach devolves into a theory that says that ‘something complex happens during task performance.’ The early memory systems approaches endorsing a small number of systems, and the transfer appropriate processing approach endorsing two well-specified forms of processing, at least had the potential of being testable, and we have seen that many findings are problematic for both approaches. For proponents of the components of processing approach the task will be to make the theories specific enough to be testable in any form. Still, the approach is congenial to both processing and structural theories and well worth developing as potentially embodying the strengths of both — if it can dispense with their weaknesses.

Concluding remarks

Patient readers still with us may think that we have included almost every conceivable study and topic under the rubric of implicit tests of memory in normal subjects. As we finish our review we are, curiously, afflicted by an opposite reaction — we worry about how much we have left out. There are some studies we could not fit into our organization, some we doubtless were unaware of, and theories that are interesting but did not form the basis for our review (e.g., Humphreys, Bain and Pike, 1989; Nelson et al., 1992). In addition, every journal that arrives seems to carry three or four more papers relevant to our review and yet we must leave off some-
Implicit memory in normal humans  Ch. 2

where, even at the risk of our conclusions about various topics quickly being rendered obsolete by
the rush of new evidence.

Greene (1992) attempted a chapter on the emerging experimental paradigm of implicit memory in
his book on paradigms in the field of human memory. Our reaction is rather like his, so we bor-
row his words. Greene noted that for many well established paradigms it ‘seems unlikely that our
basic beliefs about them will be easily overthrown’ in a short period of time. ‘However, if the research
that one is discussing is relatively new, still rather poorly understood, and being studied by a shock-
ingly large proportion of researchers in the field, one can have no such confidence. Such is the case when
one writes about implicit memory,’ on which research ‘is proceeding at a feverish pace. I doubt
that any researcher will deny that it is by far the most heavily studied area in cognitive psychology at pre-
sent’ (Greene, 1992, p. 172).

Greene (1992) is clearly correct about the rate of research on this topic, even if we believe that it is
likely more accurate to think of paradigms rather than a paradigm, for the study of implicit memory.
This review has concentrated on what is perhaps the dominant paradigm in this area, the study of prim-
ing, which of course arises in many different forms and on many different tasks. Despite the difficulties
engendered in writing about so swiftly a changing field, we certainly prefer its rapid change to the alter-
atives we can envisage. Research conducted under the rubric of implicit memory tests seems likely
to occupy many researchers’ thoughts and labora-
tories for the foreseeable future.

Acknowledgements

Preparation of this review was supported by Grant 91-0253 and Grant F49620-92-J-0437 from the Air
Force Office of Scientific Research. We thank Brad Challis, John Gardiner, Jordan Grafman, Melissa
Guynn, Larry Jacoby, Morris Moscovitch, Suparna Rajaram, Dan Schacter, Kavitha Srinivas, and
Endel Tulving for their interesting comments on earlier drafts of this chapter. Correspondence may
be addressed to the authors at the Department of Psychology, P.O. Box 1892, Rice University,
Houston, TX 77251-1892, or by electronic mail to roedige @ricevm1.rice.edu or mcdermo @ricevm1.
rice.edu.

References

Allen SW, Jacoby LL: Reinstating study context produces uncon-
scious influences of memory. Mem. Cogn.: 18; 270 – 278,
1990.
Balota DA, Neely JH: Test-expectancy and word-frequency ef-
Bartram DJ: The role of visual and semantic codes in object nam-
Basden B, Basden D, Gargano G: Directed forgetting in implicit
and explicit memory tasks: some tests of the retrieval inhibition
explanation. J. Exp. Psychol.: Learn. Mem. Cogn.: in press,
1993.
Bassili JN, Smith MC, MacLeod CM: Auditory and visual word
stem completion: separating data-driven and conceptually-
Bentin S, Moscovitch M: Psychophysiological indices of implicit
1990.
Benzing WE, Squire LR: Preserved learning and memory in amn-
Biederman I: Recognition-by-components: a theory of human
Biederman I, Cooper EE: Priming contour-deleted images: evi-
dence for intermediate representations in visual object recog-
Biederman I, Cooper EE: Evidence for complete translational
and reflectional invariance in visual object priming. Percep-
Biederman I, Cooper EE: Size invariance in visual object prim-
ing. J. Exp. Psychol.: Hum. Percep. Perform.: 18; 121 – 133,
Bjork RA: Theoretical implications of directed forgetting. In
Melton AW, Martin E (Eds), Coding Processes in Human Memory.
Bjork RA: Retrieval inhibition as an adaptive mechanism in
human memory. In Roediger HL, Craik FIM (Eds), Varieties of
Memory and Consciousness: Essays in Honour of Endel
Blaxton TA: Investigating dissociations among memory mea-
sures: support for a transfer appropriate processing frame-
work. J. Exp. Psychol.: Learn. Mem. Cogn.: 15; 657 – 668,
1989.
Blaxton TA: Dissociations among memory measures in memory-
impaired subjects: evidence for a processing account of
Block RA: Effects of instructions to forget in short-term
memory. J. Exp. Psychol.: 89; 1 – 9, 1971.
Bousfield WA: The occurrence of clustering in the recall of ran-

123
Cofer CN, Bruce DR, Reicher GM: Clustering in free recall as a function of certain methodological variations. J. Exp. Psychol.: 71; 858 – 866, 1966.
Ebbinghaus H: Memory: A Contribution to Experimental Psychology. New York: Dover. (Original work published 1885; translated 1913.)
Flexer AJ: The implications of item differences: commentary on


Khan M, Roediger HL: Unpublished data.


Implicit memory in normal humans  Ch. 2


Nairne JS: The mnemonic value of perceptual identification. J.
Paivio A, Yuille JC, Madigan SA: Concreteness, imagery and meaningfulness ratings for 925 nouns. J. Exp. Psychol.
Monogr.: 76 (1, Part 2), 1968.


Slamecka NJ: On comparing rates of forgetting: comment on


