Comparing techniques
for estimating automatic retrieval:
Effects of retention interval

DARYL E. WILSON and KEITH D. HORTON
Wilfrid Laurier University, Waterloo, Ontario, Canada

Results from implicit memory (IM) tasks suggest that automatic retrieval remains stable or decreases over time. In contrast, results from the process dissociation approach (PDA) suggest that automatic retrieval may actually increase over time. One explanation for these discrepant results is that performance on IM tasks is contaminated by controlled retrieval strategies, thereby overestimating automatic retrieval, particularly at short retention intervals, when controlled retrieval strategies are high. An alternative explanation is that automatic and controlled retrieval are positively correlated, rather than independent as assumed by the PDA. If so, the PDA would underestimate automatic retrieval, particularly when controlled retrieval strategies are high. Results from a speeded IM task suggest that a standard IM task provided an accurate estimate of automatic retrieval, whereas the PDA underestimated automatic retrieval at a short retention interval. This pattern of underestimation by the PDA supports the conclusion that automatic and controlled retrieval were positively correlated rather than independent.

A contentious issue in memory research is the degree to which automatic retrieval (the unintentional retrieval of information) and controlled retrieval (the intentional retrieval of information) contribute to memory. This paper focuses on procedures that have been used to estimate the influence of automatic retrieval on stem completion performance.

In contrast to an explicit memory (EM) task, in which subjects are instructed to intentionally retrieve studied information, subjects in a typical stem completion implicit memory (IM) task are instructed to complete stems (e.g., HOR__) with the first word that comes to mind. Priming on the IM task has commonly been attributed to automatic retrieval of studied information (Schacter, 1987) and has been assumed to provide a relatively pure index of automatic retrieval (e.g., Graf & Mandler, 1984).

This assumption of the process purity of IM tasks has been questioned by several researchers (Jacoby, 1991; Reingold & Merikle, 1988, 1990; Toth & Reingold, 1996), who have suggested that performance on some IM tasks may be affected by both automatic and controlled retrieval. For example, some subjects may become aware that they have completed some items on the IM task with studied

words and, as a result, elect to adopt a controlled retrieval strategy in an attempt to maximize performance.

The possibility that subjects may invoke controlled retrieval on an IM task limits the interpretation of performance. For example, if an experimental manipulation (e.g., depth of processing) affects IM performance, the manipulation may affect automatic retrieval, controlled retrieval, or both. If controlled retrieval contributes to IM performance, then to find that a variable such as depth of processing has an effect on IM tasks (e.g., Brown & Mitchell, 1994; Challis & Brodbeck, 1992), albeit a smaller effect than on EM tasks, could be explained by the effect of this variable on controlled retrieval rather than on automatic retrieval.

Because of this difficulty in interpreting IM task performance, Jacoby (1991) developed the process dissociation approach (PDA), which he argued yields separate estimates of the contributions of automatic and controlled retrieval to memory performance. The PDA stem completion procedure consists of two tasks. For the inclusion task, subjects are instructed to use the stem to recall a studied word. If they cannot recall a studied word, they are to respond with the first word that comes to mind. For the exclusion task, subjects are instructed to use the stem to recall a studied word, but then complete the stem with a word that has not been studied (Jacoby, 1998; Toth, Reingold, & Jacoby, 1994).

The inclusion and exclusion tasks are designed to evoke "direct retrieval" of studied items. The direct retrieval model makes two key assumptions. First, under the independence assumption, automatic and controlled retrieval are assumed to provide independent bases of responding on inclusion and exclusion tasks (Jacoby, 1998; Jacoby, Toth, & Yonelinas, 1993). Second, under the awareness assumption (Horton, Wilson, & Evans, 2001), it is assumed that

only controlled retrieval produces conscious awareness that an item has been studied and thus *only* controlled retrieval allows for intentional responding (Jacoby, 1998; Reingold & Toth, 1996). Thus, if controlled retrieval produces a studied item, the subject is capable of "including" or "excluding" that item as a response. By contrast, automatic retrieval of a studied item does not produce awareness that the item has been studied and thus does not allow for intentional responding. In sum, for the inclusion condition, responding with studied items is expected when studied items are retrieved via either automatic or controlled retrieval, whereas for the exclusion condition, responding with studied items is expected only when studied items are retrieved via automatic processes but not via controlled processes.

These assumptions are quantified in the following equations. The probability of retrieval of a studied item on the inclusion (I) and exclusion (E) tasks, respectively, is defined as

$$I = C + A(1 - C),$$
$$E = A(1 - C),$$

where C and A refer to the probability of retrieval via controlled and automatic processes, respectively. Then, an estimate of controlled retrieval may be derived as

$$C = I - E. \qquad (1)$$

An estimate of automatic retrieval conditionalized on the failure of controlled retrieval may be derived as

$$A_{|}\overline{C} = \frac{E}{(1 - C)}. \qquad (2)$$

The independence assumption necessitates (Buchner, Erdfelder, & Vaterrodt-Plünnecke, 1995) that

$$A|C = A = A|\overline{C} \qquad (3)$$

and allows the PDA to estimate unconditional automatic retrieval as

$$A = A_{|}\overline{C} = \frac{E}{(1 - C)}. \qquad (4)$$

Estimates of automatic retrieval from the PDA and from IM tasks have been compared. For example, using a typical IM task, Toth et al. (1994) found a reliable depth-of-processing effect on priming, suggesting that the variable affected automatic retrieval. By contrast, using the PDA, they found a depth-of-processing effect on the estimate of controlled retrieval but not on the automatic estimate. Toth et al. concluded that the depth effect found on the IM task was attributable to contamination by controlled retrieval.

The conclusion offered by Toth et al. (1994) is appropriate only if the PDA estimates of automatic and controlled retrieval are accurate. However, some researchers have questioned the key assumptions underlying the PDA (e.g., Curran & Hintzman, 1995, 1997; Joordens & Merikle, 1993; Mecklenbräuker, Wippich, & Mohrhusen, 1996; Richardson-Klavehn & Gardiner, 1996; Richardson-Klavehn, Gardiner, & Java, 1996; but see Jacoby, 1998; Jacoby, Begg, & Toth, 1997; Jacoby & Shrout, 1997). Concerns have been raised regarding theoretical, methodological, and empirical aspects of PDA (e.g., Bodner, Masson, & Caldwell, 2000; Buchner et al., 1995; de Houwer, 1997; Dodson & Johnson, 1996; Komatsu, Graf, & Uttl, 1995; Richardson-Klavehn & Gardiner, 1998). At least some of these issues have yet to be resolved conclusively.

If automatic and controlled retrieval are not independent but rather are positively correlated (Curran & Hintzman, 1995, 1997; Joordens & Merikle, 1993), then the relationship between automatic and controlled retrieval is described as

$$A|C > A > A|\overline{C}.$$

If a positive correlation exists such that $A > A|\overline{C}$, then Equation 3 is invalid, and estimating A using the independence assumption $A = A|\overline{C}$, as with Equation 4, is invalid. With a positive correlation, $A|\overline{C}$ will underestimate A, and this underestimation will increase as both $\overline{C}$ and the strength of the correlation increase. To illustrate how the underestimation by the PDA will increase as C increases, consider Equation 5, which defines unconditional automatic retrieval in terms of the conditional probabilities $(A|C, A|\overline{C})$ and controlled retrieval (Buchner et al., 1995):

$$A = (A|C) \times C + (A|\overline{C}) \times \overline{C}. \qquad (5)$$

As C approaches zero, A approaches $A|\overline{C}$, and thus the PDA estimate $(A|\overline{C})$ should be accurate, even when there is a positive correlation. As C approaches unity, A approaches $A|C$, and thus the PDA's estimate $(A|\overline{C})$ should underestimate A, and this underestimate will approach the difference between $A|\overline{C}$ and $A|C$. Critical to the present experiment, if there is a positive correlation of automatic and controlled retrieval, then the PDA estimate of automatic retrieval is expected to produce a significant underestimation of automatic retrieval when controlled retrieval is high, but minimal underestimation when controlled retrieval is low.

Recent work by Stolz and Merikle (2000) may be understood in light of this analysis. They used PDA to examine retention interval effects on automatic and controlled retrieval. As expected, they found that controlled retrieval decreased over retention intervals (Experiment 1: 2 min, $C = .54$; 2 days, $C = .15$; 2 weeks, $C = .04$; 2 months, $C = .01$). Unexpectedly, automatic influences increased from 2 min to 2 days, and then remained stable over the longest intervals (Experiment 1: 2 minutes, $A|\overline{C} = .13$; 2 days, 2 weeks, and 2 months, $A|\overline{C} = .24$). The increase in automatic influences over the shorter intervals is in sharp contrast to findings from IM studies in which automatic influences have either remained relatively stable (e.g., Jacoby & Dallas, 1981; Tulving, Schacter, & Stark, 1982) or decreased over time (e.g., Roediger, Weldon, Stadler, & Riegler, 1992; Sloman, Hayman, Ohta, Law, & Tulving, 1988). It should be noted that Stolz and Merikle examined automatic retrieval conditionalized on the ab-

sence of controlled retrieval, $A\,|\,\overline{C}$, rather than automatic retrieval, A. However, because we are interested in changes in A over time and the validity of PDA estimates of A, we will use the Stolz and Merikle data to derive PDA estimates of A.

There are two possible explanations for these apparently contradictory findings from IM tasks and PDA. First, it could be argued that PDA provides an accurate estimate of automatic retrieval, whereas priming on IM tasks is contaminated with controlled retrieval. Because the opportunity for, and the success of, controlled retrieval should be highest at short intervals, overestimation of automatic retrieval by an IM task should be greater for short than for long retention intervals. This differential overestimation for short and long retention intervals may result in estimates of automatic retrieval that decrease over time when, in fact, automatic retrieval remains stable, or counterintuitively, increases.

The second explanation for the contradictory findings is that IM tasks provide accurate estimates of automatic retrieval, but questions the validity of the PDA automatic retrieval estimates. If automatic and controlled retrieval are correlated rather than independent processes, then, as noted previously, $A\,|\,\overline{C}$ will underestimate A. Importantly, as C increases, the magnitude of the underestimation of A will increase (see also Russo, Cullis, & Parkin, 1998). Thus, the underestimation of automatic retrieval by PDA should be greater for short retention intervals. This differential underestimation for short and long retention intervals may result in PDA estimates of automatic retrieval that increase with time when in fact automatic retrieval remains stable or decreases. Thus, there may not be an increase in automatic retrieval over time, but rather the increase in A derived from the Stolz and Merikle (2000) PDA data may be the result of a violation of the independence assumption.

We attempted to determine which of these two explanations is correct. Note that these explanations are not mutually exclusive. Both explanations could be contributing to the discrepant results, and this possibility should be detectable with the procedure used in the present experiment. The procedure was an adaptation of that described by Horton et al. (2001; see also Richardson-Klavehn, Clarke, & Gardiner, 1999; Richardson-Klavehn & Gardiner, 1995, 1996, 1998). After either a zero or a 7-day retention interval, subjects in a speeded IM group completed two practice and two test phases of a speeded IM stem completion task. None of the stems presented during the practice phases could be completed with study words, whereas half of the stems presented during the test phases could be completed with study words. Stem completion response times (RTs) were measured as an indicator of retrieval strategy. To determine whether subjects in the speeded IM group adopted a controlled retrieval strategy on the test phases, RTs were compared, first, to a baseline group, in which controlled retrieval was improbable because none of the stems could be completed with a study word and, second, to a speeded EM group, in which controlled retrieval was instructed on the test phases.

Because automatic retrieval is generally executed faster than controlled retrieval (e.g., de Houwer, 1997; Richardson-Klavehn & Gardiner, 1995, 1998; Toth, 1996; Yonelinas & Jacoby, 1994), we expected to replicate our earlier finding (Horton et al., 2001) that the speeded EM group prompted longer RTs than did the baseline group. Longer RTs for the speeded EM group could reflect use of a more time-consuming controlled retrieval process and/or a more time-consuming postretrieval process in which subjects check whether the retrieved word was studied. Regardless, these longer RTs index use of a controlled retrieval strategy to increase completion of stems with studied words—a strategy different from that used by the baseline group. Therefore, contamination of the speeded IM group data with either of these controlled retrieval strategies should be indicated by longer RTs relative to those of the baseline group. However, we expected that RTs of our speeded IM group would not differ from those of the baseline group (Horton et al., 2001), providing no evidence for contamination.

With evidence that our speeded IM group was engaged in automatic retrieval, priming on the speeded IM task was then compared with priming from a standard IM group and with automatic estimates from a standard PDA group. The hypothesis that priming in the standard IM group was contaminated with controlled retrieval would be supported if the standard IM group produced greater priming than the speeded IM group. The overestimation was expected to be greater in the zero-delay condition, as described above. The hypothesis that the independence assumption of PDA is incorrect, and that automatic and controlled retrieval are positively correlated, would be supported if automatic estimates from the standard PDA group were less than priming in the speeded IM group. The underestimation was expected to be greater in the zero-delay condition, as noted above.

METHOD

Subjects

A total of 160 undergraduate students participated in return for course credit.

Design

Sixteen subjects were randomly assigned to each of the conditions formed by the factorial combination of the between-subjects factors of test type (standard IM, speeded IM, standard PDA, speeded EM, baseline) and delay (zero, 7-day). For the standard PDA condition, inclusion and exclusion tasks were administered within subjects.

Materials

A total of 156 words were selected in such a way that none were proper nouns and all had straightforward spellings and unique three-letter stems within the list. The 156 words were divided into four critical lists (24 words each), a filler list (24 words), a baseline study list (24 words), and 12 buffer words (4 study, 4 nonstudy, 4 baseline).

On the basis of a pilot study, each critical list had a mean stem completion baseline rate of .23. Each critical list served equally often as the study list, nonstudy list, and the two practice lists. For the standard PDA condition, the study and nonstudy lists were divided in half, with each half serving equally often as the inclusion or exclusion lists.

Study phase. Stems for the study and filler list words were first randomized and then presented in alternating order. Two study buffers

were presented at the beginning and two at the end of the study phase. For subjects in the baseline condition, the baseline study list replaced the study list and the baseline buffers replaced the study buffers.

Practice phase. Stems for words in Practice Lists 1 and 2 were randomized and presented during Practice Phases 1 and 2, respectively.

Test phase. Half of the stems for the study list and nonstudy list words were presented during Test Phase 1 and half during Test Phase 2. Presentation order was block randomized within each list, with each block containing two study and two nonstudy list stems. For the standard PDA condition, eight buffer stems (four study and four nonstudy) were presented at the beginning of Test Phase 1. For all other conditions, four buffer stems (two study and two nonstudy) were presented at the beginning of each test phase. For the baseline condition, the test phase was as described except that the "study list stems" and "study buffers" were not in fact studied.

Procedure

Subjects in the zero-delay condition completed all tasks in a single session. Subjects in the 7-day delay condition completed an unrelated nonverbal filler task, the study phase, and another filler task in the first session and then returned 7 days later to complete the practice and test phases.

Study phase. Subjects performed a semantically cued stem completion study task. Each trial consisted of presentation of the stem for a study list word along with a semantic cue (e.g., a large animal with a trunk—ELE). This study task was chosen so that at the zero delay, both automatic and controlled retrieval would be substantial, allowing us to observe controlled retrieval contamination of the standard IM task and underestimation of automatic retrieval by PDA, should either occur.[1] The stem was presented three lines below the semantic cue in the center of a computer screen. Subjects were instructed to complete the stem with a word that made sense given the semantic cue. If the subject could not provide the target word, the experimenter supplied the correct answer. Subjects provided the target for 99.1% of the stems.

Practice phase. Each trial consisted of presentation of a stem in the center of the computer screen. Subjects in the standard IM and standard PDA conditions were instructed to complete each stem with the first word that came to mind, although there was no particular emphasis on speed. Subjects in all other conditions were instructed to complete each stem with the first word that came to mind *and* to respond as quickly as possible.

Test phase. In all but the standard PDA condition, each test trial consisted of presentation of a stem, identical to trials in the practice phase. Instructions to the standard IM groups were identical for the practice and test phases—"Complete each stem with the first word that comes to mind." Instructions to the speeded IM and baseline groups were also identical for the practice and test phases—"Complete each stem with the first word that comes to mind *and* respond as quickly as possible." Subjects in the standard IM and speeded IM groups were not informed that some stems could be completed with studied words.

Immediately prior to the test phase, subjects in the speeded EM condition were informed that some stems on the test phase could be completed with study phase words. They were instructed to try to complete each stem with a word that they had generated at study, but still to respond as quickly as possible. If they could not recall a studied word, they were to respond with the first word that came to mind.

For the speeded IM, speeded EM, and baseline conditions, a voice-activated relay recorded RTs for each trial of the practice and test phases—the time from onset of the stem on the screen until initiation of a verbal response. Following each practice and test phase, subjects were shown their mean RT for that phase and any previous phases, and encouraged to go faster on the next phase.

In the standard PDA condition, each trial consisted of presentation of a stem in the center of the computer screen, with the word *old* or *new* presented three lines above. Subjects were instructed that when the word *old* appeared (inclusion), they were to try to complete the stem with a study phase word. If they could not remember a study word, they were to complete the stem with the first word that came to mind. When the word *new* appeared (exclusion), they were to respond with the first word that came to mind that was not a study phase word. The eight buffer stems were used as examples to ensure that subjects understood the instructions.

RESULTS

An alpha level of .05 was adopted for all statistical tests.

Response Times

For each subject in the speeded IM, baseline, and speeded EM conditions, median RTs were calculated for all trials of the practice phases and for the nonstudy word trials of the test phases. The RTs appear in Table 1.

Table 1
Median Stem Completion Response Times and Standard Deviations
(in Milliseconds) as a Function of Test Type, Delay, and Phase

Delay	Phase	Baseline M	Baseline SD	Speeded EM M	Speeded EM SD	Speeded IM M	Speeded IM SD
Zero delay	Practice 1	865	123	850	138	829	114
	Practice 2	877	156	810	156	820	133
	Test 1, nonstudied	838	132	1,387	514	844	180
	Test 2, nonstudied	850	172	1,220	356	817	245
	Test 1, studied	–		930	209	780	282
	Test 2, studied	–		901	178	724	130
7-day delay	Practice 1	876	145	811	102	817	153
	Practice 2	868	179	802	119	816	140
	Test 1, nonstudied	829	174	1,280	343	804	175
	Test 2, nonstudied	854	172	1,061	215	823	189
	Test 1, studied	–		995	287	788	183
	Test 2, studied	–		870	285	724	125

Note—The practice phases contained only nonstudied items. Response times for both nonstudied and studied items are provided for the test phases of the speeded explicit (speeded EM) and speeded implicit (speeded IM) conditions.

To determine whether subjects in the speeded IM condition used a controlled retrieval strategy, an analysis of variance (ANOVA) was conducted on RTs as a function of phase (Practice 1, Practice 2, Test 1 nonstudy items, Test 2 nonstudy items), test type (speeded IM, baseline, speeded EM), and delay (zero, 7-day). We included only the nonstudy word trials because study word trials were expected to show RT priming. This priming might offset, and thus make difficult to detect, an increase in RTs due to controlled retrieval contamination. The three-way interaction was not significant [$F(6,270) < 1$, $MS_e = 23,342$]. Also, the effect of delay, as well as the interactions between phase and delay and between test type and delay, was not significant (all Fs < 1), indicating that delay had no influence on RTs.

Critically, there was a significant interaction between phase and test type, and significant main effects of both phase and test type [$F(6,270) = 29.26$, $F(3,270) = 23.78$, $F(2,90) = 15.63$, respectively]. Simple main effects of test type for Practice Phases 1 and 2 were not significant [$F(2,93) = 1.27$, $MS_e = 16,595$, and $F(2,93) = 1.86$, $MS_e = 21,317$], respectively, indicating that, for the practice phases, RTs for the three test types did not differ. However, the simple main effects of test type for Test Phases 1 and 2 were significant [$F(2,93) = 33.83$, $MS_e = 80,501$, and $F(2,93) = 18.12$, $MS_e = 55,078$, respectively]. A least significant difference test for each test phase (Phase 1 LSD = 241 msec; Phase 2 LSD = 199 msec) indicated that RTs for the speeded IM (Phase 1, 824 msec; Phase 2, 820 msec) and baseline (Phase 1, 833 msec; Phase 2, 852 msec) conditions did not differ significantly, but RTs for both these conditions differed from those for the speeded EM condition (Phase 1, 1,334 msec; Phase 2, 1,140 msec).

The RT data indicate that subjects in the speeded IM, baseline, and speeded EM conditions utilized the same retrieval strategy for the practice phases. Because all three groups were given the same IM instructions for the practice phases, and because none of the stems could be completed with studied words, it is assumed that subjects responded using automatic retrieval for the practice phases, as instructed. Subjects in the baseline condition showed no change in RTs from the practice to the test phases, suggesting that these subjects continued to respond using automatic retrieval, as instructed. However, subjects in the speeded EM condition were instructed, and thus expected, to switch to a controlled retrieval strategy for the test phases. This switch to a controlled retrieval strategy resulted in an increase in RTs for the test phases, thereby replicating previous data (Horton et al., 2001; Richardson-Klavehn & Gardiner, 1995, 1998) in showing that controlled retrieval takes longer to execute than automatic retrieval. The critical result was that RTs for the speeded IM condition did not show an increase in RTs for the test phases, as occurred for the speeded EM condition. Indeed, RTs for the speeded IM condition did not change from the practice to the test phases mirroring those in the baseline

condition for all phases. This indicates that subjects in the speeded IM condition continued to respond using automatic retrieval on the test phases. These results support the conclusion that responding on the speeded IM task was not contaminated by controlled retrieval.

An additional analysis comparing RTs for nonstudied items with RTs for successfully retrieved studied items was conducted on the speeded IM group to determine whether there was significant RT priming. An ANOVA was conducted on RTs as a function of test phase, study history (studied correct, nonstudied), and delay. The three-way interaction was not significant [$F(1,30) < 1$, $MS_e = 12,216$]. Also, the effect of delay, along with the interactions between test phase and delay and between study history and delay, was not significant (all Fs < 1), indicating that delay had no influence. Neither the interaction between test phase and study history nor the effect of test phase was significant [$F(1,30) = 2.05$, and $F(1,30) = 1.32$, $MS_e = 24,594$, respectively], indicating that test phase had no influence. However, there was a significant effect of study history [$F(1,30) = 8.78$, $MS_e = 16,649$], indicating that stem completion RTs were faster for studied than for nonstudied items.

Estimates of Automatic Retrieval

An initial analysis of nonstudied target scores in the standard PDA condition (Table 2) as a function of instructions (inclusion, exclusion) and delay revealed no significant effects (all Fs < 1.04). Thus, an average nonstudied PDA score for each subject was used in the second analysis. The second analysis of nonstudied scores as a function of test type (speeded IM, standard IM, standard PDA) and delay revealed no significant effects of test type [$F(2,90) = 1.12$, $MS_e = 0.01$], delay ($F < 1$), or their interaction ($F < 1$). Because nonstudied scores did not differ across test type or delay conditions, estimates of automatic retrieval were not corrected for nonstudied scores (Toth et al., 1994).

Because the RT analysis indicated that responding in the speeded IM condition provided an uncontaminated estimate of automatic retrieval, the speeded IM estimates of automatic retrieval were compared with those from the standard IM and standard PDA conditions. An ANOVA was conducted on automatic retrieval estimates as a function of test type (speeded IM, standard IM, standard PDA) and delay. For subjects in the speeded IM and standard IM conditions, automatic retrieval was estimated by the proportion of test phase study stems completed with studied words. For subjects in the standard PDA condition, an estimate of automatic retrieval was calculated using Equation 4. The estimates are shown in Table 2.

The analysis revealed a significant interaction of test type and delay, and a significant main effect of test type [$F(2,90) = 11.63$, $MS_e = 0.013$, and $F(2,90) = 19.03$, respectively]. The main effect of delay approached significance [$F(1,90) = 3.17$, $p = .08$]. The simple main effect of test type at the zero delay was significant [$F(2,45) = 23.41$, $MS_e = 0.017$], whereas the simple main effect of

Table 2
Automatic Estimates and Nonstudied Completion Rates (and Standard Deviations)
for the Speeded Implicit (Speeded IM), Standard Implicit (Standard IM),
and Standard Process Dissociation Approach (PDA) Groups as a Function of Delay

Delay	Memory Estimate	Speeded IM		Standard IM		Standard PDA	
		M	SD	M	SD	M	SD
Zero delay	Inclusion					.64	.13
	Exclusion					.09	.07
	Controlled					.55	.13
	Automatic	.48	.11	.49	.11	.21	.06
	Nonstudied	.24	.07	.28	.12	.27	.06
7-day delay	Inclusion					.46	.15
	Exclusion					.27	.12
	Controlled					.19	.21
	Automatic	.36	.08	.36	.09	.33	.11
	Nonstudied	.23	.09	.24	.07	.27	.08

Note—Inclusion scores, exclusion scores, and controlled estimates are also shown for the standard PDA group.

test type at the 7-day delay was not significant $[F(2,45) < 1, MS_e = 0.009]$. A LSD test (LSD = .11) indicated that at the zero delay, the estimate of automatic retrieval for the speeded IM (.48) and the standard IM (.49) conditions did not differ significantly, but the automatic estimates for both these conditions differed from that of the standard PDA condition (.21).

These results indicate that, for the 7-day delay, estimates of automatic retrieval from the standard IM and standard PDA conditions are both valid. This result is not surprising because conscious awareness is low for the 7-day delay (.17). When conscious awareness is low, contamination of the standard IM task with controlled retrieval should be minimal. Furthermore, when conscious awareness is low, even if automatic and controlled retrieval are positively correlated, underestimation in the standard PDA condition should be minimal. By contrast, for the zero delay, the results indicate that the automatic retrieval estimate from the standard IM task is uncontaminated and valid, but that the estimate from the standard PDA condition underestimates automatic retrieval.

Finally, simple main effects of delay for the standard IM and standard PDA conditions were conducted to determine whether our results were consistent with those of previous studies. The simple main effect of delay for the standard IM condition was reliable $[F(1,30) = 13.48, MS_e = 0.01]$. Target completion decreased from the zero delay (.49) to the 7-day delay (.36), consistent with results from previous studies (Roediger & Blaxton, 1987; Roediger et al., 1992; Sloman et al., 1988). The simple main effect of delay for the standard PDA condition was also reliable $[F(1,30) = 5.74, MS_e = .019]$. Automatic retrieval estimates increased from the zero (.21) to the 7-day delay (.33), consistent with the findings of Stolz and Merikle (2000).[2]

Estimates of Controlled Retrieval

Estimates of controlled retrieval, C, were calculated for each subject in the standard PDA condition using Equa-

tion 1 (Table 2). As expected, controlled retrieval was greater in the zero than in the 7-day delay $[t(30) = 5.88, SEM = 0.062]$.

DISCUSSION

According to our data from the standard PDA group, automatic retrieval increased from the short to the long delay, consistent with Stolz and Merikle (2000). By contrast, the data from the standard IM group replicated previous IM data in showing a decrease in automatic retrieval over time (Roediger et al., 1992; Sloman et al., 1988). To provide an explanation for these contrasting findings, we adopted the speeded IM procedure, which utilizes RT as an indicator of retrieval strategy (Horton et al., 2001; Richardson-Klavehn & Gardiner, 1995, 1998). We concluded that the speeded IM group performed on the basis of strictly automatic retrieval.

It has been suggested to us that even with our RT evidence, there may still be some contamination in the speeded IM task. For example, one could speculate that the speeded IM group used a controlled retrieval strategy different from that used by the speeded EM group, thus producing different RT patterns. However, to account for our results, one would have to further hypothesize that the controlled retrieval strategy used by the speeded IM group produces a similar RT pattern as the automatic retrieval strategy used by the baseline group. We find our interpretation that similar RT patterns are produced by the same retrieval strategy—automatic retrieval—more parsimonious than the interpretation that different retrieval strategies produce the same RT pattern.

Furthermore, our speeded IM task was expected to minimize controlled retrieval contamination to a much greater degree than previous IM studies. First, subjects are strongly encouraged to respond quickly. Second, they receive two practice phases in which none of the stems can be completed with studied words, thus encouraging development of an automatic retrieval strategy. With contamination

minimized, it would be difficult to assert that the PDA's automatic retrieval estimate for the zero delay of .21 is accurate, and that because of controlled retrieval contamination, the speeded IM group estimate of .48 overestimates true automatic retrieval by more than double. Within the conditions of our experiment, it is doubtful that contamination could produce such a large overestimation of automatic retrieval.

Having concluded that the speeded IM estimates of automatic retrieval were not contaminated with controlled retrieval, we compared them to the automatic estimates from the standard IM and standard PDA groups. There was no evidence that the standard IM task was contaminated by controlled retrieval, because the standard IM group provided a valid index of automatic retrieval at both delays. In contrast, the standard PDA group underestimated automatic retrieval when conscious awareness was high (zero delay), but not when conscious awareness was low (7-day delay). This pattern of underestimation by the standard PDA group was predicted by and supports the hypothesis that automatic and controlled retrieval are positively correlated (Curran & Hintzman, 1995, 1997; Russo et al., 1998). Furthermore, these results indicate that the increase in automatic retrieval over time derived from Stolz and Merikle's (2000) PDA data and found in our standard PDA data are artifactual. These artifactual increases in automatic retrieval over time appear to be a consequence of the PDA's incorrect assumption that automatic and controlled retrieval independently contribute to PDA task performance, when their contributions are in fact positively correlated.

There are several reasons why a positive correlation is not surprising and should be expected. First, as explained below, it is probable that subjects used a generate/recognize strategy on some inclusion and exclusion tasks. The relational assumption underlying generate/recognize theories is redundancy (Joordens & Merikle, 1993), in which $A|C = 1$. Except for extreme cases, $A|\overline{C}$ will be less than 1, making $A|\overline{C} < A|C$, thus making generate/recognize theories consistent with a positive correlation. Second, Curran and Hintzman (1997) have suggested that automatic and controlled retrieval may share some processing stages. For example, it is plausible that automatic and controlled retrieval share a lexical access stage. Any processing at study that facilitates this common lexical access would be expected to facilitate both automatic and controlled retrieval, causing a positive correlation. Third, even if automatic and controlled retrieval do not use the same lexical access stage, they may be expected to each have their own lexical access stage. Again, any study processing that facilitates lexical access would be expected to facilitate both automatic and controlled retrieval, causing a positive correlation.

This third point illustrates the extreme situation required for automatic and controlled retrieval to make independent contributions to inclusion and exclusion task performance. Even if automatic and controlled retrieval were separate retrieval processes with no shared components, any processing done at study that similarly affected automatic and controlled retrieval would be expected to yield a positive correlation in their contributions to inclusion and exclusion task performance. So, although there may be some processes that are independent and could be accurately indexed with the PDA, a positive correlation between automatic and controlled retrieval processes seems likely to be the typical situation.

Whereas we have suggested that the PDA's underestimation is attributable to a faulty independence assumption, Horton et al. (2001) have identified a second explanation for this underestimation (see also Richardson-Klavehn & Gardiner, 1996, 1998). To produce a valid estimate of automatic retrieval, *both* the independence and awareness assumptions must be valid. It is possible that automatic and controlled retrieval are independent, but that the awareness assumption is flawed in such a way that automatic retrieval can produce conscious awareness that an item has been studied, a phenomenon termed "involuntary conscious awareness" (Richardson-Klavehn & Gardiner, 1996). To the extent that automatically retrieved studied items are not used as responses on exclusion trials because of involuntary conscious awareness, the PDA estimate of automatic retrieval will underestimate true automatic retrieval. If it is reasonably assumed that involuntary conscious awareness is greater at a short than at a long retention interval, then our finding that the PDA underestimates automatic retrieval at a short but not at a long retention interval is readily explained.

Toth and Reingold (1996) have asserted that the PDA and the IM tasks, should produce the same result if the "two paradigms are measuring the same construct and the implicit test is uncontaminated by conscious uses of memory" (p. 56). Results from the present study suggest that the IM task is uncontaminated by controlled (conscious) uses of memory. Accordingly, the PDA and the IM tasks must be measuring different constructs. If the independence assumption were valid, then Equation 4 would be valid, and the two procedures would index the same construct, automatic retrieval. However, our results indicate that automatic and controlled retrieval are positively correlated, and thus the discrepancy in IM and PDA results seems to occur because the PDA is not truly estimating automatic retrieval: It is measuring a different construct—namely, automatic retrieval conditionalized on controlled retrieval (Equation 4).

The PDA requires a direct retrieval strategy in order to satisfy the requirements of the independence assumption (Jacoby, 1998; Jacoby et al., 1997). However, a generate/recognize strategy is also possible: Subjects could generate (automatic retrieval) a word completion for each stem and then use controlled processes to determine whether the generated word has been studied. Curran and Hintzman (1995) suggested that, similar to how researchers using IM tasks must concern themselves with the potential for contamination with controlled retrieval, researchers using the PDA must concern themselves with the potential for contamination with generate/recognize retrieval

strategies. With our speeded IM procedure, we believe we are able to minimize contamination of the IM task with controlled processes and, critically, provide evidence that performance on the speeded IM task truly reflects automatic retrieval only. In contrast, with the PDA, it may be difficult to eliminate contamination with generate/recognize strategies, and it may be difficult to detect this contamination when it does occur (Bodner et al., 2000; Richardson-Klavehn & Gardiner, 1995).

We do not suggest that the independence assumption is violated and automatic retrieval is underestimated for all PDA tasks, or that no IM tasks are contaminated by controlled retrieval. However, it is plausible that the independence assumption is violated and automatic retrieval is underestimated for other PDA experiments that contain inclusion and exclusion tasks cued with stems. The faster RTs for the baseline group relative to the speeded EM group indicate that generating words for stems (automatic retrieval) is easier than trying to complete stems with studied words (controlled retrieval). The ease and speed of automatic retrieval may encourage PDA subjects to rely on generate/recognize strategies, causing a positive correlation between automatic and controlled retrieval. Furthermore, the ease and speed of automatic retrieval might encourage IM subjects to follow IM instructions and discourage contamination by a slower controlled retrieval strategy.

Finally, we have been concerned that, in some situations, implementation of an independence-based PDA produces misleading interpretations of the effects of variables on automatic retrieval. The typical PDA conclusion is that a variable affects controlled retrieval but has no impact on automatic retrieval. On the basis of our studies, we conclude that some variables (e.g., retention interval, depth of processing, full vs. divided attention) have similar effects on automatic and controlled retrieval, but that the effects on automatic retrieval are less than the effects on controlled retrieval (Horton et al., 2001).

REFERENCES

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

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

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

Challis, B. H., & Brodbeck, D. R. (1992). Level of processing affects priming in word fragment completion. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **18**, 595-607.

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

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

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

Dodson, C. S., & Johnson, M. K. (1996). Some problems with the process-dissociation approach to memory. *Journal of Experimental Psychology: General*, **125**, 181-194.

Gardiner, J. M., Dawson, A. J., & Sutton, E. A. (1989). Specificity and generality of enhanced priming effects for self-retrieved studied words. *American Journal of Psychology*, **102**, 295-305.

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

Horton, K. D., & Nash, B. D. (1999). Perceptual transfer in stem-completion and fragment-completion tests. *Canadian Journal of Experimental Psychology*, **53**, 203-219.

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

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

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

Jacoby, L. L., Begg, I. M., & Toth, J. P. (1997). In defense of functional independence: Violations of assumptions underlying the process-dissociation procedure? *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **23**, 484-495.

Jacoby, L. L., & Dallas, M. (1981). On the relationship between autobiographical memory and perceptual learning. *Journal of Experimental Psychology: General*, **110**, 306-340.

Jacoby, L. L., & Shrout, P. E. (1997). Toward a psychometric analysis of violations of the independence assumption in process dissociation. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **23**, 505-510.

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.

Joordens, S., & Merikle, P. M. (1993). Independence or redundancy? Two models of conscious and unconscious influences. *Journal of Experimental Psychology: General*, **122**, 462-467.

Komatsu, S., Graf, P., & Uttl, B. (1995). Process dissociation procedure: Core assumptions fail, sometimes. *European Journal of Cognitive Psychology*, **7**, 19-40.

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.

Reingold, E. M., & Merikle, P. M. (1988). Using direct and indirect measures to study perception without awareness. *Perception & Psychophysics*, **44**, 563-575.

Reingold, E. M., & Merikle, P. M. (1990). On the interrelatedness of theory and measurement in the study of unconscious processes. *Mind & Language*, **5**, 9-28.

Reingold, E. M., & Toth, J. P. (1996). Process dissociations versus task dissociations: A controversy in progress. In G. D. M. 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 & 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, & Cognition*, **24**, 593-609.

Richardson-Klavehn, A., Gardiner, J. M., & Java, R. I. (1996).

Memory: Task dissociations, process dissociations and dissociations of consciousness. In G. D. M. Underwood (Ed.), *Implicit cognition* (pp. 85-158). New York: Oxford University Press.

ROEDIGER, H. L., III, & BLAXTON, T. A. (1987). Effects of varying modality, surface features, and retention interval on priming in word-fragment completion. *Memory & Cognition, 15, 379-388.*

ROEDIGER, H. L., III, WELDON, M. S., & CHALLIS, B. H. (1989). Explaining dissociations between implicit and explicit measures of retention: A processing account. In H. L. Roediger III, & F. I. M. Craik (Eds.), *Varieties of memory and consciousness: Essays in honour of Endel Tulving* (pp. 3-41). Hillsdale, NJ: Erlbaum.

ROEDIGER, H. L., III, WELDON, M. S., STADLER, M. L., & RIEGLER, G. L. (1992). Direct comparison of two implicit memory tests: Word fragment and word stem completion. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 18, 1251-1269.*

RUSSO, R., CULLIS, A. M., & PARKIN, A. J. (1998). Consequences of violating the assumption of independence in the process dissociation procedure: A word fragment completion study. *Memory & Cognition, 26, 617-632.*

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

SLOMAN, S. A., HAYMAN, C. A. G., OHTA, N., LAW, J., & TULVING, E. (1988). Forgetting in primed fragment completion. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 14, 223-239.*

STOLZ, J. A., & MERIKLE, P. M. (2000). Conscious and unconscious influences of memory: Temporal dynamics. *Memory, 8, 333-343.*

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. (1996). Beyond perception: Conceptual contributions to unconscious influences of memory. In G. D. M. Underwood (Ed.), *Implicit cognition* (pp. 41-84). New York: Oxford University Press.

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, & Cognition, 20, 290-303.*

TULVING, E., SCHACTER, D. L., & STARK, H. A. (1982). Priming effects in word-fragment completion are independent of recognition memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 8, 336-342.*

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.*

NOTES

1. High automatic priming was expected because subjects generate the studied word from a word stem at both study and test. According to the transfer appropriate processing approach, the more similar the operations at study and test, the greater will be priming (Roediger, Weldon, & Challis, 1989). For example, Gardiner, Dawson, and Sutton (1989) reported that on an IM task in which subjects generated a word from a word fragment at test, priming was significantly greater if at study they generated words from the same (though semantically cued) word fragment as opposed to simply reading the target word at study (see also Horton & Nash, 1999).

2. There is some debate as to whether subjects who complete no stems on the exclusion task with a study word should be removed from the analyses (e.g., Curran & Hintzman, 1995; Jacoby et al., 1993). In the Results section, we report analyses that include data from all subjects, including those who scored zero on the exclusion task. However, all of the analyses in the sections on estimates of automatic retrieval and estimates of controlled retrieval were also conducted with the data from these subjects removed. This resulted in the removal of 4 PDA subjects, all in the zero-delay condition. Results and conclusions for these analyses were identical to those reported, with one exception. The significant simple main effect of delay for the standard PDA condition indicating that PDA estimates of automatic retrieval increased from the zero (.21) to the 7-day delay (.33) was not found when the 4 subjects scoring zero on the exclusion task were removed. With these subjects removed, automatic retrieval does not change from the zero (.28) to the 7-day (.33) delay. Note that regardless of the analysis used, the key conclusion is still that the PDA estimate at the zero delay (all subjects included, .21; subjects with exclusion scores of zero removed, .28) significantly underestimated our uncontaminated estimate of automatic retrieval (.48) from the speeded IM condition.