2005

A comparison of techniques for measuring automatic retrieval in conceptual priming

Jennifer C. Major
Wilfrid Laurier University

Running Head: AUTOMATIC CONCEPTUAL PRIMING

A Comparison of Techniques for Measuring Automatic Retrieval in Conceptual Priming

by

Jennifer C. Major

Honours Bachelor of Arts, Memorial University of Newfoundland, 2002

THESIS

Submitted to the Department of Psychology

in partial fulfillment of the requirements for

Master of Science

Wilfrid Laurier University

2004

Your file Votre référence
ISBN: 0-494-04876-X
Our file Notre référence
ISBN: 0-494-04876-X

Abstract

The speeded response technique has provided pure estimates of automatic retrieval in perceptual memory tasks. The present study was designed to investigate whether pure automatic retrieval could also be evidenced in a conceptual task. Subjects were encouraged to generate category exemplars using strictly automatic retrieval by presenting practice trials that did not allow responding with previously studied items and by encouraging speed of responding. This speeded condition was compared to a baseline condition in which conscious retrieval was not possible and to an explicit condition in which conscious retrieval was required. Average RTs in the speeded condition were the same as those in the baseline condition and were faster than those of the explicit condition, supporting the use of an automatic retrieval strategy in the speeded group. Semantic study did not increase target completion rates for the speeded or implicit groups, but it did for the explicit group, suggesting that conscious, but not automatic retrieval benefits from semantic encoding. Estimates of automatic retrieval obtained using PDP formula were identical to automatic estimates obtained using target completion from the speeded group. The idea of using higher frequency target category exemplars in a future study to further clarify present findings is explored.

A Comparison of Techniques for Measuring Automatic Retrieval in Conceptual Priming

The question of whether we can acquire and recall information without the awareness of doing so has become an important question in memory research since Graf and Schacter (1985) first defined implicit memory as a memory phenomenon that operates without conscious or intentional recollection processes. In a typical experiment exploring implicit memory, subjects perform a task under incidental memory instructions with both previously studied and unstudied stimuli, without being informed of the relation between the study and test phases. For instance, in a perceptual implicit stem completion task, subjects are asked to complete 3-letter word stems (e.g., STU_____) with the first word that comes to mind. This procedure contrasts with an explicit stem completion task in which subjects are asked to complete the stems with words they had previously studied. Similarly, in an implicit perceptual identification test, subjects are asked to identify previously studied stimuli (e.g., words) that are presented for very short time intervals without being informed that some of the stimuli have been previously studied. This procedure contrasts with an explicit perceptual identification task in which subjects are instructed to respond with previously studied stimuli.

In contrast to implicit perceptual tests, whereby performance depends to a large extent on the perceptual features of the studied stimuli, performance on conceptual implicit tests depends to a large extent on the semantic properties of the studied stimuli. For example, in a category cued association test, subjects are asked to supply an exemplar - some of which have been previously studied - for a given category label. The primary observation of implicit memory tasks is that they demonstrate repetition priming: In the absence of a goal to recall, people will more frequently recognize or more quickly perceive a stimulus that was encountered some time before the second exposure (Graf & Schacter, 1985). Prior exposure to an item has been found to facilitate the retrieval of studied items on both perceptual and conceptual tasks (Horton, Wilson, & Evans, 2001; Weldon & Coyote, 1996; Vaidya, Keane, Gutierrez-Rivas, Gabrieli, Monti, & Zarella, 1997).

Repetition priming obtained on implicit memory tests has commonly been conceptualized as the automatic retrieval of studied items (Schacter, 1987) and has been described as a relatively pure measure of automatic retrieval in that it is not contaminated by conscious retrieval (Graf & Mandler, 1984), with conscious retrieval defined as the deliberate retrieval of studied items. In contrast, the better recall of studied items in explicit tasks – or in some implicit tasks – has been proposed as evidence for conscious retrieval. Whereas conscious retrieval processes are said to operate with intention and with the use of conscious awareness strategies, automatic retrieval processes are said to operate without intention and without the use of conscious awareness strategies, such as consciously attempting to retrieve items (Bower, 1996). Distinguishing between incidences of pure automatic retrieval - where priming in implicit memory tests is facilitated purely through unintentional automatic memory processes - and contaminated automatic retrieval - where priming is contaminated by an intention to retrieve studied words - has become a prominent concern in the implicit memory literature (Butler & Berry, 2001).

Techniques for obtaining pure automatic estimates

Retrieval intentionality criterion. The retrieval intentionality criterion (RIC) was one of the first approaches designed to detect explicit contamination on implicit memory tests (Schacter, Bowers, & Booker, 1989). The approach demands that explicit and implicit conditions be matched in all respects except for the test instructions. Any differences that then arose in test performance could be attributed to the divergent test instructions. More specifically, proponents of the RIC contended that differential performance between an implicit instruction group from that of an explicit instruction group indicates use of automatic retrieval strategies in the implicit group. Alternatively, equivalent performance by the two groups on a memory test indicates that both groups were using conscious retrieval strategies.

Although Schacter's RIC continues to be implemented in some experiments of implicit memory (Butler & Berry, 2001), it is usually used in conjunction with other procedures designed to detect and control for explicit contamination. Assuming that automatic retrieval is the basis of performance strictly because incidental memory instructions are given is not an assumption that has been accepted by many researchers, who argue that divergent test instructions alone are not enough to control for explicit contamination (Horton et al., 2001; Reingold & Toth, 1996). Even if performance dissociations are obtained between implicit and explicit instruction groups in a memory test, it is not a definitive indication that subjects in the implicit group did not use conscious retrieval strategies (Mace, 2003). Subjects given incidental memory instructions can develop the awareness that studied items can be used to facilitate implicit memory test and subsequently use this awareness to consciously retrieve items. Just as the use of conscious retrieval strategies cannot be inferred solely on the fact that a subject is told that items that are to appear on a memory test have been previously studied - since awareness does not necessitate conscious recollection (Richardson-Klavehn, Gardiner, & Java, 1994; Roediger & McDermott, 1993) – the use of automatic retrieval strategies cannot be inferred solely on the fact that a subject is not told that items on a memory test have been previously studied.

Process dissociation procedure. Another approach that has been used to discriminate automatic and conscious components of memory performance is the process dissociation procedure (PDP) (Jacoby, 1991, 1998; Toth, Reingold, & Jacoby, 1994). "Process pure" measures of automatic retrieval are derived from performance on two different tasks in the test phase of a memory experiment, referred to as inclusion and exclusion tasks. In an inclusion task, subjects are instructed to use a cue to recall a previously studied word. If they cannot recall a previously studied word, they are instructed to respond with the first word that comes to mind. In an exclusion task, subjects are instructed to use the memory test cue to recall a previously studied word, just like on the inclusion task, but then they are to respond with a word other than the studied word. Thus, for both tasks, *direct retrieval* instructions are given, such that subjects are asked to recall an item from a study phase and then respond according to either the inclusion or exclusion instructions.

Performance on inclusion and exclusion tasks is used to derive independent, "process pure" estimates of automatic and conscious retrieval based on two critical assumptions. First, PDP assumes that, although both automatic and conscious retrieval processes are triggered upon seeing a test cue, automatic and conscious retrieval processes operate independently. In other words, conscious and automatic retrieval processes are not correlated. If, however, conscious and automatic retrieval processes are correlated – a notion that has received support in the literature (Curran & Hintzman, 1995; Wilson & Horton, 2002) – then PDP will inaccurately estimate the contribution of conscious and automatic retrieval processes. Secondly, PDP assumes that the conscious awareness that an item has been studied can occur through conscious responding but never through automatic responding (Jacoby, 1998; Reingold & Toth, 1996). In the context of an implicit memory test, this assumption implies that pure automatic retrieval will never produce the awareness that a test item has been studied, and thus, automatic retrieval will never lead to intentional responding. With respect to the exclusion task, this "awareness assumption" (Horton et al., 2001) implies that studied items will be used as responses only when retrieved through automatic processes since there is no conscious awareness of the previous study episode. With respect to the inclusion condition, responding with studied items can occur through either automatic or controlled processes.

The contributions of automatic and conscious retrieval to inclusion and exclusion performance can be quantified using these assumptions. Given that C and A denote the probability that retrieval occurs through conscious and automatic retrieval processes, respectively, the probability of retrieving a studied item on an exclusion or an inclusion task, respectively, is defined as

$$\mathrm{E} = A(1 - C) \tag{1}$$

$$I = C + A(1 - C) \tag{2}$$

Estimates of conscious retrieval can then be defined as

$$C = \mathrm{I} - \mathrm{E} \tag{3}$$

The independence assumption prescribes that

$$A \mid C = A = A \mid \overline{C} \tag{4}$$

Crucially, the estimate of automatic retrieval that is conditionalized on the failure of conscious retrieval is then defined as

$$A = A \mid \overline{C} = \frac{E}{1 - C} \tag{5}$$

Thus, performance on the exclusion task provides a conditional probability of automatic retrieval which is assumed to equal the unconditional probability of automatic retrieval (A).

As with Schacter's RIC, estimates of automatic retrieval derived from Jacoby's PDP are only valid if the assumptions underlying the procedure are valid. The key assumptions that define Jacoby's PDP have proven to be controversial (Curran & Hintzman, 1995, 1997; Richardson-Klavehn, & Gardiner, 1996; Richardson-Klavehn et al., 1994; Wilson & Horton, 2002), with at least some of the controversy yet to be resolved. With regards to the awareness assumption, it has been suggested that automatic retrieval can give rise to conscious awareness (Richardson-Klavehn & Gardiner, 1996; Wilson & Horton, 2002). If this is the case, then this "involuntary conscious awareness" may function to exclude automatically retrieved studied items on the exclusion task, thus leading the PDP to underestimate the occurrence of automatic retrieval (Wilson & Horton, 2002).

Some research also suggests that automatic and conscious retrieval processes do not function independently, but instead are positively correlated (Curran & Hintzman, 1995, 1997; Wilson & Horton, 2002). To reiterate, according to the independence assumption,

$$A \mid C = A = A \mid \overline{C} \qquad (4)$$

But if there is a positive correlation between automatic and conscious retrieval, then the relationship between automatic and conscious retrieval processes is defined as

$$A \mid C > A > A \mid \overline{C} \qquad (6)$$

A more generic, atheoretic definition of automatic retrieval was provided by Buchner, Erdfelder, and Vaterrodt-Plűnnecke (1995) as

$$A = (A \mid C) \times C + (A \mid \overline{C}) \times \overline{C} \qquad (7)$$

Assuming independence, as quantified in equation 4, the use of $A \mid \overline{C}$ to estimate A will result in an increasingly large underestimate of A as the strength of the correlation between automatic and conscious retrieval processes increases and also as C increases.

Recall that direct retrieval instructions are administered to subjects in the PDP group (e.g., Jacoby, 1998; Toth et al., 1994). Thus, subjects are told to use the test cue to recall a studied item for both exclusion and inclusion trials, and then to exclude or include the studied item depending on the cue (i.e., old: include; new: exclude). Given these instructions, it is assumed that subjects only respond with the first word that comes to mind if they are unable to recall a studied word. However, the extent to which subjects in exclusion conditions follow direct retrieval instructions is a matter of debate. If subjects in the exclusion condition do not perform using direct retrieval instructions, either by responding with the first word that comes to mind or by excluding studied words only if a studied word is the *first* word that comes to mind, then subjects are engaging in a generate-recognize strategy. The use of a generate-recognize strategy violates two assumptions of PDP (e.g., Jacoby, 1998). First, it violates the assumption that conscious and automatic retrieval operate independently, because conscious retrieval can only occur following automatic retrieval. Consciously recognized responses can become a subset of those that are automatically retrieved, leading subjects to reject responses that were automatically retrieved, thus creating an underestimate in automatic retrieval estimates. As well, it violates the assumption that responding with studied items on the exclusion task occurs following automatic retrieval. Familiarity alone might motivate a subject to exclude a word as studied, and some have argued that familiarity may induce subjects to exclude all words that come to mind automatically, whether they are studied or unstudied (Richardson-Klavehn, Gardiner, & Ramponi, 2002). Since automatic estimates from PDP are thought to be a product of the studied words that are used in the exclusion task, then excluding words that are familiar (but not recalled as part of the study list) will reduce the automatic estimate obtained.

If automatic and conscious retrieval are positively correlated rather than independent, the PDP should underestimate the contribution of automatic retrieval more extensively for short compared to long retention intervals, since the probability of using, as well as the probability of successfully using, conscious retrieval is highest at short retention intervals. A study by Wilson and Horton (2002) provides evidence for this. Subjects were first exposed to a study trial whereby they completed word stems with words that made sense given a semantic cue (e.g., a large animal with a trunk – ELE). Following completion of the study trial, subjects were given a practice test whereby they completed word stems with the first word that came to mind *and* as quickly as possible. The practice test exactly mimicked the format of the critical test except that on the practice test, none of the word stems could be completed with previously studied words. Importantly, since none of the word stems that appeared on the practice trial could be completed with previously studied words, the subjects were assumed to be engaging in automatic retrieval since there was no basis for using conscious retrieval. Results from the critical test indicated that subjects were equally as fast in completing the word stems in a condition in which they could switch to a conscious retrieval strategy despite incidental retrieval instructions (since some stems could be completed with previously studied items) compared to the practice test in which they could not benefit from using a conscious retrieval strategy (since no stems could be completed with previously studied items). As well, subjects given an explicit stem completion task after completing an implicit stem completion task exhibited longer RTs compared to subjects in the speeded condition. Critically, these RT data indicate that the speeded group used an automatic retrieval strategy in the stem completion task, despite the fact that they could theoretically use a conscious retrieval strategy, thus providing empirical evidence for the use of an automatic retrieval strategy.

Following a 7 day delay between the study phase and test phase, estimates of automatic retrieval obtained from priming estimates from the speeded group were virtually identical to the estimates of automatic retrieval obtained from the PDP formula for the PDP group. Even if conscious and automatic retrieval processes are positively correlated rather than independent, this result if not surprising: Because the successful use of a conscious retrieval strategy is expected to be low following a long delay, estimates of automatic retrieval using the PDP formula should not substantially underestimate automatic retrieval despite the independence assumption, since low values of C will leave the value of A derived by the formula $A \mid \overline{C}$ relatively unchanged. In contrast, for a condition in which the study and critical test phases were given in the same experimental session (zero delay), estimates of automatic retrieval obtained using PDP formula

$$A = A \mid \overline{C} = \frac{E}{1-C} \tag{5}$$

were significantly lower for the PDP group compared to those obtained from the speeded group, as indicated by greater priming for the speeded group. Intuitively, the opportunity for conscious retrieval is expected to be higher following short delays between study and critical test phases. Thus, it is not surprising that the estimate of automatic retrieval using the PDP formula was low since $A \mid \overline{C}$ will underestimate A more substantially when C is high. Thus, PDP's estimation of automatic retrieval varies as a function of conscious retrieval, with underestimations increasing as the incidence of conscious retrieval increases. It seems that accurate estimations of automatic retrieval will only occur when the estimate of conscious retrieval approaches zero (Wilson & Horton, 2002). Horton et al. found that estimates of automatic retrieval derived using the PDP formula were lower than those obtained in the speeded group, but only when the estimates of conscious retrieval were relatively high ($C = .55$ compared to $C = .19$). When conscious estimates were low, automatic estimates derived using PDP formula did not differ from those obtained from a speeded group. Wilson and Horton's (2002) study supports the hypothesis that automatic and conscious retrieval processes are actually positively correlated as opposed to independent.

Proponents of PDP have assumed that automatic retrieval is unaffected by depth manipulations, a claim that they propose is validated if it can be demonstrated that automatic estimates do not change following depth manipulations. Studies have found this null effect of depth on automatic estimates using PDP in both perceptual (e.g., Toth et al., 1994; Richardson-Klavehn & Gardiner, 1996), and conceptual (e.g., Mecklenbräuker, Wippich, & Mohrhusen, 1996; Toth et al., 1994) tests. For example, using a category exemplar generation task, both McBride and Shoudel (2003) and Mecklenbräuker et al. showed that estimates of conscious retrieval increased following semantic study relative to nonsemantic study, but estimates of automatic retrieval processes showed the exact opposite pattern. Mecklenbräuker et al. suggested that since larger automatic estimates were not obtained following semantic study, then conceptual priming is not a product of automatic uses of memory. However, automatic estimates were actually found to be lower than the average target baseline following the semantic encoding condition in both studies, implying that subjects were not using a direct retrieval strategy as instructed, but were instead using a generate-recognize strategy. As well, McBride and Shoudel found that subjects were more likely to correctly guess the target on exclusion than on inclusion tasks; this strategic difference is not predicted by proponents of PDP, and can be taken as evidence of a violation of the independence of automatic and conscious processes: Producing a target response through guessing should be equally likely in exclusion and inclusion conditions if subjects are using a direct retrieval strategy and not a generate-recognize strategy. After correcting for this difference by using a generate-recognize model to derive memory estimates, McBride and Shoudel found higher automatic estimates following semantic study compared to nonsemantic study. This result illustrates the inadequacy of the PDP for obtaining memory estimates when it appears as though subjects are using a strategy other than direct retrieval to respond to test items. Importantly, the finding that automatic retrieval estimates increased following semantic study implies that automatic estimates can be influenced by conceptual depth manipulations, contrary to the predictions of PDP. Other studies have found that automatic estimates do increase following conceptual elaboration at study for both perceptual (e.g., Horton, Wilson, Vonk, Kirby, & Nielsen, 2004; Richardson-Klavehn & Gardiner, 1998) and conceptual (Bergerbest & Goshen-Gottstein, 2002; McBride & Shoudel, 2003) memory tests when memory estimates are obtained through means other than PDP formula.

Speeded response technique. The speeded response technique is another approach that has been used to investigate automatic retrieval. Instead of relying on questionable assumptions that underlie RIC and PDP, it provides converging evidence of the incidence of automatic retrieval by making use of a commonly observed phenomenon that occurs in implicit and explicit memory tests. Specifically, automatic retrieval is executed faster than conscious retrieval (e.g., de Houwer, 1997; Richardson-Klavehn & Gardiner, 1995, 1996, 1998; Toth, 1996; Yonelinas & Jacoby, 1994). Shorter response times for retrieving studied items is often taken as evidence that a subject is engaging in automatic retrieval strategies (Mace, 2003; MacLeod & Masson, 2000). Researchers have contended that explicit retrieval of studied items is less likely to occur in experiments in which subjects are encouraged to respond as quickly as possible (Gold, Beauregard, Roch Lecours, & Chertkow, 2003; Horton et al., 2001; Richardson-Klavehn & Gardiner, 1996, 1998). Requiring fast responses reduces the probability that subjects will become study test aware, thereby reducing the use of conscious retrieval.

A study by Neely (1977) illustrates how requiring fast responses can decrease the use of conscious retrieval processes in a conceptual memory test. Subjects were told that when a category label BODY or BUILDING appeared on a computer screen, they were to determine if the letter string that followed it was a word or a nonword, with half of the letter strings denoting words and half denoting nonwords. Critically, they were further told that if the word BODY appeared that they were to expect to see a building-related word, and when the word BUILDING appeared they were to expect to see a body-related word. However, on one third of the trials, a related word was given at test (e.g., BUILDING – window), contrary to the expectation subjects had established. Since the words that subjects were told to expect following the category labels were not the natural associates of the category, it was hypothesized that the subjects would have to consciously attempt to overcome the tendency to think of naturally associated words in order to think of the category exemplars that they were instructed to expect. This hypothesis was supported: Subjects verified words as real words faster when they were natural associates of the category at the short stimulus onset asynchrony (SOA) of 250ms but verified words faster when they were not natural associates of the category (i.e., they were from the "expected" category) at the long SOA of 2,000ms. Apparently, short SOA's did not allow subjects sufficient time to engage in conscious implementation of the instruction that they were to expect a category label to be followed by a word from a different category (e.g., a building-related as opposed to body-related word following the category label BODY), whereas long SOA's did allow subjects to override preexisting associations and to expect what they were told to expect.

Using the speeded response technique, estimates of pure automatic retrieval are derived by comparing the response times (RTs) of three groups (speeded, baseline, explicit). The groups are first exposed to identical study conditions. In this study phase, no group is informed that their memory for the stimuli will later be tested. All three groups are then exposed to an identical practice test under incidental memory instructions and are told that they are to respond as quickly as possible to whatever task they are given. For instance, subjects in Horton et al. (2001) were given word stems (e.g., ELE) and asked to complete the word stems with the first word that came to mind as quickly as possible. In order to optimally encourage quick responding, subjects are told that they will be shown their average RT following the completion of each task, and that they are to try to improve on that time on subsequent tests. All subjects in the practice test are assumed to be using an automatic retrieval strategy because previously studied items cannot serve as responses, thus providing no basis for using a conscious retrieval strategy.

The critical memory test occurs following completion of the practice test. The three groups are exposed to different conditions in order to obtain converging evidence that automatic retrieval is occurring. One group (explicit) is given intentional memory instructions, whereby they are instructed to use the test cue to recall a word that was previously studied. Importantly, asking subjects to intentionally retrieve items for a memory test after they had been given incidental memory instructions for an identical task has been found to produce significant increases in RTs compared to a group only given incidental instructions (Horton et al., 2001; Horton et al., 2004; Richardson-Klavehn & Gardiner, 1995, 1996, 1998; Wilson & Horton, 2002), thus providing converging evidence for the claim that conscious retrieval operates more slowly than automatic retrieval. For instance, using the speeded response technique, Horton et al. (2001) found that on a critical stem completion task, subjects in an explicit group showed a significant increase in RT compared to subjects in a speeded group following a practice test in which both groups of subjects were given incidental (implicit) instructions and both had identical RTs. The fact that the subjects given intentional retrieval instructions on the critical test exhibited longer RTs compared to earlier performance when incidental instructions were given indicates that these subjects switched to conscious retrieval strategies as instructed (Richardson-Klavehn & Gardiner, 1995, 1996, 1998) and provides support for the idea that conscious retrieval operates more slowly than automatic retrieval (de Houwer, 1997; Richardson-Klavehn & Gardiner, 1995, 1998; Toth, 1996; Yonelinas & Jacoby, 1994).

The two other groups (baseline and speeded) are, as in the practice test, given incidental instructions for the critical memory test, but only one of the groups (speeded) has a basis to use previously studied items as responses, since the other group (baseline) is never given previously studied items on the critical memory test. Because the baseline group is not exposed to items corresponding to previously studied items, they have no basis for switching from an automatic to a conscious retrieval strategy on the critical test. By contrast, the speeded group is exposed to items corresponding to previously studied items and therefore has the opportunity to engage conscious retrieval on the critical test. If, however, RTs in the speeded group do not differ from those of the baseline group, and are faster than those in the explicit group, then we would have evidence that the speeded group also did not switch to a conscious retrieval strategy on the critical test.

The first evidence to support the use of automatic retrieval comes from a comparison of RTs on the critical test of those subjects given incidental instructions (speeded and baseline) to those given intentional instructions (explicit). If subjects in the explicit group have significantly faster RTs compared to subjects in the speeded and baseline groups, then there is evidence that those subjects in the explicit group switched to a conscious retrieval strategy as instructed - supporting the assumption that conscious retrieval takes longer to execute than automatic retrieval - as indicated by the longer RTs. The second piece of evidence to support the use of automatic retrieval comes from comparing RTs of groups given incidental instructions (speeded and baseline) but where only one group had a basis for using a conscious retrieval strategy (speeded group - since only that group received stems corresponding to previously studied items). If RTs of these two groups are the same, then there is evidence that the group who could theoretically switch to a conscious retrieval strategy did not.

Other researchers have used similar procedures to estimate automatic retrieval (e.g., Richardson-Klavehn & Gardiner, 1995, 1996, 1998; Toth, 1996). With regards to priming scores, research has yielded data similar to those obtained using Horton's speeded response technique, indicating that priming on perceptual implicit memory tasks is guided by automatic retrieval (Horton et al., 2001; Horton et al., 2004; Richardson-Klavehn & Gardiner, 1995, 1996, 1998;

Wilson & Horton, 2002). However, the RTs obtained from Horton's procedure have been much shorter than those from other research (e.g., 0.8-s – 0.9-s per item in Horton et al. (2001) compared to 2.5-s and 2.6-s per item in Richardson-Klavehn & Gardiner, 1998; Richardson-Klavehn, Clarke, & Gardiner, 1999). Long RTs increase the probability that subjects will resort to using conscious retrieval strategies, thereby reducing the probability that the subjects were using pure automatic retrieval strategies in these other studies.

A critical instructional difference in Richardson-Klavehn and Gardiner's approach likely mediated the heightened potential for explicit contamination as indicated by the long RTs. Their subjects were instructed that "the purpose of timing them was not to rush them, but that it was necessary to measure how long it took to perform the remaining tasks" (Richardson-Klavehn & Gardiner, 1998, p. 598). Thus, the instructions do not optimally discourage the potential for the use of conscious retrieval strategies, and the relatively long RTs obtained using these instructions (e.g., 2.5s and 2.6s per item in the incidental conditions in Richardson-Klavehn and Gardiner, 1998, and Richardson-Klavehn et al., 1999) *may* reflect the use of conscious retrieval strategies. In contrast, Horton's subjects were given instructions designed to emphasize speed of responding as the primary task: Subjects were told that their responses were being timed before beginning the practice test and are also shown their average RTs after each practice test and critical test so that they may try to improve their speed on subsequent tests. The relatively shorter RTs obtained using Horton's procedure (e.g., median RT's of 0.9s for non-studied items and 0.8s and 0.7s for studied items) provide stronger evidence for the conclusion that subjects were using pure automatic retrieval compared to Richardson-Klavehn and Gardiner's subjects.

Perceptual priming

Based on the processing demands of a memory test, implicit memory tests can be classified into two major groups - perceptual and conceptual (Vaidya et al., 1997). Test stimuli in perceptual memory tests are presented very rapidly or in fragmented form and test performance requires analysis of the stimulus form. A prime example of a perceptual task is word stem completion, in which only the first three letters (typically) of a previously studied word are provided in a test phase and subjects are asked to complete the stem with the first word that comes to mind. In the explicit version of the stem completion task, subjects are asked to complete the word stem with a previously studied word. In contrast, in the implicit version of the stem completion task, subjects are asked to complete the word stem with the first word that comes to mind.

In contrast to performance on explicit perceptual memory tests, performance on implicit perceptual tasks is heavily reliant on the degree of similarity between the perceptual features of the study and test stimuli (Mecklenbräuker et al., 1996). Thus, changing the modality (visual – auditory) or the symbolic form (picture – word) of test items relative to study items decreases priming obtained in perceptual tests (Park & Gabrieli, 1995; Srinivas, 1993; Weldon, 1991; Vaidya, 1997). Although semantic manipulations in the study phase of a memory experiment, such as deep processing or generating words from semantic cues, usually increases perceptual priming on explicit tests, these depth manipulations do not always increase the amount of priming on implicit perceptual tasks (Vaidya et al., 1997; Richardson-Klavehn & Gardiner, 1996; but see Horton et al., 2001; Horton et al., 2004). Instead, priming on perceptual implicit memory tasks appears to be mediated primarily by perceptual feature overlap, where performance benefits most strongly by attending to the nonsemantic, physical properties of a stimulus, as opposed to semantic feature overlap of study and test items as in perceptual explicit memory tasks. Findings that depth manipulations can increase priming obtained in perceptual implicit tests suggests that automatic retrieval can be enhanced following conceptual elaboration, but does not undermine the fact that the priming on perceptual tests is mediated primarily through perceptual processing demands of the memory test.

Perceptual automatic priming

Many studies have demonstrated that priming in implicit perceptual memory tests reflects the use of automatic retrieval (Horton et al., 2001; Horton et al., 2004; Richardson-Klavehn & Gardiner, 1996, 1998; Richardson-Klavehn et al., 1999; Wilson & Horton, 2002). The concern that priming obtained on these perceptual implicit tasks may reflect a combination of automatic and intentional retrieval processes as opposed to solely automatic processes is not supported in experiments using the speeded response technique (Horton et al., 2001; Horton et al., 2004; Richardson-Klavehn & Gardiner, 1996, 1998; Richardson-Klavehn et al., 1999; Wilson & Horton, 2002). Some have argued that implicit perceptual tests do not share as many processes with explicit tests as conceptual implicit tests (Butler & Berry, 2001). Attending to perceptual features of study and test stimuli arguably requires a lesser degree of attention than attending to conceptual features. Since attending to perceptual - not semantic - features of study and test stimuli is the primary mediator of priming in perceptual tests (because changes in the modality or symbolic form of the test stimuli relative to the study stimuli greatly reduces priming), researchers have contended that priming on perceptual tests is less likely to be contaminated by conscious retrieval strategies.

Conceptual priming

If the processing demands of a memory test includes processing stimulus meaning, or otherwise require semantic analysis of test stimuli, then the memory test is classified as conceptual (Vaidya et al., 1997). Priming on an implicit conceptual test relies on prior encoding of conceptual information. For example, in category exemplar generation tests, subjects are first exposed to exemplars of different categories. Then, at test, they are given category labels and are asked to provide the first exemplars of that category that come to mind. Responding with more exemplars that were previously studied than nonstudied items indicates priming. In category cued association tests, subjects are first exposed to words that are associated in some way with words that will later appear on a critical memory test. Responding with more of the previously studied words as the first word that comes to mind indicates priming. In conceptual tests of general knowledge, subjects are exposed to a list of stimuli (usually words) that are potential responses to general knowledge questions given at test. Responding with more of these studied stimuli compared to nonstudied stimuli indicates priming. The physical similarity of study and test items does not affect priming in either of the aforementioned conceptual implicit tests or others, such as category verification, where priming is indicated when subjects verify category membership faster for previously studied exemplars. For example, presenting study stimuli in auditory form and presenting the test stimuli in visual form does not lower conceptual priming relative to conditions in which study and test stimuli are presented in the same modality in such implicit conceptual tasks as category cued association (Srinivas & Roediger, 1990), category verification, and category exemplar generation tests (Vaidya et al., 1997).

Relative to perceptual implicit tasks, the study of conceptual implicit tasks has only recently been pursued (Butler & Berry, 2001). Researchers in the past maintained that implicit memory was unaffected by conceptual manipulations that enhance the semantic encoding of study stimuli, such as generating study words from sentences or verifying some semantic quality of a study word (Graf & Mandler, 1984; Jacoby & Dallas, 1981), contending that only explicit tests such as recognition and free recall were sensitive to conceptual manipulations performed during encoding. A classic study conducted by Jacoby (1983) demonstrated this once widely held notion by having subjects vocalize words under three different conditions: Reading the word in isolation without a context (e.g., *xxx-cold*), reading a word embedded in a meaningful context (e.g., *hot - cold*), or generating a word from a given context (e.g., *hot - ????*). Following this study phase, subjects were given an implicit perceptual identification test in which they identified words presented for very brief intervals. In contrast to what would be expected if the implicit task was conceptually based, results indicated that reading words in isolation produced greater priming compared to the generation condition – the opposite of what was found in an explicit test. The authors suggested that only explicit memory tests are conceptually based since priming levels did not increase following conceptual encoding manipulations whereas explicit memory test performance did increase.

An alternative explanation - one that does not rule out the possibility that implicit tests can be influenced by conceptual manipulations - can also account for the findings of Jacoby (1983). According to the theory of transfer appropriate processing (TAP), priming on implicit memory tests will be possible when operations performed at study are similar to the operations necessary to produce the best performance on the later memory test. Unlike explicit perceptual tests, performance on implicit perceptual tests depends primarily on analysis of stimulus physical form, and is usually not enhanced by semantic encoding of study items. Since the memory test used in Jacoby's study demanded use of perceptual information, it is not surprising that the conceptual manipulation at study did not benefit implicit test performance. More recently, levels of processing manipulations that increase conceptual processing during encoding have been found to influence priming on conceptual implicit tests that demand the use of conceptual information (Blaxton, 1989; Hamilton & Rajaram, 2001; Srinivas & Roediger, 1990), providing evidence that conceptual processes can mediate priming.

Conceptual priming is sometimes improved following conceptual elaboration of study stimuli such as deep semantic encoding or generating study stimuli from semantic information for conceptual tests of general knowledge (Hamilton & Rajaram, 2001), category exemplar generation (Keane, Monti, Fleischman, Gabrieli, Cantor, & Noland, 1997; McBride & Shoudel, 2003; Mulligan, Guyer, & Beland, 1999; Mulligan, 2002; Weldon & Coyote, 1996), and category cued association (Srinivas & Roediger, 1990; Vaidya et al., 1997). For instance, deep processing of words using a pleasantness rating task mediated higher priming levels for a word cued association task when the cue target associations were weak (Vaidya et al., 1997, Exp. 3). Depth effects were not found when the cue target associations were strong (but see Vaidya et al., 1997, Exp. 4), however, suggesting that depth effects in conceptual implicit tests of association can be hard to detect when the probability of generating a strongly related cue is high prior to the study phase. Even when depth effects were absent for the high frequency words, priming levels did not decrease following visual or auditory study when the test was only presented in one modality, suggesting that perceptual processes were not mediating the priming effect. Another semantic encoding manipulation, one that entails generating critical test items from semantic information in the study phase, has been found to improve conceptual priming on tests of general knowledge (Blaxton, 1989) and category exemplar generation (Mulligan, 2002).

If conceptual implicit tasks are dependent on semantic elaboration that occurs at study, then interrupting cognitive resources that are available at encoding should decrease priming. Although some studies have shown significant conceptual priming despite lowered attention during encoding (e.g., Isingrini, Vazou, & Leroy, 1995), dividing attention during study has been shown to decrease conceptual priming in tests of category association (Mulligan, 1997; Mulligan & Hartman, 1996; Schmitter-Edgecombe, 1999) and category exemplar generation (Light, Prull, & Kennison, 2000; Mulligan, 1997). Studies that have shown conceptual priming to be unaffected by divisions of attention at study have prompted suggestions that performance on conceptual implicit memory tests reflects automatic encoding processes (Graf & Mandler, 1984; Isingrini et al., 1995; Jacoby, Toth, & Yonelinas, 1993). Alternatively, studies that have shown conceptual priming to be negatively affected by divisions of attention at study have prompted exactly the opposite conclusion: If, as is typically assumed by proponents of TAP (e.g., Craik, 1983), semantic encoding requires attention, then it follows that attention during encoding is critical for supporting conceptual memory in either implicit or explicit tests (Mulligan & Hartman, 1996).

How, then, can the discrepancy in results of divided attention studies be explained? It seems that the strength or difficulty of the divided attention manipulation determines whether conceptual priming will be affected (Mulligan, 1997; Mulligan & Brown, 2003; Wolters & Prinsen, 1997). During encoding, Isingrini et al. (1995) had subjects listen to a tape and signal whether they heard the letter "B" or "G", while Mulligan and Hartman (1996) had subjects listen to a tape and signal whether they heard a sequence of three consecutive odd numbers. The divided attention manipulation administered by Mulligan and Hartman is arguably more difficult, since unlike in Isingrini et al., subjects had to maintain digits in memory in order to perform the task. Mulligan and Hartman's subjects were also more accurate in performing the divided attention task, perhaps because they were signaled during encoding if they got a sequence incorrect, thus causing them to focus more on the secondary task compared to Isingrini et al.'s subjects who were not signaled if they mistakenly supplied an incorrect letter, again suggesting that the divided attention manipulation in the Mulligan and Hartman study was stronger. Thus, it seems that conceptual information is not encoded automatically, and that attention to semantic features is crucial for mediating conceptual priming. However, this does not mean that priming on conceptual implicit memory tests cannot occur automatically. In order for either automatic or conscious retrieval to be successful, the critical test phase items must be adequately processed in the study phase, be it through automatic or conscious encoding.

Semantic verification test: A conceptual test

Priming in the category verification task is indexed by how much faster subjects are able to verify studied versus unstudied exemplars of a category. The task has both perceptual and conceptual features: It is perceptual in that study words are re-presented at test and it is conceptual in that the task of verifying an exemplar as a member of a category requires conceptual analysis of the test stimuli. Evidence indicates, however, that priming on the category verification task is not mediated by the perceptual features of the task, leading many researchers to argue that priming found on implicit verification tasks is conceptually mediated (Bowers & Turner, 2003; Vaidya et al., 1997; Vriezen, Moscovitch, & Bellos, 1995).

A key finding indicating that priming on the category verification task is not perceptually driven is that changes in perceptual features of study and test stimuli do not affect the response latency for verifying category membership of studied items. For instance, Vaidya et al. (1997) presented subjects at study with category exemplars in either visual or auditory form but in the test phase presented exemplars presented in only visual form. Analysis of median RTs indicated that subjects were equally fast in verifying category membership for exemplars previously presented in auditory form as those previously presented in visual form. Since perceptual changes in stimuli - such as the modality manipulation used here - reduce perceptual, but not conceptual, overlap, the failure to find a reduction in priming for items studied in auditory form is evidence that the priming was conceptual in nature (Srinivas & Roediger, 1990).

Equivalent cross symbolic form priming for pictures and words has also been found in studies that, like the category verification task, index priming by response latency. Bruce, Carson, Burton, and Ellis (2000) found equivalent picture to word priming of exemplars using a semantic decision task whereby subjects had to verify whether pictured exemplars were manufactured or naturally occurring (Exp. 3). They also found that two different semantic classification tasks performed at study - determining whether pictured exemplars were manufactured or whether they were more frequently found inside or outside of a house - equally primed the naming of the same or different pictorial depictions of the studied exemplars in the memory test (Exp. 1). The lack of sensitivity to the perceptual characteristics of the exemplar indicates that naming the object in the test phase - something that may seem to rely on perceptual features of the words themselves - indicates priming is controlled conceptually.

Other evidence that priming on classification tests is conceptually based comes from studies that have found priming to be unaffected when different semantic classifications were made at study and test. Bowers and Turner (2003) found equivalent priming for low frequency words when the study phase required either a classification regarding the size or state (natural or manufactured) of the referent of a word and the test required a classification of only one criterion, indicating that subjects retained the conceptual features of the words necessary to complete different classifications tests as quickly as they could complete the same classification made at study. This priming was not modality specific for either low frequency regular or irregular words. Since irregular words cannot be read by relying solely on sublexical grapheme-phoneme correspondences (i.e., low-level perceptual information), Bowers and Turner hypothesized that irregular words should mediate greater within-modality priming compared to cross-modality priming, a hypothesis postulated based on the notion that perceptual priming relies on access to lexical-orthographic representations. The finding that cross-modality priming (and within-modality priming) was significant for irregular as well as regular words provides strong support for the notion that priming in the categorization task is conceptually rather than perceptually mediated.

The wealth of studies showing that modality manipulations do not influence priming on categorization and verification tasks, category exemplar generation, and category cued association tests (e.g., Vaidya et al., 1997) strongly suggests that performance on these implicit tests are not dependent on perceptual feature overlap of study and test stimuli. Instead, the evidence suggests that the conceptual tasks of categorizing items as exemplars of categories or verifying exemplars as category members are conceptual tasks that are dependent upon the *semantic* overlap of study and test processing - and independent of perceptual feature overlap - as theories such as TAP predict.

Dissociations among conceptual implicit tests of memory

Just as tests of implicit and explicit memory are dissociable based on several factors, tests of conceptual implicit memory dissociate on several factors. The finding that factors that normally cause a change in performance on some conceptual tasks actually do not for others has led some researchers to question the validity of the latter tests as conceptual. Some conceptual implicit tests that measure priming in terms of response latency, such as tests of category verification and classification tasks (whereby the classifications being tested have been ascertained in a previous study phase) appear to be consistently insensitive to semantic encoding manipulations that often increase priming on other tests of conceptual implicit memory such as category exemplar generation, category cued association, and implicit general knowledge tests. For instance, Vaidya et al. (1997) found that priming following deep encoding (judging whether the referent to a word was man-made) was no greater than priming following shallow encoding (whether the word contained the letter "a") in a category verification task. In contrast, subjects given a category cued association test showed more priming for words that had been deeply encoded. The same pattern of results was found with priming in an abstract/concrete classification task: Subjects were no faster in judging if a previously studied word was concrete or abstract if the words had been deeply encoded during a prior study phase. Similarly, Light et al. (2000) found that rating the pleasantness of exemplars on a five-point scale produced priming levels in a category verification test that were no greater than those produced following a letter counting task. Thus, semantic encoding manipulations dissociate tests of implicit memory: Category exemplar production, category cued association, and word cued association tests have all been found to benefit from semantic encoding manipulations (e.g., Hamann, 1990; Keane et al., 1997; Mulligan et al., 1999; Mulligan, 2002; Srinivas & Roediger, 1990; Vaidya & Gabrieli, 2000) whereas verification and classification tasks do not (Vaidya & Gabrieli, 2000; Vaidya et al., 1997).

A second way in which conceptual implicit memory tests have been found to dissociate is in regards to memory for pictures and words following different encoding tasks. The picture superiority effect is the phenomenon that the pictures of objects are often recalled better than the written names of the same objects. Although the picture superiority effect is found in some tests of conceptual implicit memory, such as category cued association (e.g., Vaidya & Gabrieli, 2000; Weldon & Coyote, 1996) and category exemplar production (McBride & Dosher, 2002), it is not evident in category verification tests (Vaidya & Gabrieli, 2000). Thus, these two conceptual tasks differ in sensitivity to encoding manipulations, with category verification being insensitive to semantic encoding manipulations that mediate the picture and word superiority effects that occur in tests of category generation and category cued association.

A model that distinguishes between competitive and non-competitive access of conceptual knowledge has been proposed to account for the insensitivity of some implicit tests to conceptual encoding manipulations (Vaidya et al., 1997). It has been proposed that the reason why implicit tests, such as category verification tests, are insensitive to conceptual manipulations is because the cue provided at test is identical to the cue provided during the study phase, thus providing direct access to the relevant conceptual knowledge of the test cue since there are no competing response alternatives. For example, a subject participating in a category verification task may see the word "dog" in the study phase and then in the critical test phase will be asked "is this an animal....dog." Thus, an implicit test in which the study cue is also given at test provides subjects with direct, non-competitive access to the conceptual knowledge (i.e., "dog" is an exemplar of the category "animal") that is needed to respond. In contrast, implicit tests that are sensitive to conceptual encoding manipulations provide a test cue that reveals only a semantic criterion - a criterion that evokes competitive access among categorical information - for which the subject has to use in order to generate one of several legitimate responses. For example, a subject participating in a category exemplar generation task may see the word "dog" in the study phase and then in the critical test phase will be asked to give an example of the category "animal." Thus, an implicit test in which the test cue is representative of the category of the study cue provides subjects with only the indirect access to the conceptual knowledge (i.e., the link between "dog" and "animal" was not explicitly primed during study) that is needed to respond. Thus, subjects must choose among competing alternatives in order to respond (i.e., the test category "animal" triggers several exemplars among which subjects must select). It seems that only implicit tests that elicit competitive access to categorical information are sensitive to conceptual elaboration that occurs at study.

Automatic conceptual priming

Priming has been found on implicit memory tests that require conceptual analysis of test stimuli in both normal (Graf, Shimamura, & Squire, 1985; Mace, 2003; Weldon & Coyote, 1996) and amnesiac subjects (Graf et al., 1985; Keane et al., 1997; Vaidya, Gabrieli, Keane, &

Monti, 1995). Research on amnesiacs has usually shown intact priming on implicit memory tasks but impaired performance on explicit memory tests (Keane et al., 1997; Warrington & Weiskrantz, 1968, 1974). For instance, Graf et al. (1984) found that amnesiac subjects exhibited normal stem completion priming but impaired recognition and free recall of the same words. Keane et al. (1997) found that amnesiac subjects exhibited normal category exemplar generation priming (as well as depth effects) but impaired explicit recall performance. Since amnesiacs fail to explicitly recall or recognize words even after a short delay, it follows that priming exhibited by these subjects on implicit memory tests is not likely contaminated by conscious retrieval strategies, leading many researchers to assume that performance on implicit memory tests is guided by automatic retrieval processes. Incidental instructions were then assumed to induce automatic retrieval processes in normal subjects. However, subjects with intact conscious retrieval abilities may nonetheless opt to use conscious retrieval strategies in an implicit test in order to facilitate test performance through a developed awareness of the study test relationship. Thus, although evidence of intact priming in amnesiacs shows that priming can occur through automatic retrieval, it is nonetheless important to control for the use of conscious retrieval strategies in normal subjects.

Deep semantic encoding of study stimuli has been shown to increases conceptual priming in amnesic patients (Keane et al., 1997), suggesting that depth manipulations affect automatic forms of memory. Keane et al. found greater category exemplar generation priming following deep processing at encoding (deciding whether a study word was man-made or natural) compared to a shallow encoding condition (deciding whether a study word appeared in upper or lower case). The effect of the depth manipulation was equivalent in normal and amnesiac subjects, indicating that the effect could have been mediated automatically for both normal and amnesiac subjects. Following an encoding task in which subjects rated how much they liked exemplars of categories, Graf et al. (1985) found that normal and amnesiac subjects exhibited virtually identical priming in a category exemplar generation task. Taken together with the finding that the amnesiac subjects' explicit memory performance was far below that of normals, the virtually identical priming for normal and amnesiac subjects could indicate that the priming was mediated by automatic memory processes for both groups of subjects.

It is important to note that, although virtually identical conceptual priming levels in amnesiac and normal subjects strongly suggests that both groups used automatic retrieval, assuming that this is *always* the case may be a mistake. Implicit test performance of subjects who are capable of engaging in intentional retrieval strategies may at least partially reflect conscious retrieval, with the use of conscious retrieval arising from a developed awareness of the relationship between the study and critical test phase of the memory experiment. Since target completion rates are higher following the use of conscious retrieval strategies compared to the use of automatic retrieval strategies (Wilson & Horton, 2002), the finding of higher target completion rates in normal subjects compared to amnesiac subjects would imply the use of a conscious retrieval strategy in the normal subjects. Indeed, this result has been found: Carlesimo (1994) found that although priming levels on an implicit category association test were significant for both normal and amnesiac subjects, the overall priming levels for the amnesiacs was half that found for the normals.

The procedures implemented in many tests of implicit memory may encourage the use of conscious retrieval strategies in normal subjects. For instance, in many studies of conceptual implicit memory, most of which use the category exemplar generation test, subjects are given a substantial amount of time on the critical test in which to generate responses. Increasing the amount of time allotted to subjects to complete an implicit memory test can serve to increase the probability that test awareness will arise. For instance, in Graf et al. (1985), subjects generated the first eight exemplars that they could think of when the category label appeared in order "to ensure that exemplars other than the very common ones would be produced" (p. 392). It seems likely that the time it takes to generate eight exemplars for a category might provide subjects a reasonable opportunity to think back to the study list to facilitate memory performance. Similarly, Mace (2003) found that normal subjects exhibited conceptual priming in a category exemplar generation task but he also gave subjects more than sufficient time (20 seconds to list one exemplar for a category) to switch from an automatic retrieval strategy to a conscious one. Thus, although evidence from amnesiacs shows that conceptual priming can occur through automatic memory mechanisms, it is important to control for the potential use of conscious retrieval in normal subjects. By implementing procedures such as the speeded response technique, conscious contamination that may occur in priming on implicit memory tests for normal subjects can be detected.

Conceptually mediated priming is only automatic if the implicit test is free of explicit contamination

In order to assert that conceptual manipulations improve priming on implicit conceptual tests, it is necessary to show that the priming found on implicit tests is a product of automatic retrieval processes and not controlled retrieval processes (McBride & Shoudel, 2003). Semantic encoding may increase the chance that subjects will become study test aware and use conscious retrieval strategies, especially on cognitively demanding tasks such as category exemplar generation. It has been suggested that conceptual implicit tasks share more of the same processes in common with explicit memory tasks than do other implicit tests, such as implicit perceptual tasks (Butler & Berry, 2001). Thus, many researchers contend that performance on implicit conceptual tests is more prone to contamination from controlled retrieval strategies compared to implicit perceptual tests (Butler & Berry, 2001).

Arguments that only explicit tests are influenced by conceptual encoding manipulations (Jacoby & Dallas, 1981; Craik, Moscovitch, & McDowd, 1994) have been used to differentiate conscious and automatic memory processes. In particular, McBride and Dosher (2002) used Jacoby's PDP to argue that it is only explicit, consciously controlled memory that is influenced by conceptual manipulations, and that implicit, automatically controlled memory is only influenced by perceptual manipulations. When depth manipulations are found to influence priming in implicit tests (e.g., Weldon & Coyote, 1996; Wippich, Melzer, & Mecklenbräuker, 1998), proponents of PDP have argued that conscious retrieval strategies may have contaminated the automatic retrieval strategies that mediate "process pure" priming (McBride & Dosher, 2002). Arguments of this sort are largely based on theoretical assumptions postulated regarding the types of processing that occur in automatic and conscious memory (Toth et al., 1994) and on questionable assumptions of PDP, as opposed to relying on convergent evidence of conscious contamination of automatic implicit memory.

Researchers have defended priming on conceptual implicit tests as automatic and free of explicit contamination because subjects were naive as to the purpose of the study (e.g., Gold et al., 2003). However, at least one study that used posttest questionnaires to assess study test awareness in an implicit category exemplar generation test has shown increased depth effects following increases in level of awareness for the study test relationship (Mace, 2003). In this study, out of the 49 subjects who were not informed of the study test relationship, 33 indicated on a posttest questionnaire that they became aware of the study test relationship sometime during the critical category exemplar test. Subjects who were not told but became aware (test uninformed - aware) exhibited stronger depth effects (though the depth effect found was only marginally significant, $p = .065$) than subjects who did not realize (test uninformed – unaware) that previously studied exemplars could be used to facilitate category exemplar generation performance. Subjects who were informed that "they may sometimes experience words coming to mind from the previous experiment (i.e., the study phase)" but that "they should try not to think about this" (Mace, 2003, p. 285) exhibited significantly greater depth effects than both groups. Since intentional memory instructions often facilitate greater depth effects compared to incidental memory instructions (Horton et al., 2004), it is not surprising that a developed awareness of the study test relationship can also create depth effects that resemble those found under intentional recall instructions, if the developed awareness leads to the use of conscious retrieval strategies. However, if depth effects are not found in implicit memory tests, despite awareness of the study test relationship, then this would imply that conscious retrieval does not necessarily follow from awareness, and priming in implicit memory tests can occur through automatic retrieval. Mace may have found an increase in depth effects for the test aware subjects because the conditions of his study facilitated the use of conscious retrieval strategies. For example, subjects in his study were given 20 seconds to respond with as many category exemplars as possible in the critical test, making it more likely for subjects to adopt a conscious retrieval strategy than if subjects were encouraged to give only one response as quickly as possible.

Current experiment

The speeded response technique (Horton et al., 2001) can be used to provide converging evidence that explicit contamination may occur on tasks designed to measure automatic retrieval for conceptual memory tasks by comparing the RTs of subjects who are using automatic retrieval to those who are using conscious retrieval. Specifically, it was expected that, if subjects in the speeded group adopted a conscious retrieval strategy on at least some percentage of the critical test trials, then RTs for this group should be longer than those for subjects in the baseline group. If subjects in the speeded group used an automatic retrieval strategy, then the RTs should be faster than those of a baseline group, since the speeded group should benefit from the previous study trial by showing priming for these items; the baseline group does not receive the study trial, so they cannot benefit from this priming. Research conducted by Richardson-Klavehn and Gardiner (1998) that used a very similar speeded response procedure to the one employed in the current experiment and found faster average RTs in a speeded group compared to the average RTs in a baseline group in the critical trial. The RTs in the speeded group may not be as long as those in the explicit group because the former subjects have no basis for adopting conscious retrieval in the early trials of the critical test (Horton et al., 2001). Thus, the speeded response technique was used to attempt to evaluate whether automatic retrieval processes can support priming on conceptual tasks, and in particular to this study, the category exemplar generation task. As in previous experiments exploring automatic memory using the speeded response technique, speed was emphasized as the primary task for the speeded, explicit, and baseline groups.

As well as comparing the average RTs of the speeded, explicit, and baseline groups, a comparison was conducted on the RTs of the individual target items generated by the speeded, explicit, and baseline groups. If the speeded group was responding based on an automatic retrieval strategy, then their RTs should be faster than those of the explicit group on the critical test. As well, RTs for generating target items should be faster in the speeded group compared to the baseline group because the speeded group had been primed with the target response whereas the baseline group had not.

A standard implicit and a PDP group were included in the study in order to compare automatic estimates obtained from these groups to the automatic estimates obtained using the speeded response technique. In typical implicit memory tests, an implicit group is given incidental memory instructions, like the instructions given to the speeded group, but unlike the speeded group, implicit subjects are not instructed to respond as quickly as possible. Automatic estimates obtained from previous work comparing automatic estimates of implicit and speeded groups have indicated equal automatic estimates for both groups (Horton et al., 2004; Wilson & Horton, 2002), suggesting that performance of the implicit subjects, like the performance of the speeded subjects, is not contaminated by conscious retrieval. If automatic estimates obtained from the implicit group in the current study are again the same as automatic estimates obtained from the speeded group, then it can be more strongly concluded that performance of subjects given standard implicit instructions can be controlled automatically. If RTs are longer for the implicit group compared to the speeded group, then evidence for conscious contamination in the implicit group will be implicated (given evidence that the speeded group used automatic retrieval through comparisons of RT data).

A PDP group is included in order to assess whether this technique of indexing automatic retrieval gives the same estimates as the speeded response technique. As previously stated, some studies have indicated that estimates of automatic retrieval using PDP are lower than those obtained using the speeded response technique, particularly when estimates of conscious retrieval are high (Wilson & Horton, 2002). Other studies have indicated that automatic estimates using PDP are actually higher for unstudied items than for studied items (McBride & Dosher, 2002), a puzzling result that brings into question the reliability of the PDP's use in measuring automatic retrieval in some contexts. If automatic estimates obtained from the PDP group are different than those obtained from the speeded group, then this could indicate that estimates using the PDP are flawed because of the questionable assumptions that are used to compute the estimates (i.e., that automatic and conscious retrieval processes operate independently and that awareness never accompanies automatic retrieval).

The category exemplar generation test was selected for this experiment. Subjects should be able to generate responses fairly quickly, as the test is relatively simple and there is very little problem solving to be done in comparison to other conceptual tests. Minimizing problem solving in a test phase may increase the chance that performance will reflect automatic processes (MacLeod & Masson, 2000). As well, a pilot study using a category exemplar generation task produced RTs in the baseline and speeded groups that were long in comparison to those found in previous work using the speeded response technique in the context of perceptual tasks (e.g., Horton et al., 2001; Horton et al., 2004; Wilson & Horton, 2002), indicating that the task required too much problem solving. However, subjects in the pilot study had the additional task of supplying an exemplar of a category that began with an initial letter that was given, which increased the difficulty of the task given that, in many instances, there were very few possible responses available. The current study does not impose such restrictions on responses, which should substantially reduce the problem solving involved and decrease RTs, as well as decrease the probability that subjects will develop study test awareness and elect to use an intentional retrieval strategy to facilitate memory test performance.

Since conceptual encoding manipulations have been found to increase conceptual priming on most conceptual implicit tasks, including the category exemplar generation task employed in the present experiment, subjects in the present study received a depth of processing manipulation at study in order to determine whether depth effects would appear given automatic retrieval. Given the evidence reported by Mace (2003), it was predicted that depth effects would be more prominent in those subjects who are aware of the study test relationship. Depth effects were expected to be greater in the explicit group compared to the speeded and implicit groups, given that those subjects in the speeded and implicit group claimed that they did not use intentional retrieval strategies on an awareness questionnaire administered after the completion of the trials. The awareness questionnaire was given to the speeded and implicit memory groups in order to examine the relationship of awareness to depth effects and overall priming levels to provide a stronger test of the hypothesis, as well as to the baseline group to check for false claims of awareness. In order to prevent subjects from trying to conform to perceived expectations, open ended questions were first used on the questionnaire. Since amnesiac subjects benefit from deep encoding of study stimuli in implicit category exemplar production tasks

(Carlesimo, 1994; Graf et al., 1985; Keane et al., 1997), it was hypothesized that deep encoding would benefit automatic conceptual priming.

Repetition priming is often reduced for words with a high baseline frequency generation rate relative to words with low baseline frequency generation rates and is sometimes eliminated (Bowers & Turner, 2003). Conceptual priming levels obtained using the category cued association task have often been greater following semantic processing for low compared to high frequency exemplars (e.g., Hamman, 1990), however, other studies have found that both high and low frequency exemplars benefit from semantic processing (Vaidya et al., 1997). Because the primary goal of the present study was to assess the automaticity of conceptual priming found on the category exemplar generation task, only low frequency exemplars were used in conjunction with depth manipulations because low frequency exemplars more consistently produce priming.

Encoding questions appeared on a computer screen along with the exemplar until the experimenter typed a response after which time both the question and the exemplar disappeared. Previous studies have implemented procedures in which the exemplar appeared after the appearance of the encoding question and disappeared after only 1 second. A longer time frame was allowed in the present study in an attempt to better ensure that subjects think about the exemplar and process its meaning more fully than would be possible in a shorter time frame.

Method

Participants

A total of 162 students from Wilfrid Laurier University participated in return for course credit.

Design and materials

A 5 x 2 x 2 x 2 incomplete mixed factorial was used, with the group variable (speeded, explicit, baseline, implicit, PDP) and depth of processing (semantic, nonsemantic) manipulated between subjects and test (practice, critical) manipulated within subjects. For the PDP group, inclusion and exclusion conditions were manipulated within subjects.

One hundred category labels were selected and paired with a low frequency exemplar selected from Battig and Montague's (1969) and Shapiro and Palermo's (1970) category norms. The low frequency exemplars were defined as those exemplars ranked 10 to 17. Caution was taken to insure that no exemplar was strongly related to any other category name or exemplar in the list. The category labels and exemplars are presented in Appendix 1.

Low frequency exemplars and their associated category labels were randomly assigned on a subject by subject basis to studied items, practice items, nonstudied items on the critical test such that each list contained 32 exemplars and category labels (64 nonstudied items were used on the critical test for the baseline group). For subjects given the semantic encoding task, the question directed subjects to rate the pleasantness of the exemplar on a scale from one to seven. For subjects given the nonsemantic encoding task, the question directed subjects to decide how many vowels were in the given word. For both encoding tasks, subjects had to vocalize the exemplar that appeared before making a decision. Two buffer items appeared at the beginning and end of each list. Critical test lists contained the 32 category labels corresponding to the studied items plus an equal number of nonstudied labels.

The awareness questionnaire contained five questions, the first three of which were modified from a questionnaire used in Geraci and Rajaram (2004) and the latter two of which were modified from a questionnaire used in Mace (2003). The complete awareness questionnaire is presented in Appendix 2. The first question asked subjects to describe any characteristics that they noticed about the category labels or the words that they encountered in the experiment. The second question asked subjects to describe anything that may have helped them to supply an exemplar on the category association task. The third question asked subjects to state what they thought was the purpose of the study. The fourth question more directly probed the existence of study test awareness, asking if subjects had ever become aware *during* the experiment that some of the category labels that appeared in the final phase of the experiment were representative of some of the exemplars encountered in the study phase. If subjects answered yes to the fourth question, they were asked to identify whether they had ever tried to deliberately retrieve studied items in order to complete the final task.

Procedure

Study phase. Subjects were tested individually with an experimenter present while facing a computer. After subjects had read and indicated understanding of the instructions, the study phase was initiated by the experimenter pressing a designated key on the computer keyboard. An exemplar appeared in the center of the screen along with either a semantic or nonsemantic category encoding question at the top of the screen. Subjects read the exemplar out loud and then vocalized their response to the question. The experimenter keyed the subject's response into the computer, after which the next exemplar appeared with the same encoding question. The experimenter's presence also verified that subjects were completing the task correctly. The experimenter stopped the study trial to provide more instruction, if required.

Practice test. Subjects in the speeded, explicit, and baseline groups were told that the experiment was designed to explore how fast people are able to process English words. At no time during the practice test were they told that their memory for the exemplars encountered in the study phase would be assessed.

The three groups of subjects were told that a category label would appear in the center of the computer screen and their task was to say the first member of that category that came to mind as quickly as possible. Subjects in the speeded, explicit, and baseline groups were further told that they would be shown their average RTs so that they may try to improve their speed in subsequent tests. Thus, at the end of the practice test, these subjects were shown their average RT in milliseconds. Subjects in the speeded and baseline groups were encouraged to try to improve on that time in the next test.

Subjects in the implicit and PDP groups were told that the experiment was designed to investigate how people process English words, with no particular emphasis on speed indicated to these groups. These subjects were given a famous name distracter task, whereby they were asked to vocalize a surname when shown a first name to make a famous name. The famous name task was designed to take about as long to complete as the practice test.

Critical test. The critical test was formatted exactly like the practice test except that for some of the subjects (i.e., speeded, explicit, implicit), some of the category labels corresponded to exemplars that had been previously studied. Subjects in three groups (speeded, explicit, baseline) pressed the spacebar to begin the critical test. The baseline and speeded groups were told that their task on this trial was exactly the same as on the previous (practice) trial. Subjects in the explicit group were given a category cued recall test. This group was informed that half of the categories were previously studied and that they should use the previously studied exemplars as responses where possible. If they could not think of a previously studied exemplar, they were to provide the first response that came to mind. All subjects in the speeded, explicit and baseline groups were reminded to respond as quickly as possible without sacrificing accuracy.

Subjects in the implicit group pressed the spacebar to begin the critical test after reading the instructions that appeared on the computer screen following the famous name task. Unlike in the speeded, explicit, and baseline groups, speed was not emphasized as a factor, and subjects were instead only told that they were to respond with the first word that came to mind.

Subjects in the PDP group were instructed that each trial consisted of the presentation of a category label in the center of the screen with the word *old* or *new* three lines above. Subjects were told that, when the word "*old*" appeared (inclusion), they were to use the category label as a cue to recall an exemplar of that category that had previously been studied. If they could not remember a study word, they were to give the first member of the category that came to mind. When the word "*new*" appeared (exclusion), they were again told to use the category label as a cue to recall an exemplar of that category that had been previously studied, but if they recalled a studied word, they were to give an example of the category that was not previously studied. If they could not recall a word that had been previously studied, they were to respond with the first word that came to mind. Thus, direct retrieval instructions were given to subjects in the PDP group. An example of this procedure was shown to the subjects before the critical test began in order to ensure that they understood the instructions.

Awareness questionnaire. Subjects in the speeded, implicit, and baseline groups were given the awareness questionnaire after completion of the critical test.

Results

An alpha level of .05 was used for all statistical tests. Twelve categories were removed from the analysis because at least 15% of subjects provided a response that was not from the designated category or indicated misunderstanding of the category (part of speech, type of metal, chemical, denomination of Canadian currency, member of clergy, natural earth formation, part of a boat, a piece of lab equipment, a substance to flavour food, type of cloth, a type of human dwelling). Another 18 categories were removed from the analysis because the responses within those categories overlapped (for example, subjects could respond to both "a type of music" and "a type if dance" with "jazz"). The 18 removed categories and the overall frequencies of overlapping responses are listed in Appendix 3. Thus, a total of 30 categories were removed from the analysis and responses from 70 categories were assessed. Data from six subjects were discarded because the subjects indicated either prior to or after the experiment that their first language was not English. Data from an additional 2 subjects from the implicit group were not used because the subjects indicated the use of intentional retrieval in the critical test despite incidental instructions.

Response times for nonstudied items on the practice and critical test

The RT data appear in Table 1. To determine whether subjects in the speeded group used a conscious retrieval strategy, an analysis of variance was conducted on the RTs as a function of group (speeded, explicit, baseline) for both the practice and critical tests. Only nonstudied items were used because studied items were expected to show RT priming. This priming might offset, and thus make difficult to detect, an increase in RTs due to conscious retrieval. There were no main effects of group for the practice test, $F < 1$, $MSE = 45456$. Importantly, there was a main effect of group for the critical test, $F(2, 94) = 28.02$, $MSE = 40706765$, with a least significant difference test (LSD = 5.26) indicating that RTs of the explicit group were reliably slower than those of the speeded and baseline groups, which did not differ (Horton et al., 2001; Richardson-Klavehn & Gardiner, 1995, 1996, 1998; Wilson & Horton, 2002).

The increase in RTs between the practice test and the critical test for subjects in the explicit group is evidence that this group switched to a conscious retrieval strategy, as instructed (Horton et al., 2001; Richardson-Klavehn & Gardiner, 1995, 1996, 1998). Combined with the finding that RTs in the speeded group did not differ from those of the baseline group on the critical test, these data indicate that the speeded group did not switch to a conscious retrieval strategy, but instead relied on an automatic strategy.

Response times for target items on critical test

Response time data for target items on the critical test (independent of condition) and the number of subjects included in each analysis appear in Table 2. Response times for the speeded group should be at least as fast as those of the baseline group for designated target items if RT priming occurs through automatic retrieval (Horton et al., 2001; Richardson-Klavehn & Gardiner, 1998). The RTs could conceivably be faster in the speeded group, since unlike the baseline group, automatic priming effects should bring previously studied responses to mind more readily. Subjects who did not generate at least two target exemplars in the critical test were removed from the analysis. A marginally significant group effect was found for RTs for target items in the speeded, explicit, and baseline groups, $F(2, 44) = 2.93$, $MSE = 781997$, $p = .064$. A least significant difference test (LSD = 2.92) indicated that the explicit group responded significantly slower than the speeded group, but not the baseline group, while the speeded and baseline groups responded equally as fast. When this analysis was repeated using data from subjects who responded with at least one target exemplar (leading to the inclusion of 11 additional baseline subjects), the same outcome was found.

Apparently, prior exposure to targets did not mediate an increase in speed of responding for the speeded group relative to the baseline group who did not receive such exposure. However, this result should be interpreted with caution since only two subjects in the baseline group generated more than one target exemplar in the critical test. Using items with a higher baseline frequency may have allowed for a more meaningful priming analysis since baseline subjects would be expected to respond with more target exemplars, making any differences in RTs for target exemplars between groups easier to detect.

Target completion for practice and critical test

Priming was significant for the speeded and implicit groups, $F(1,63) = 58.74$, $MSE = 63.26$, since both groups responded with more target exemplars in the critical test than in the practice test. Target completion data for the speeded, explicit, baseline, and implicit groups are reported in Table 3 and in Table 4 for the PDP group. Since we have evidence that the speeded group responded based on automatic retrieval, a comparison of target completion rates for the speeded and implicit groups can indicate whether the implicit group also responded based on automatic retrieval (Horton et al., 2004). Contamination in the implicit group with conscious retrieval would be evidenced by higher target completion rates compared to the speeded group. An analysis of variance on target completion rates on the critical trial for the speeded, explicit, and implicit groups revealed a significant group effect, $F(2, 95) = 28.53$, $MSE = 4.99$. A least significant difference test (LSD = 6.99) confirmed that the explicit group completed more test items with target (studied) exemplars than did the speeded and implicit groups, whose target completion rates did not differ. A comparison of means for the speeded and implicit groups suggested a trend for greater target completions in the implicit group compared to the speeded group, a trend that has been evidenced in every other study conducted using the speeded response technique (Horton et al., 2001, 2004; Vonk & Horton, 2004; Wilson & Horton, 2002), albeit with a stem completion task.

Target completion for critical test as a function of study condition

Next, an analysis of target completion data for the critical test for semantically and nonsemantically studied items was conducted. As expected, semantic study led to an increase in target completions for the explicit group, $t(31) = 4.90$. An analysis of variance conducted for the PDP group across inclusion and exclusion conditions showed a significant interaction of depth and condition, $F(1,30) = 23.77$, $MSE = 2.33$, suggesting that semantic encoding led the inclusion group to more accurately include items and the exclusion group to more accurately exclude items relative to nonsemantic encoding. A test of least significant differences indicated that semantic study led to an increase in target completions relative to nonsemantic study in the inclusion condition (LSD = 6.55) but not the exclusion condition (LSD = .157), a result that has been found in other studies (e.g., Mecklenbräuker et al., 1996). An analysis of variance indicated that semantic study did not lead to an increase in target completions for the speeded and implicit groups, as indicated by the lack of a main effect of group and a nonsignificant interaction, both F's < 1. This result has been obtained before (Richardson-Klavehn & Gardiner, 1998) but opposite results have also been found (Horton et al., 2004; Keane et al., 1997; McBride & Shoudel, 2003; Mulligan et al., 1999). However, it should be noted that target completion rates were relatively low in comparison to those found in other studies, possibly lowering the chance for detecting differences in target completion rates between the two encoding conditions. Alternatively, given the RT data reported earlier indicating that the speeded group used automatic retrieval, and given the finding that the explicit and PDP groups retrieved more target items than the speeded and implicit groups, and took longer to do so, it is possible that depth effects do not occur in automatic retrieval for conceptual tests of category exemplar generation when low frequency target exemplars are used and when subjects are only exposed to one exemplar per category during encoding.

Response times for target exemplars as a function of study condition

The means and standard deviations of RTs and number of subjects included are reported in Table 1. An analysis was conducted to assess whether semantic study led to faster RTs for target items compared to nonsemantic study for the speeded and implicit groups. Subjects who responded with fewer than two target items in the critical test were removed from the analysis. A significant main effect of depth was found, $F(1, 31) = 5.83$, $MSE = 223275$. No other significant effects were found. The results indicate that semantic encoding decreased RTs for target items compared to nonsemantic encoding for the speeded and implicit groups, a result that has been found in other studies (Horton et al., 2004). Thus, although the target completion data indicate no depth effects for subjects in the speeded and implicit groups, the fact that semantic study mediated faster RTs than the nonsemantic condition implies that there are some depth effects. However, caution should be exercised when interpreting these results because of the small number of subjects from the speeded nonsemantic study condition who were included in the analysis.

Conscious and automatic estimates in PDP group as a function of study condition

Memory estimates are reported in Table 5. Estimates of conscious retrieval were computed using the formula C = I – E, where I and E refer to target completion rates on the inclusion and exclusion tasks, respectively. As expected, conscious estimates for semantic study were higher than for nonsemantic study in the PDP group, $t(30) = 3.19$. This result parallels results found previously (Horton et al., 2004; Jacoby et al., 1993). Estimates of automatic retrieval were computed using the formula $A = \frac{E}{1-C}$, where C refers to the conscious estimate. When subjects with exclusion scores equal to zero were included in the analysis, a *t*-test revealed no differences in the automatic estimate for semantic or nonsemantic study in the PDP group, $t(30) = .50$, $SEM = .42$. However, a number of subjects had exclusion test scores of zero (and thus had automatic estimates of zero). In this case, the automatic estimate contributing to performance is underestimated since the independence assumption – the assumption that both conscious and automatic memory independently contribute to performance – is violated. Thus, it is common practice to remove scores for subjects with exclusion scores equal to zero from analyses conducted on automatic estimates (Jacoby, 1998). Removing subjects with perfect exclusion scores creates an underestimate in automatic retrieval, because a perfect exclusion score (or more generally, an exclusion score below baseline) implies that subjects are rejecting exclusion responses on familiarity (i.e., automatic retrieval) and recollection (i.e., conscious retrieval) rather than strictly recollection, as specified by the direct retrieval instructions. However, when these subjects were removed, a *t*-test again revealed no differences in the automatic estimates for semantic and nonsemantic study, $t(15) = .26$, $SEM = .70$. Thus, as hypothesized, automatic estimates for the PDP group revealed no depth effect, both when exclusion scores were included and when they were removed. Numerically, automatic estimates increased when subjects with exclusion scores equal to zero were removed from the analysis. This replicates results found previously (Horton et al., 2004; Jacoby et al., 1993).

Comparison of automatic estimates from PDP group and speeded group

Automatic estimates obtained for the PDP group were compared to automatic estimates obtained from target completion data for the speeded group. It was hypothesized that automatic estimates derived from PDP would produce an underestimate in automatic retrieval compared to estimates derived from target completion data in the speeded group, assuming that conscious estimates were measurably above zero (Wilson & Horton, 2002). An analysis of variance showed no differences for automatic estimates between the PDP and speeded group, $F(1, 57) = 1.98$, $MSE = 0.92$. No differences were found between automatic estimates for the PDP and implicit groups either, $F(1, 58) = 3.49$, $MSE = 0.89$. However, a comparison of proportion conscious estimates (computed by dividing the conscious estimate by the total number of items in each condition) indicated that the conscious estimates were not above zero (semantic, .19; nonsemantic, .04). When conscious estimates are low, the contamination of automatic estimates by conscious retrieval should also be low. This result has been found in other studies (Horton et al., 2004; Wilson & Horton, 2002). Thus, the finding that automatic estimates obtained using the PDP group mimicked those obtained using the speeded group is not surprising.

Target completion rates as a function of awareness

Target completion rates as a function of awareness are reported in Table 6. Commencing the completion of the critical test, subjects who cited that they were aware that some responses they generated were studied items were deemed test aware. A *t*-test revealed that subjects in the speeded and implicit groups who cited awareness (n = 37) responded with significantly more target exemplars than subjects who did not cite awareness (n = 28), $t(63) = 3.63$, $SEM = .32$. Importantly, all of the subjects included in this analysis also denied the use of intentional retrieval strategies, claiming instead that the target responses they generated were simply the first responses that came to mind upon exposure to the category label in the critical test. However, greater awareness was evident in the semantic than in the nonsemantic condition, as expected.

Target completion rates as a function of depth and awareness

In order to assess whether depth effects occurred as a product of awareness, an analysis of variance was conducted to compare target completion rates for aware and unaware subjects in the speeded and implicit groups. Aware subjects exhibited a higher target completion rate than unaware subjects, $F(1, 57) = 9.97$, $MSE = 1.57$. The interaction between awareness and depth was nonsignificant, indicating that aware and unaware subjects showed equivalent depth effects. Contrary to this result, Mace (2003) found a marginal effect of depth for the test aware ($p = .065$). The absence of an interaction in the present study suggests that despite awareness, subjects were not using conscious retrieval strategies to respond on the critical category exemplar test. If subjects used conscious retrieval strategies when they were aware, aware subjects should have generated more targets following semantic study than the unaware subjects, just as subjects who are instructed to use conscious retrieval strategies show depth effects. Rather, the awareness that occurred was likely that of unintentional conscious awareness, and thus, the higher target completion rate present in the aware group was what triggered their awareness. Mace found a reliable interaction between awareness and condition when only the second half of the test was analyzed, when all subjects indicated test awareness.

Only two subjects reported using intentional retrieval in the critical test, so the hypothesis that subjects using intentional retrieval in the speeded and implicit groups would have higher target completion rates compared to those subjects who did not claim to use intentional retrieval strategies was not explored.

Discussion

Evidence for automatic retrieval

The speeded response technique has provided pure estimates of automatic retrieval in perceptual memory tasks (Horton et al., 2001; Wilson & Horton, 2002). The RT data reported here extend this finding by showing that the speeded response technique also provides pure estimates of automatic retrieval in conceptual tasks, where some researchers have argued that conscious contamination is more likely (Butler & Berry, 2001). Subjects in a speeded group were encouraged to generate category exemplars using strictly automatic retrieval by presenting a practice test that did not allow responding with previously studied items and by encouraging speed of responding. As in previous studies, performance of the speeded group was compared to that of both an explicit group for whom conscious retrieval was required and a baseline condition for whom conscious retrieval was not possible (since category labels could not be completed with studied items).

The time to generate category exemplars on the critical category exemplar test was reliably longer in the explicit group than in either the speeded group or the baseline group, which did not differ. The longer RTs in the explicit group compared to the speeded and baseline groups confirm that conscious retrieval takes longer to complete than automatic retrieval (Richardson-Klavehn & Gardiner, 1995, 1998). The average RTs in the speeded group did not differ from those in the baseline group and were faster than those in the explicit group during the critical category exemplar test, indicating that subjects in the speeded group did not switch to a conscious retrieval strategy, but instead used an automatic retrieval strategy when presented with category labels for which exemplars were previously studied.

Comparisons of automatic estimates

With evidence that the speeded group provided a relatively pure measure of automatic retrieval, target completions for the speeded group were compared to those from an implicit group and with automatic estimates obtained from a PDP group. The lack of a difference in the target completion rates of the speeded and implicit groups suggests that the implicit group also used automatic retrieval on the critical category completion test. Estimates of automatic retrieval obtained in the PDP group did not differ from those obtained from the speeded group. Notably, conscious estimates derived from the PDP group were low (.19), and therefore the lack of difference in the automatic estimates of the PDP group compared to those of the speeded and implicit groups is expected. Other studies that have spawned lower conscious retrieval estimates through the use of long study and test lists (Horton et al., 2004), divided attention (Horton et al., 2004), or long delays between study and critical tests (C = .19, Wilson & Horton, 2002) have found the same result. By contrast, when conscious estimates are high, as occurs with full attention, (Horton et al., 2004) or with short delays between study and critical tests (C = .55, Wilson & Horton, 2002), PDP tends to underestimate automatic retrieval relative to automatic estimates obtained from the speeded group.

In the current study, inclusion and exclusion instructions were designed to follow the prescriptions of PDP, whereby the category label was first used to recall an exemplar from the study phase, and then subjects output a response in accordance with either the inclusion or exclusion criteria. If subjects used a retrieval strategy other than direct retrieval, then automatic and conscious estimates obtained using PDP are incorrect (Jacoby, 1998; Bodner, Masson, & Caldwell, 2000). Recall that the use of a generate recognize strategy violates the assumption that automatic and conscious retrieval strategies operate independently to produce memory: Instead of trying to recall studied items, subjects generate responses using automatic retrieval and then perform recognition checks using conscious processes, thus making consciously recognized responses redundant with automatically retrieved responses. The use of this strategy also may lead subjects to reject automatically retrieved items on the basis of familiarity, when the independence assumption requires that retrieved items are rejected only if they are recalled as being a part of the study list. There are two signatures of a generate recognize strategy: Target completion rates are lower than baseline in the exclusion condition and a greater number of unstudied targets are generated in the inclusion compared to the exclusion condition (Bodner et al., 2000). These signatures were not found in the present study, likely because the target exemplars used were very low in frequency, and were rarely generated if they were not studied (only two of 32 baseline subjects responded with more than two target exemplars in the critical test). Thus, it would be difficult to detect higher target completion rates in the baseline condition compared to the exclusion condition or a higher frequency of unstudied targets in the inclusion condition. Indeed, studies that have found these signatures have used higher frequency items as target items. For example, Richardson-Klavehn and Gardiner (1998) used the most frequent response to a word stem when the word had not been previously studied as target items in an perceptual implicit stem completion task, and uncovered the signatures of a generate recognize strategy despite direct retrieval instructions.

It has been argued that even when the signs of a generate recognize strategy are not present it may still be employed. Bodner et al. (2000) argue that giving strict response criteria in the exclusion task, such that subjects are highly motivated to provide no response in the exclusion task if they are unable to think of a response other than the studied item, determines whether signatures of a generate recognize strategy appear. In the first experiment conducted by Bodner et al., subjects were encouraged to exclude studied items in the exclusion task by directing them to complete as many of the test items as possible but to "not use an earlier-presented word" (p. 272). Following these instructions, the signatures of a generate recognize strategy were present, despite giving direct retrieval instructions at test. Firstly, the proportion of perfect exclusion scores increased, with 11 out of 24 subjects demonstrating perfect performance, which led to higher baseline target completion rates in inclusion compared to the exclusion condition. Secondly, since mean exclusion scores were low, baseline subjects generated more targets than subjects in the exclusion condition (despite the fact that baseline subjects did not study any targets) such that automatic estimates obtained from exclusion scores were reliably below baseline. However, in a follow up study in which exclusion instructions encouraged subjects to exclude studied items more strongly than in the first study (directing subjects to respond with the first word that came to mind unless they were sure that the first word they generated had been previously studied; if subjects were not sure if the first item generated had been studied then they were to respond with that word), the signatures of a generate recognize strategy disappeared despite giving generate recognize instructions. Baseline target completion rates were the same in inclusion and exclusion conditions, and baseline subjects did not have higher target completions compared to exclusion subjects since the proportion of perfect exclusion scores substantially decreased.

In the present study, almost half of the subjects had perfect exclusion scores (15 out of 32), and the remaining subjects had very low exclusion scores. Subjects may also have been highly motivated to exclude studied items since the experimenter was present and since the experimenter reminded subjects of the instructions throughout the critical test if it appeared that the subject was not following instructions. However, signatures of a generate recognize strategy may not have occurred simply because the baseline frequency of target exemplars was extremely low, which may have prevented exclusion scores from falling below the baseline target completion rate. Thus, a future study that employs the use of higher frequency responses to category labels may make the signatures of a generate recognize strategy appear, as it has in other studies (Richardson-Klavehn & Gardiner, 1998).

The difficulty of the inclusion and exclusion tasks themselves may also motivate subjects in the PDP condition to use a retrieval strategy other than direct retrieval. Recall that subjects were not told during the study phase that their memory for the stimuli would be tested later. Thus, upon exposure to the critical trial, subjects invariably found that they were unable to recall many studied items. This frustration may have motivated subjects to ignore the direct retrieval instructions and respond with studied items in the few instances in which they could retrieve such items, even on the exclusion task, since they were frustrated in their inability to recall studied words. Subjects often commented on their failure to follow instructions, indicating that there were times when they realized immediately after voicing a response in the exclusion task that the item had been studied and should have been excluded. Thus, other than the problematic assumptions associated with PDP, the difficulty of the PDP task itself and the frustration the task elicits for subjects can lead to inaccurate automatic retrieval estimates using PDP.

Comparisons of depth effects

The reliable depth effect shown in the explicit group is consistent with an abundance of previous literature (e.g., Craik et al., 1994; Horton et al., 2001; Richardson-Klavehn & Gardiner, 1998). The reliable depth effect found in the inclusion task and the absence of a depth effect in the exclusion task for the PDP group is also consistent with past literature (Horton & Vaughan, 1999; Mecklenbräuker et al., 1996). Thus, it appears that subjects in the PDP group are much better able to include previously studied items following semantic compared to nonsemantic encoding, but they are no better able to exclude items as previously studied following semantic encoding. However, the lack of a depth effect for the exclusion task likely occurred because the very low exclusion scores obtained following nonsemantic study made it difficult to detect further decreases in target completions following semantic study.

The lack of a depth effect in target completion rates for the speeded and implicit groups implies that automatic conceptual priming does not benefit from semantic encoding, at least when very low frequency category exemplars are used. Because semantic encoding increased conscious retrieval in the explicit and the PDP test, semantic encoding should have increased the influence of conscious retrieval, and thus target completion rates, in the speeded and implicit groups if the conceptual priming was controlled by conscious processes. Thus, the lack of a depth effect in the speeded and implicit groups may be interpreted as further evidence that priming in these groups occurs through automatic retrieval.

Although the lack of a depth effect in the present study supports the hypothesis that conceptual priming occurs through automatic retrieval, inherent problems associated with shallow processing conditions of many studies may be the reason why depth effects have appeared in many tests of implicit memory. Specifically, depth effects that have been found in implicit memory tests may result from a failure of subjects in nonsemantic conditions to lexically encode study items (Richardson-Klavehn & Gardiner, 1998). Much evidence is available to warrant the claim that lexical processing is a necessary precursor to priming. For instance, when subjects are only required to make decisions about letter positions within words, and are not required to verbalize the word, or are not given sufficient exposure to the word before making an encoding decision, priming does not occur (Hayman & Jacoby, 1989). Many studies that have uncovered a priming advantage for semantically encoded words over nonsemantically encoded words have used nonsemantic encoding similar to this, whereby processing the stimulus as a lexical unit – also called whole word processing – was not required (e.g., Bergerbest & Goshen-Gottstein, 2002; Mulligan et al., 1999; Toth et al., 1994; Vaidya et al., 1997). Thus, advantages of semantic encoding may have appeared because, unlike in semantic conditions where lexical processing was required in order to make semantic decisions, lexical processing was not required in nonsemantic conditions, and in turn, subjects may not have processed the lexical unit, which would make priming for those items very unlikely. Indeed, subjects who are not required to process the lexical unit during nonsemantic study have often expressed this lack of attention to the actual words during debriefing (Bergerbest & Goshen-Gottstein, 2002; Richardson-Klavehn & Gardiner, 1998).

In a study designed to test the lexical processing hypothesis, Richardson-Klavehn and Gardiner (1998) included two nonsemantic encoding conditions and a semantic encoding condition. The first nonsemantic task was graphemic, whereby subjects counted enclosed spaces of letters in a word. This graphemic nonsemantic task could be completed without processing the word as a lexical unit. The second nonsemantic task was phonemic, whereby subjects counted word syllables, a task that is argued to be easiest to complete by pronouncing the entire word (due to the complex nature of spelling to sound correspondence) thereby promoting lexical access. The semantic task involved both lexical and conceptual processing, and required subjects to rate pleasantness of the object described by a word. The predictions of the lexical processing hypothesis were supported – priming levels were equivalent in the semantic and phonemic study conditions, but there was less priming found in the graphemic study condition. A follow up study verified that the lack of lexical processing in the graphemic study condition was the reason for the lack of priming: Adding a lexical processing requirement to the task (having subjects in all three groups verbalize each word prior to performing the other task) re-established priming in the graphemic condition, at levels equal to those of the other two groups. A study conducted by Vaidya and Gabrieli (2000) also showed no depth effects when a shallow processing condition ensured lexical processing in a category exemplar generation task. In this study, target completion rates were not dependent on encoding task when the shallow processing task involved word reading, a task that ensured that lexical processing occurred, and the deep processing task involved word categorization.

Given these results in support of the lexical processing hypothesis, the finding of no depth effects for the implicit and speeded groups is not surprising. Subjects in the current study did lexically process the study words, since they were required to verbalize the words before commencing the nonsemantic task. However, it should be noted that depth effects are sometimes found even when lexical access to target items is required during study. For example, a study conducted by Horton et al. (2004) found a depth effect in a perceptual implicit stem completion task, such that generating associates for targets during study led to higher target completions than counting consonants for targets following target verbalization. Importantly, RT evidence from speeded, baseline, and explicit groups was like that obtained in the present study, and thus supported the use of an automatic retrieval strategy. Thus, although the depth effect was substantially reduced compared to a condition in which lexical access of study items was not assured (since subjects did not have to verbalize words before making encoding decisions), the presence of a depth effect implies that depth effects can occur through automatic retrieval. However, the overall low target completion rates in the present study may have made differences in target completion rates for semantic and nonsemantic study too small to detect. The use of higher frequency target exemplars may have increased the target completion rates, in turn increasing the automatic estimates obtained for the speeded and implicit groups.

The relatively low target completion rates compared to other studies that used the speeded response technique may have created a situation in which depth effects were difficult to detect. The design of the current study may not have been optimal for creating depth effects. Depth effects may be more likely to occur if subjects are exposed to more exemplars per category, since this increases the salience of categorical information (Mace, 2003; Mulligan et al., 1999). As well, at least one study has found stronger depth effects using a category exemplar generation task when the encoding condition was presented in a blocked as opposed to a random fashion in the study list, such that several exemplars from the same category were presented in succession as opposed to in a random order during encoding (Mulligan et al., 1999). In the critical test of this study, subjects were required to generate several exemplars to a given category label. A future study that employs both higher frequency target items and has a study list that presents more than one exemplar per category in a blocked fashion, and has a lexical processing requirement in the shallow processing condition may very well find a depth effect for automatic estimates on automatic estimates for speeded and implicit groups.

Priming effects for RTs

Response times for the speeded group should be at least as fast as those of the baseline group for designated target items if RT priming occurs through automatic retrieval (Horton et al., 2001; Richardson-Klavehn & Gardiner, 1998). This result was obtained in the present study, providing further evidence that the speeded group used an automatic retrieval strategy to retrieve exemplars in the critical category exemplar test. Automatic priming effects did not bring previously studied responses to mind more readily in the speeded group, as indicated by the finding that the speeded and baseline groups provided exemplars equally as fast when critical category labels were given. It is possible that priming effects on RTs may not be strong enough to observe, at least with the present procedures. Other studies that have employed the use of the speeded response technique also did not find priming effects on RTs (Horton et al., 2001). Although the absence of RT priming does not undermine the conclusion that subjects in the speeded group were responding based on automatic retrieval, other possible reasons for the lack of a priming effect for RTs in the speeded group will be examined.

At least one other study has shown faster RTs in a speeded group compared to the baseline group (Richardson-Klavehn & Gardiner, 1998). However, there is at least one reason why this study would be more likely to find priming for RTs compared to the current study. The study that revealed this effect used the most frequent response for a word stem (the task used was an implicit perceptual stem completion test). Thus, baseline subjects responded more often with the target item than in the current study (an average of 30% of the time as opposed to less than 1%) possibly allowing priming of RTs for studied items to be easier to detect. It is also possible that Richardson-Klavehn and Gardiner's inclusion of the stem of the target word in the critical test made the priming effect on RTs more prominent, since the inclusion of a specific part of the target word may better trigger a studied response compared to giving a cue that allows for less constrained responding, such as giving only a category label (without, for example, the first letter of the target response) as in the present study. Thus, it is possible that if more specific response criteria had been used in the critical test of the current study, priming for RTs would have been found. However, the method of supplying subjects with more specific response criteria on the critical test slowed down responding for baseline and implicit subjects (median RTs of 3 and 2.3 seconds respectively) in the Richardson-Klavehn and Gardiner (1998) study as well as in a pilot study conducted prior to the present experiment. Because the primary goal of the present study was to examine automatic retrieval in memory, it is important to assure that responses in the critical test can occur quickly. The more time a subject has to respond, the greater the opportunity for conscious retrieval strategies to contribute to memory performance. Indeed, in the pilot work, RTs for the speeded group and the baseline group were no different from those of the explicit group, suggesting that all three groups may have been using conscious retrieval strategies.

Priming effects for RTs as a function of depth

Although the target completion data indicate that deep encoding did not mediate depth effects in the speeded and implicit groups, the RT data indicate that there was a priming effect for semantically encoded stimuli relative to nonsemantically encoded stimuli, with semantic study mediating faster RTs than nonsemantic study for both groups. This finding may further support the hypothesis that the reason why a depth effect was not obtained for target completion rates was due to the overall low target completion rates (which likely occurred because of low frequency of target exemplars) and not because depth effects do not occur with automatic retrieval.

Awareness, target completion, and depth effects

Results from a questionnaire given following the completion of the test phase revealed that more subjects became aware that some responses they used in the critical test had been previously studied following semantic study than following nonsemantic study. This result replicates previous findings (Mace, 2003) and can be explained by the theory of transfer appropriate processing: Semantic encoding of a study item should create a better memory for that item in a memory test that demands semantic processing. Subjects in the speeded and implicit groups who cited awareness on the posttest questionnaire (but claimed that they did not use intentional retrieval strategies) also had higher target completion rates compared to subjects who did not cite awareness. However, aware and nonaware subjects generated equal numbers of targets following semantic and nonsemantic study. Thus, although awareness led to overall greater target completion rates, awareness did not mediate depth effects. Contrary to this result, Mace (2003) found a marginal effect of depth for the test aware ($p = .065$). As in the current study, test aware subjects in Mace's study were those subjects who cited awareness of the study test phase relationship but who denied the use of intentional retrieval strategies on the critical test.

Although there was not a significant depth-awareness interaction in the current study, the numerical trend in the data suggests that if there had been higher target completion rates, a depth awareness interaction may have been found. Mean target completion rates were higher for those subjects who cited awareness. Mace attributed the finding that depth effects were only found for aware subjects to involuntary aware memory: That is, depth effects were enhanced by involuntary aware memory, but were not contaminated by explicit retrieval. However, another study has shown that depth effects remain robust after subjects who claim both the use of intentional retrieval strategies and awareness are removed from the analysis when the exemplars in the study list were organized in a blocked fashion (Mulligan et al., 1999). Thus, depth effects are not always enhanced through involuntary awareness. A future study that finds depth effects will be able to explore the relationship between awareness and depth effects more fully.

Is awareness more likely to occur during conscious retrieval?

It has been proposed that recognition during short response deadlines can only occur through automatic familiarity process, a relatively faster process than consciously controlled recollection process that can potentially control recognition for longer response deadlines (Gardiner, Ramponi, & Richardson-Klavehn, 1999; Jacoby, 1991). In a study partially designed to explore differences in recognition memory performance whereby subjects performed a recognition test in either a short (500 ms) or long (1500 ms) deadline, it was found that subjects could determine whether an item had been previously studied equally well for short and long response deadlines (Gardiner et al., 1999; see also Toth, 1996). Equivalent performance on the recognition test implies that conceptual processing can operate automatically. As well, depth effects were equal for both the short and long response deadlines; that is, semantic study led to an increase in recognition scores for both the short and long response deadlines.

The results of the current study can be likened to these results in a couple of ways. In the Gardiner et al. (1999) study, more subjects recognized items as previously studied following a semantic encoding phase, just as occurred in the present study. The emphasis on responding quickly (although the mean RTs were not quite as fast as they were in the Gardiner et al. study) also implies that this conceptual processing operated automatically. In the current study, RT data from the speeded, implicit, and baseline groups provided empirical evidence that the increase in awareness following semantic study for the speeded and implicit groups was not a side effect of conscious retrieval. Instead, awareness may be as equally likely to arise during automatic retrieval processes as it is in conscious retrieval processes. Recall that the awareness assumption postulated by Jacoby maintains that awareness that an item has been studied can occur only following conscious retrieval and never following automatic retrieval. However, Richardson-Klavehn et al. (1994) argue that the automatic retrieval of a studied item may also trigger the conscious awareness of the items prior appearance. If this is the case, then this is further evidence to dispute the assumption that automatic and conscious retrieval processes operate independently, in that the conscious awareness that an item has been studied can not only occur following conscious retrieval, but also following automatic retrieval. Results of the current study and the Gardiner et al. study support the hypothesis that involuntary conscious awareness can occur following automatic retrieval.

In conclusion, the results of the current study support the hypothesis that conceptual priming can occur automatically, despite the fact that researchers have claimed that conscious contamination of automatic retrieval is more likely to occur in conceptual compared to perceptual implicit memory tests. Although there were no depth effects for the automatic estimates in the current study, it is possible that using target exemplars with a higher baseline frequency would provide an opportunity for this effect to be realized, as it has in past research (Horton et al., 2004). Automatic estimates obtained from PDP should also underestimate the contribution of automatic retrieval relative to estimates obtained using the speeded response technique if conscious estimates increase through the use of higher frequency target exemplars, as it has in past research (Richardson-Klavehn & Gardiner, 1998). Importantly, this study illustrates that priming in conceptual memory tests can function through automatic retrieval. Future studies should be designed with higher frequency target items to provide a better test of the preceding hypotheses.

References

Battig, W.F., & Montague, W.E. (1969). Category norms for verbal items in 56 categories: A replication and extension of the Connecticut category norms. *Journal of Experimental Psychology Monograph, 80*, 1-46.

Bergerbest, D., & Goshen-Gottstein, Y. (2002). The origins of levels-of-processing effects in a conceptual test: Evidence for automatic influences of memory from the process-dissociation procedure. *Memory & Cognition, 30,* 1252-1262.

Blaxton, T.A. (1989). Investigating dissociations among memory measures: Support for a transfer appropriate processing framework. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 657-668.

Bodner, G.E., Masson, M.E., & Caldwell, J.I. (2000). Evidence for a generate-recognize model of episodic influences on word-stem completion. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*, 267-293.

Bower, G.H. (1996). Reactivating a reactivation theory of implicit memory. *Consciousness and Cognition, 5*, 27-72.

Bowers, J.S., & Turner, E.L. (2003). In search of perceptual priming in a semantic classification task. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*, 1248-1255.

Brown, A.S., & Mitchell, D.B. (1994). A reevaluation of semantic versus nonsemantic processing in implicit memory. *Memory & Cognition, 22*, 533-541.

Bruce, V., Carson, D., Burton, A.M., Ellis, A.W. (2000). Perceptual priming is not a necessary consequence of semantic classification of pictures. *The Quarterly Journal of Experimental Psychology, 53A*, 289-323.

Buchner, A., Erdfelder, E., & Vaterrodt-Plünnecke, B. (1995). Toward an unbiased measurement of conscious and unconscious memory processes within the process dissociation procedure. *Journal of Experimental Psychology: General, 124*, 137-160.

Butler, L.T., & Berry, D.C. (2001). Implicit memory: Intention and awareness revisited. *Trends in Cognitive Sciences*, *5*, 192-197.

Carlesimo, G.A. (1994). Perceptual and conceptual priming in amnesiac and alcoholic patients. *Neuropsychologia, 32*, 903-921.

Casey, P.J. (1992). A reexamination of the roles of typicality and category dominance in verifying category membership. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*, 823-834.

Craik, F.I.M. (1983). On the transfer of information from temporary to permanent memory. *Philosophical Transactions of the Royal Society of London: Series B, 302*, 155-158.

Craik, F.I.M., Moscovitch, M., & McDowd, J.M. (1994). Contributions of surface and conceptual information to performance on implicit and explicit memory tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*, 864-875.

Curran, T., & Hintzman, D.L. (1995). Violations of the independence assumption in process dissociation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 531-547.

Curran, T., & Hintzman, D.L. (1997). Consequences and causes of correlations in process dissociation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*, 496-504.

de Houwer, J. (1997). Differences in intentional retrieval during inclusion and exclusion tasks. *Memory, 5*, 379-400.

Gardiner, J.M., Ramponi, C., & Richardson-Klavehn, A. (1999). Response deadline and subjective awareness in recognition memory. *Consciousness and Cognition, 8*, 484-496.

Geraci, L., & Rajaram, S. (2004). The distinctiveness effect in the absence of conscious recollection: Evidence from conceptual priming. *Journal of Memory and Language, 51*, 217-230.

Gold, D., Beauregard, M., Roch Lecours, A., & Chertkow, H. (2003). Semantic category differences in cross-form priming. *Journal of the International Neuropsychological Society, 9*, 796-805.

Graf, P., & Mandler, G. (1984). Activation makes words more accessible but not necessarily more retrievable. *Journal of Verbal Learning and Verbal Behavior, 23*, 553-568.

Graf, P., & Schacter, D.L. (1985). Implicit and explicit memory for new associations in normal amnesiac subjects. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 11*, 501-518.

Graf, P., Squire, L.R., & Mandler, G. (1984). The information that amnesiac patients do not forget. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10*, 164-178.

Graf, P., Shimamura, A.P., & Squire, L.R. (1985). Priming across modalities and priming category levels: Extending the domain of preserved function in amnesia. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 11*, 385-395.

Hamilton, M., & Rajaram, S. (2001). The concreteness effect in implicit and explicit memory tests. *Journal of Memory and Language, 44*, 96-117.

Hamann, S.B. (1990). Level-of-processing effects in conceptually driven implicit tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*, 970-977.

Hayman, C.A.G., & Jacoby, L.L. (1989). Specific transfer as a measure of processing in the word-superiority paradigm. *Memory & Cognition, 17*, 125-133.

Horton, K.D., & Vaughan, D. (1999). Analyzing estimates of automatic and conscious retrieval. *Behaviour Research Methods, Instruments, & Computers, 31*, 347-352.

Horton, K.D., Wilson, D.E., & Evans, M. (2001). Measuring automatic retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*, 958-966.

Horton, K.D., Wilson, D.E., Vonk, J., Kirby, S.L., & Nielsen, T. (2004). A comparison of implicit memory and the process dissociation procedure with the speeded response procedure: Depth of processing and attentional effects on automatic retrieval. Manuscript submitted.

Isingrini, M., Vazou, F., & Leroy, P. (1995). Dissociation between implicit and explicit memory tests: Effects of age and divided attention on category exemplar generation and cued recall. *Memory & Cognition, 23*, 462-467.

Jacoby, L.L. (1983). Remembering the data: Analyzing interactive processes in reading. *Journal of Verbal Learning and Verbal Behavior, 22*, 485-508.

Jacoby, L.L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language, 30,* 513-541.

Jacoby, L.L., & Dallas, M. (1981). On the relationship between autobiographical memory and perceptual learning. *Journal of Experimental Psychology: General, 3*, 306-340.

Jacoby, L.L., Toth, J.P., & Yonelinas, A.P. (1993). Separating conscious and unconscious influences of memory: Measuring recollection. *Journal of Experimental Psychology: General, 122*, 139-154.

Jacoby, L.L. (1998). Invariance in automatic influences of memory: Toward a users guide for the process-dissociation procedure. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*, 3-20.

Keane, M.M., Monti, L.A., Fleischman, D.A., Gabrieli, J.D.E., Cantor, J.M., & Noland, J.S. (1997). Intact and impaired conceptual memory processes in amnesia. *Neuropsycholgy, 11*, 59-69.

Kiran, S., & Thompson, C.K. (2003). Effect of typicality on online category verification of animate category exemplars in aphasia. *Brain and Language, 85*, 441-450.

Light, L.L., Prull, M.W., & Kennison, R.F. (2000). Divided attention, aging, and priming in exemplar generation and category verification. *Memory & Cognition, 28*, 856-872.

Mace, J.H. (2003). Study–test awareness can enhance priming on an implicit memory task: Evidence from a word completion task. *American Journal of Psychology, 116,* 257-279.

MacLeod, C.M., & Masson, M.E.J. (1997). Priming patterns are different in masked word identification and word fragment completion. *Journal of Memory and Language 36*, 461-483.

MacLeod, C.M., & Masson, M.E.J. (2000). Repetition priming in speeded word reading: Contributions of perceptual and conceptual processing episodes. *Journal of Memory and Language, 42*, 208-228.

McBride, D.M., & Dosher, B.A. (2002). A comparison of conscious and automatic memory processes for picture and word stimuli: A process dissociation analysis. *Consciousness and Cognition, 11*, 423-460.

McBride, D.M., & Shoudel, H. (2003). Conceptual processing effects on automatic memory. *Memory & Cognition, 31*, 393-400.

Mecklenbräuker, S., Wippich, W., & Mohrhusen, S.H. (1996). Conscious and unconscious influences of memory in a conceptual task: Limitations of a process-dissociation procedure. *Swiss Journal of Psychology, 55*, 34-48.

Mulligan, N.W., Guyer, G.S., & Beland, A. (1999). The effects of levels-of-processing and organization on conceptual implicit memory in the category exemplar production test. *Memory & Cognition, 27*, 633-647.

Mulligan, N.W., & Brown, A.S. (2003). Attention and implicit memory. In L. Jimenez (Ed.), *Attention and implicit learning*. (p.297-334). Amsterdam: John Benjamins Publishing Company.

Mulligan, N.W., & Hartman, M. (1996). Divided attention and indirect memory tests. *Memory & Cognition, 24*, 453-465.

Mulligan, N.W. (1997). Attention and implicit memory tests: The effects of varying attentional load on conceptual priming. *Memory & Cognition, 25*, 11-17.

Mulligan, N.W. (2002). The effects of generation on conceptual implicit memory. *Journal of Memory and Language, 47*, 327-342.

Neely, J.H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited capacity attention. *Journal of Experimental Psychology: General, 106*, 226-254.

Park, S.M., & Gabrieli, J.D.E. (1995). Perceptual and conceptual components in implicit memory for pictures. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 1583-1594.

Reingold, E.M., & Toth, J.P. (1996). Process dissociations versus task dissociations: A controversy in progress. In G. Underwood (Ed.), *Implicit cognition* (pp. 159-202). New York: Oxford University Press.

Richardson-Klavehn, A., Clarke, A.J.B., & Gardiner, J.M. (1999). Conjoint dissociations reveal involuntary "perceptual" priming from generating at study. *Consciousness and Cognition, 8*, 271-284.

Richardson-Klavehn, A., & Gardiner, J.M. (1995). Retrieval volition and memorial awareness in stem completion: An empirical analysis. *Psychological Research, 57*, 166-178.

Richardson-Klavehn, A., & Gardiner, J.M. (1996). Cross-modality priming in stem completion reflects conscious memory, but not voluntary memory. *Psychonomic Bulletin & Review, 3,* 238-244.

Richardson-Klavehn, A., & Gardiner, J.M. (1998). Depth-of-processing effects on priming in stem completion: Tests of the voluntary contamination, conceptual processing, and lexical processing hypotheses. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*, 593-609.

Richardson-Klavehn, A., Gardiner, J.M., & Java, R.I. (1994). Involuntary conscious memory and and the method of opposition. *Memory, 2*, 1-29.

Richardson-Klavehn, A., Gardiner, J.M., & Ramponi, C. Level of processing and the process-dissociation procedure: Elusiveness of null effects on estimates of automatic retrieval. *Memory, 10*, 349-364.

Roediger, H.L., III, & McDermott, K.B. (1993). Implicit memory in normal human subjects. In H. Spinnler & F. Boller (Eds.), *Handbook of neuropsychology* (Vol. 8, pp. 63-131). New York: Elsevier.

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

Schacter, D.L., Bowers, J., & Booker, J. (1989). Intention, awareness, and implicit memory: The retrieval intentionality criterion. In S. Lewandowsky, J.C. Dunn, & K. Kirsner (Eds.), *Implicit memory: Theoretical issues* (pp. 47-65). Hillsdale, NJ: Erlbaum.

Schmitter-Edgecombe, M. (1999). Effects of divided attention on perceptual and conceptual memory tests: An analysis using a process dissociation approach. *Memory & Cognition, 27*, 512-525.

Shapiro, S.I., & Palermo, D.S. (1970). Conceptual organization and class membership: Normative data for representatives of 100 categories. *Psychonomic Monograph Supplements, 3*, 105-127.

Srinivas, K. (1993). Perceptual specificity in nonverbal priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*, 582-602.

Srinivas, K., & Roediger, H.L., III. (1990). Classifying implicit memory tests: Category association and anagram solution. *Journal of Memory and Language, 29*, 389-412.

Toth, J.P. (1996). Conceptual automaticity in recognition memory: Levels-of-processing effects on familiarity. *Canadian Journal of Experimental Psychology, 50*, 123-138.

Toth, J.P., Reingold, E.M., & Jacoby, L.L. (1994). Toward a redefinition of implicit memory: Process dissociations following elaborative processing and self generation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*, 290-303.

Vaidya, C.J., & Gabrieli, J.D. (2000). Picture superiority in conceptual memory: Dissociative effects of encoding and retrieval tasks. *Memory & Cognition, 28*, 1165-1172.

Vaidya, C.J., Keane, M.M., Gabrieli, J.D.E., & Monti, L.A. (1995). Perceptual and conceptual memory processes in global amnesia. *Neuropsychology, 9*, 580-591.

Vaidya, C.J., Gabrieli, J.D.E., Keane, M.M., Monti, L.A., Gutierrez-Rivas, H., & Zarella, M.M. (1997). Evidence for multiple mechanisms of conceptual priming on implicit memory tests. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*, 1324-1343.

Vonk, J., & Horton, K.D. (2004). The role of conscious and automatic retrieval processes in directed forgetting. Submitted.

Vriezen, E.R., Moscovitch, M., & Bellos, S.A. (1995). Priming effects in semantic classification tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 933-946.

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

Warrington, E.K., & Weiskrantz, L. (1968). Amnesic syndrome: Consolidation or retrieval? *Nature, 228*, 628-630.

Weldon, M.S. (1991). Mechanisms underlying priming on perceptual tests. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17*, 526-541.

Weldon, M.S., & Coyote, K.C. (1996). Failure to find the picture superiority effect in implicit conceptual memory tests. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 670-686.

Wilson, D.E., & Horton, K.D. (2002). Comparing techniques for estimating automatic retrieval: Effects of retention interval. *Psychonomic Bulletin & Review, 9*, 566-574.

Wippich, W., Melzer, A., & Mecklenbräuker, S. (1998). Picture or word superiority effects in implicit memory: Levels of processing, attention, and retrieval constraints. *Swiss Journal of Psychology, 57*, 33-46.

Wolters, G., & Prinsen, A. (1997). Full versus divided attention and implicit memory performance. *Memory & Cognition, 25*, 764-771.

Yonelinas, A.P., & Jacoby, L.L. (1994). Dissociations of processes in recognition memory: Effects of interference and of response speed. *Canadian Journal of Experimental Psychology, 48*, 516-534.

Table 1

Mean of Median Response Times (RTs, in ms) as a Function of Group (Speeded, Explicit, Baseline), Study Task (Semantic, Nonsemantic), and Test Type (Practice, Critical)

	Practice test	Critical test		
Group		Nonstudied	Semantic	Nonsemantic
Speeded				
M	1233	1152	1094	1553
SD	210	214	249	315
N	32	32	11	4
Explicit				
M	1235	3086	1708	2032
SD	233	211	438	1387
N	33	33	17	16
Baseline				
M	1199	1152		
SD	194	213		
N	32	32		

Note: RTs for the studied and nonstudied items on the critical category exemplar test are shown separately for the speeded and explicit groups. Median RTs for semantic and nonsemantic items for target items on the critical test do not include subjects who generated fewer than 2 targets.

Table 2

Mean of Median Response Times (RTs, in ms) for target items as a Function of Group (Speeded, Explicit, Baseline)

	RT for target item excluding subjects who generated at least 2 target exemplars	RT for target item including subjects who generated at least 1 target exemplar
Group		
Speeded		
M	1217	1347
SD	331	566
N	15	22
Explicit		
M	1858	1865
SD	1059	1012
N	30	33
Baseline		
M	1149	1554
SD	591	775
N	2	13

Table 3

Target Completion Rates on the Critical Category Exemplar Test as a Function of Group (Speeded, Explicit, Baseline, Implicit) and Encoding Task (Semantic, Nonsemantic, Nonstudied). Maximum Target Completions = 32.

	Encoding task		
Group	Semantic	Nonsemantic	Nonstudied
Speeded			
M	2.00	.94	.38
SD	1.51	.93	.61
Explicit			
M	7.41	3.07	.55
SD	3.12	1.73	.75
Baseline			
M			.47
SD			.62
Implicit			
M	2.06	1.88	.27
SD	1.43	1.36	.57

Table 4

Target Completion Rates for PDP Group as a Function of Study Condition (Semantic, Nonsemantic) and Test Type (Inclusion, Exclusion) on Critical Test. Maximum Target Completions = 16.

	Semantic	Nonsemantic
Inclusion		
M	4.06	1.56
SD	2.05	1.21
Exclusion		
M	1.06	1.00
SD	1.29	1.16

Note: n = 16 for all conditions

Table 5

Memory Estimates for the Process Dissociation Group (PDP) as a Function of Study Condition (Semantic, Nonsemantic). Standard deviations in parentheses

Memory Estimate	Semantic	Nonsemantic
Conscious	3.00	.56
	(2.56)	(1.67)
	n = 16	n = 16
Automatic	.64	.66
(exclusion score =	(1.43)	(.89)
0 included)	n = 16	n = 16
Automatic	1.13	1.31
(exclusion score =	(1.79)	(.84)
0 excluded)	n = 9	n = 8

Table 6

Target Completion Rates on the Critical Category Exemplar Test as a Function of Group (Speeded, Implicit), Encoding Task (Semantic, Nonsemantic), and Awareness (Aware, Nonaware)

	Encoding task			
Group	Semantic		Nonsemantic	
	Aware	Nonaware	Aware	Nonaware
Speeded				
M	2.36	1.20	1.17	.80
SD	1.57	1.10	.98	.92
N	11	5	6	10
Implicit				
M	2.46	.75	2.43	1.44
SD	1.40	.50	1.62	1.01
N	13	4	7	9

Appendix 1

Categories and Target Exemplars

Category	Target Exemplar	Category	Target Exemplar
Gardening Tool	Rake	Drug	Penicillin
Insect	Beetle	City	Rome
Alcoholic Beverage	Champagne	Planet	Uranus
Tree	Willow	Disease	Malaria
Measure of Weight	Ton	Kind of Restaurant	Bistro
Colour	Pink	Flower	Orchid
Fuel	Methane	Weapon	Grenade
Bird	Hawk	Carpenter's Tool	Crowbar
Vegetable	Zucchini	Precious Stone	Opal
Type of Crime	Larceny	Military Title	Admiral
Dessert	Pudding	Type of Nut	Pecan
Season	Autumn	Surgical Tool	Forceps
Article of Clothing	Jacket	Continent	Antarctica
Occupation	Nurse	Geometric Shape	Rhombus
Relative	Grandmother	Type of Movie	Drama
Reading Material	Journal	Religion	Judaism
Fruit	Plum	Reptile	Iguana
Article of Furniture	Stool	Unit of Time	Decade
Musical Instrument	Saxophone	Found in a Circus	Trapeze

Category	Target Exemplar	Category	Target Exemplar
Cosmetic	Mascara	Piece of Jewelry	Pendant
Ice Cream Flavour	Butterscotch	Measure of Distance	Millimeter
Eating Utensil	Chopsticks	Type of Footgear	Clogs
Electronic Device	Pager	Holiday	Labour Day
Part of a Bike	Handlebar	Type of Bread	Multi-grain
Fish	Tuna	Cleaning Instrument	Sponge
Electrical Appliance	Oven	Type of Exercise	Walking
Building Material	Plaster	Bakery Product	Muffin
Month of the Year	November	Camping Equipment	Lantern
Member of Royalty	Duchess	Piece of Cookware	Frying Pan
Foreign Country	Brazil	Vehicle	Train
Sports Equipment	Puck	Farm Animal	Goat
Water Sport	Canoeing	Grooming Device	Nail Clippers
Type of Building	Hospital	Farm Crop	Hay
Non-Alcoholic Beverage	Cocoa	Dairy Product	Sour Cream
Form of Communication	Sign Language	Photographic Equipment	Tripod

Study Items for Baseline Group

Study Item	
Cactus	Multiplication
Sitcom	Monday
Triceratops	Shakespeare
Number	Volts
Matches	Einstein
Night	Democracy
Shilling	Waterloo
Homework	Magnet
Cord	Diet
Altar	Highway
Trick	Pain
Party	Baby
Aristotle	Fee
Plug	George Bush
Picasso	Diaper
Multiplication	Election

Appendix 2

Questionnaire

As a final task in this experiment, please answer the following questions. The questions are designed to give the experimenter information on how you experienced the task. There are no right or wrong answers to the questions, but it is important that you be honest when giving your responses. This final task will take no more than a minute or 2 to complete.

1. Describe any characteristics at all that you noticed about the category labels or the words that you gave as responses in the final task.
2. Describe everything that you did to help you generate responses to the category labels in the final task.
3. What do you think was the purpose of the experiment?
4. When you were responding to the category labels, did you at any time find yourself generating words that were encountered during the first task when you made (pleasantness/vowel counting) judgments about words?
5. At any time when you were responding to the category labels, did you *deliberately attempt* to recall words for which you had previously gave (pleasantness/vowel counting) judgments?

Appendix 3

Removed Categories, Overlapping Responses, and Percentages of Responses Used Applicable to One or Both Categories for Speeded and Implicit Groups

Removed Categories	Examples of Overlapping Response	Percentage of Responses Used Applicable to One or Both Categories
Natural Disaster, Weather Phenomenon	Hurricane, Tornado	76%
School Subject, Language	English, Greek, French	69%
Item of Office Supplies, Writing Utensil	Pencil, Pen, Marker	38%
Facial Expression, Emotion	Sad, Happy	34%
Game, Sport	Soccer, Hockey	29%
Type of Dance, Type of Music	Jazz	26%
Part of a House, Part of a Car	Door, Window	22%
Part of a Body, Part of a Face	Nose, Eyes, Ears	16%
Breakfast Food, Type of Meat	Bacon	16%