

The role of attention in automatization: Does attention operate at encoding, or retrieval, or both?

CONSUELO B. BORONAT and GORDON D. LOGAN
University of Illinois at Urbana-Champaign, Urbana, Illinois

In this research, we investigated whether attention operates in the encoding of automatized information, the retrieval of automatized information, or in both cases. Subjects searched two-word displays for members of a target category in focused-attention or divided-attention conditions that were crossed with block (training vs. transfer). To see whether subjects encoded all available items or only attended items, we compared performance for subjects in different training conditions but in the same transfer condition. Subjects encoded attended items. To see whether subjects retrieved all the items they had in memory, or only items associated with that to which they were attending at retrieval, we compared performance for subjects in the same training conditions but in different transfer conditions. Subjects retrieved attended items. Attention was found to operate at both encoding and retrieval. These findings support the instance theory of automaticity, which predicts the role of attention at encoding and retrieval.

In this paper, we investigate the question of when attention operates in the learning that leads to automatic, or skilled, performance. Automaticity is central to many everyday tasks, such as reading and driving, and is also prominent in skill acquisition (e.g., Bryan & Harter, 1899). Logan and Etherton (1994), as discussed below, showed that attention is important in the learning that leads to automatic performance, but did not indicate whether attention operates at encoding, at retrieval, or at both. The present experiment addresses this issue directly; we examined whether attention selects the material to be encoded, the material to be retrieved at test, or whether it performs both functions. The instance theory of automaticity (Logan, 1988, 1990, 1992), a memory-based theory of automatic performance, assumes that attention operates at both of these points. The present research therefore provides a direct test of instance theory's encoding and retrieval assumptions, which are two of the theory's three key assumptions. We also examine the question of how attention might operate at encoding and retrieval.

The question of when attention operates is important for the fields of attention and memory. Part of the memory literature concerns the role of attention in encoding and retrieval (see, e.g., A. Cohen, Ivry, & Keele, 1990; Craik & Lockhart, 1972; Kirsner & Dunn, 1985; Morris, Bransford, & Franks, 1977; Nissen & Bullemer, 1987; Roediger, 1990; Stadler, in press). Although there is general agreement in this work that attention is important for

encoding and retrieving information, there are few detailed processing accounts of how attended items are encoded. The attention literature has largely focused on the processes and mechanisms that select information, and has avoided issues of memory by using novel stimuli. We used the instance theory of automaticity (Logan, 1988, 1990, 1992) to examine the relationship of attention and memory because it provides a processing model and set of predictions about this relationship.

Instance Theory

The instance theory is a memory-based model of automaticity. Its three main assumptions are that (1) there is *obligatory encoding* of material to which one attends, (2) there is *obligatory retrieval* of instances that are associated in memory with material to which one is attending, and (3) memory representations are *instance representations*. We test the first two assumptions in the present experiment.

The instance theory asserts that automatic performance is based on the retrieval of past solutions from memory (Logan, 1988, 1990, 1992). According to this theory, the development of automaticity is a result of a person's acquisition of exemplars that are stored in memory. The more experience people have with a task, the more exemplars (or instances) from the task they will have in memory. Skilled performers of a task have more instances in memory than do novices, and this difference is what produces the performance differences between them.

One of the hallmarks of automaticity is a power function speed-up, which characteristically is an initial, sharp drop-off in the time it takes to do a task (i.e., a speed-up in processing), and then a gradual leveling off of performance, until performance reaches an asymptote (Anderson, 1982; J. D. Cohen, Dunbar, & McClelland, 1990; Lo-

This research was supported by Grant SBR 94-10406 from the National Science Foundation, principal investigator G. D. L. Correspondence should be addressed to C. B. Boronat or G. D. Logan, Psychology Department, University of Illinois at Urbana-Champaign, 603 E. Daniel St., Champaign, IL 61820 (e-mail: cboronat@s.psych.uiuc.edu or glogan@s.psych.uiuc.edu).

gan, 1988; Newell & Rosenbloom, 1981). Automaticity also shows a parallel decrease in performance variation over time. The equation for the power function is

$$RT = a + bN^{-c},$$

where RT is the time subjects take to do the task, N is the number of practice trials, and a , b , and c are constants. A represents the asymptote; this is the minimum amount of time in which subjects make a response and so is an indicator of what they can eventually learn with practice. B represents the amount to be learned: It is the difference between initial and asymptotic performance. C is the rate of learning.

The instance theory accounts for power function fits to reaction time (RT) data (both means and standard deviations) using a horse-race model of processing. The horse-race metaphor refers to the fact that performance is based on the first instance retrieved from memory: The first instance is the winner. The more practice people have with a task, the more instances they will encode in memory. The more instances they have, the greater the probability that one will be retrieved quickly. The increase in the probability of a quick retrieval with practice is reflected in the power function speed-up.

According to the obligatory encoding assumption, one necessarily encodes material to which one attends. The sense of "encoding" used here is the sense found in the memory literature: Encoding is a process that results in storage in long-term memory. "Encoding" is not used here as it is in the attention literature, where it is used to describe a process resulting in activation of a long-term memory representation that permits identification. This second sense of encoding is a more transitory sense than the sense intended in the obligatory encoding assumption.

The strong version of the obligatory encoding assumption is that only material to which one attends is encoded in memory; the weak version is that some unattended material might also be encoded. There is great difficulty in fully controlling subjects' attention (a requirement for testing the strong version of this assumption; see, e.g., Hollender, 1986, and commentary), so research has focused on the weak version. Logan and Etherton (1994) found support for a weak version of this assumption.

According to the obligatory retrieval assumption, one necessarily retrieves from memory material that is associated with the material to which one is attending. The contents of attention act as a retrieval cue that pulls material out of memory. The strong version of this claim is that only attended items act as retrieval cues; the weak version is that unattended items may also act as retrieval cues. Again, because of the difficulty in fully delimiting what subjects pay attention to, extant research has tested the weak assumption (Hollender, 1986).

The instance representation assumption holds that representations in memory are representations of the co-occurrences in a processing episode. The elements of a co-occurrence may include a subject's goals, relevant stimuli, interpretation of the stimuli, and response (Logan, 1988, 1990; Logan & Etherton, 1994). Each attended en-

counter with a stimulus is a processing episode and results in the storage in memory of an instance (Logan, 1988).

Testing the Obligatory Encoding and Retrieval Hypotheses

Logan and Etherton (1994) found support for the instance theory's obligatory encoding and instance representation assumptions. They trained subjects in a categorization task on 32 pairs of words that were repeated in 16 or 64 blocks. Subjects were given a category to look for and had to indicate its presence or absence on every trial. Logan and Etherton (1994) manipulated whether subjects were to pay attention to only one word in each pair (focused-attention condition; cued word in green) or to both words (divided-attention condition; no cue). The cue was successful in focusing subjects' attention: Focused subjects were initially faster than divided subjects, suggesting that the former had been paying attention to fewer stimuli than the latter. Subjects remained in the same attentional condition (i.e., focused or divided) for training and transfer. The experiment succeeded in producing automaticity: The difference between focused and divided subjects decreased over training blocks (suggesting a decrease in load effects); and subjects' RTs decreased as a power function of practice. These two measures, a decrease in load effects and a power function speed-up, were used as automatization benchmarks in the present experiment.

At transfer, Logan and Etherton (1994) re-paired the words from training and compared performance on the re-paired pairs with performance on the last block of practice. They found that divided-condition subjects were slower on the re-paired pairs than on the old pairs, suggesting that they had learned specific pairings and that they were sensitive to disruptions of the co-occurrences that they had learned. (This difference between re-paired and old pairs of words is referred to here as *cost*.) Logan and Etherton found, however, no cost for focused-condition subjects, suggesting that these subjects had not learned specific pairings during training and so were not disrupted when pairs changed at transfer.

Logan and Etherton's (1994) experiments show that attention is potentially important at encoding (i.e., training), at retrieval (i.e., transfer), or at both points. Because they did not manipulate attentional demands in both of these segments (subjects had the same attentional demands throughout the experiment), it is not possible to tell from their experiments when attention is working. In the present experiment, we examine when attention operates.

Training and transfer. Because we were examining assumptions of a theory of automaticity, subjects performed a task in which performance has shown characteristics of automaticity: a 16-block training segment followed by a 1-block transfer segment. Subjects searched two-word displays for a member of a target category (e.g., metals). Word pairs remained the same for all training trials to allow subjects many opportunities to encode particular pairs. Each word appeared in only one pair and thus uniquely identified a pair. Each pair consisted of ei-

ther a target or a nontarget category word (e.g., metals or countries) and a word from one of the two distractor categories (e.g., vegetables or furniture). At transfer, subjects saw both old (from training) and re-paired (re-pairings of old pairs) pairs of words. The key experimental indicator was cost. The greater the sensitivity of subjects to specific pairings, the greater the disruption when pairs changed, and the higher the cost. The present experiment tested how cost might differ depending on whether subjects focused or divided their attention at, separately, training and transfer.

Attentional demands were manipulated during training and transfer with focused- and divided-attention conditions. In the focused-attention condition, the position of the target category was cued by the color green. On each trial, subjects saw a green word and a white word. Subjects were told to concentrate on the green word and disregard the white word, because the white word would never be a member of the target category. The purpose of this condition was to direct attention to a single word in a pair of words, and to see if subjects encoded or retrieved only (or primarily) that word and its associated response (i.e., *yes* or *no*). Such findings would support the obligatory encoding and retrieval assumptions by showing that subjects encoded or retrieved attended material and did not (or did to a lesser extent) encode or retrieve unattended material. For example, subjects might see the pair *steel-broccoli*, with *steel* in green letters and *broccoli* in white letters (words were presented one above the other). If subjects were searching for *metals*, they would respond "yes," and would, the instance theory predicts, store an instance representation composed of the co-occurrence of the attended word (i.e., *steel*) and the response (i.e., "yes") (Table 1).

In the divided-attention condition, the position of the target item was not cued: Both words in each pair appeared in white. At least initially, therefore, subjects might attend to both words to make a category decision. Predictions from the obligatory encoding and retrieval assumptions were that since subjects were attending to both words, they would encode or retrieve both words and their associations. During training, subjects could have encoded three types of associations (Table 1): (1) the target category word *steel* and "yes"; (2) the stimulus pair *steel-broccoli* and "yes"; and (3) the distractor word *broccoli* and "yes."

Table 1
Training Example Under Focused- and Divided-Attention Conditions

Condition	Stimuli	Association
Focused	<u>Steel</u> *	Steel → "Yes"
	Broccoli	
Divided	Steel	Steel → "Yes"
	Broccoli	Broccoli → "Yes"
		[Steel Broccoli] → "Yes"

*The word *steel* appeared in green on the computer monitor in the focused-attention condition, and without underlining. All other stimuli in this example appeared in white.

There were three checks on the efficacy of the attentional manipulation. First, focused subjects were predicted to be initially faster than divided subjects in training. If focused subjects were paying attention only to the cued item, they should be faster than divided subjects, who were assumed to be paying attention to both items. This difference should disappear as load effects decreased for the divided condition. Second, focused subjects were predicted to be faster than divided subjects at transfer because they would benefit from attending to a cued item. Third, performance was examined on a surprise recall task that subjects took after completing the category search task. Subjects were asked to write down all of the words that they experienced during training and transfer. Focused subjects, because they were focusing primarily on the cued item, were predicted to recall fewer distractors than were divided subjects.

Separating encoding and retrieval effects. The two key independent variables in the present experiment were attentional manipulation in training (focused or divided) and attentional manipulation at transfer (focused or divided). This created four conditions (training condition-transfer condition): (1) focused-focused, (2) focused-divided, (3) divided-focused, and (4) divided-divided. Crossing these conditions allowed manipulation of what subjects might encode and retrieve, and thus permitted tests of the obligatory encoding and retrieval hypotheses.

Obligatory encoding prediction. According to the obligatory encoding assumption, what one pays attention to gets into memory. To test this assumption, we compared cost at transfer for a pair of conditions in which attentional demands at training varied, but in which transfer demands were the same: the focused-divided and divided-divided conditions (Table 2). Because the retrieval condition was the same in these conditions (i.e., divided attention), we could compare the effect of the encoding conditions, which were different (i.e., focused vs. divided).

Subjects in the focused-divided condition were in the focused condition during training and in the divided condition during transfer. When they saw a pair such as *steel-broccoli* in training, with *steel* cued, they focused on *steel* and primarily encoded it and its response (i.e., "yes") (Table 2). The unattended distractor, *broccoli*, would have been encoded less often, if at all. At transfer, subjects saw old and re-paired pairs. According to the encoding assumption, if *steel* were the only word encoded during training, subjects should show no cost at transfer: RT at transfer for the old pair (i.e., *steel-broccoli*) should be the same as that for a re-paired pair (e.g., *steel-carrot*), because both pairs retrieve the same association of *steel* and "yes." The distractor, *carrot*, would not have been attended to in training and would not be associated with a response.

Subjects in the divided-divided condition were in the divided attention condition for both training and transfer. During training, such subjects could have encoded three associations between a word pair and the correct response. For the pair *steel-broccoli*, these associations would be: *steel* and "yes," *broccoli* and "yes," and the pair *steel-*

Table 2
Associations at Transfer by Condition for Old and Re-Paired Pairs

Training Pair	Old Pair			Re-Paired Pair		
	Transfer Item	Retrieval Cue	Association	Transfer Item	Retrieval Cue	Association
			Focused-Focused			
<u>Steel</u> Broccoli	<u>Steel</u> Broccoli	Steel	Steel→"Yes"	<u>Steel</u> Carrot	Steel	Steel→"Yes"
			Focused-Divided			
<u>Steel</u> Broccoli	Steel Broccoli	Steel Broccoli	Steel→"Yes"	Steel Carrot	Steel Carrot	Steel→"Yes"
			Divided-Focused			
Steel Broccoli	<u>Steel</u> Broccoli	Steel	Steel→"Yes"	<u>Steel</u> Carrot	Steel	Steel→"Yes"
			Divided-Divided			
Steel Broccoli	Steel Broccoli	Steel Broccoli	Steel→"Yes" Broccoli→"Yes" [Steel Broccoli]→"Yes"	Steel Carrot	Steel Carrot	Steel→"Yes" Carrot→"No"

Note—Words that were cued in training or transfer appear underlined in this table.

broccoli and "yes." At transfer, subjects could respond appropriately using any of their encoded associations when they saw an old pair.

The other possible comparison in which training varied but transfer was constant was focused-focused versus divided-focused. This, however, was not an appropriate test of the encoding assumption because these subjects were in the focused condition at transfer, and focused attention at transfer might "screen out" associations related to the word that was unattended at transfer. That is, even if subjects had encoded associations to both words in training, it might not be possible to detect associations to the word that was unattended at transfer, and so one might underestimate what had been originally encoded.

Obligatory retrieval prediction. According to the obligatory retrieval assumption, one obligatorily retrieves instances from memory that are associated with the stimulus to which one is attending: The object of one's attention acts as a retrieval cue for retrieving associated instances from memory. We tested this assumption by comparing cost for two sets of conditions in which attentional demands in training were held constant, but varied at transfer: (1) focused-focused and focused-divided and (2) divided-focused and divided-divided (Table 2). In each set, subjects encoded the same material during training, and so had the same associations in memory for the transfer block. If the object of one's attention acts as a retrieval cue, subjects who can focus on just the cued item at transfer (i.e., in the focused transfer condition) should retrieve only instances associated with the cued item, whereas those who attend to both items (i.e., those in the divided transfer condition) should retrieve instances associated with both. Because divided transfer condition subjects attend to both words, they should show difficulty with conflicting responses (Table 2). Therefore, subjects in the focused condition at transfer in each set (i.e., focused-focused in Set 1 and divided-focused

in Set 2) were predicted to show less cost than those in the divided transfer condition in each set (i.e., focused-divided in Set 1 and divided-divided in Set 2).

The prediction for the first comparison was that subjects in the focused-focused condition would show less cost than those in the focused-divided condition. In both conditions, subjects attended to a cued word during training, and so encoded few if any distractor words. If, for example, these subjects learned the pair *steel-broccoli* during training, their association to it would be, *steel* and "yes." At transfer, however, the conditions were predicted to differ. If subjects in the focused-focused condition saw the re-paired pair *steel-carrot* at transfer, they should be as fast for this pair as for the old pair, *steel-broccoli*, because both retrieved the association of *steel* and "yes," and subjects were only attending to *steel* at transfer. At transfer, subjects in the focused-divided condition would be attending to both pair members. In attending to both items, they would retrieve associations to any distractors to which they might have attended during training, and this would slow them down on re-paired pairs. Therefore, the focused-divided condition was predicted to be slower at transfer than the focused-focused condition.

The prediction for the second comparison was that subjects in the divided-focused condition would show less cost than those in the divided-divided condition. In both conditions, subjects attended to both words during training, and so encoded distractor words. If, for example, these subjects learned the pair *steel-broccoli* during training, their associations to it would be *steel* and "yes," *broccoli* and "yes," and the pair *steel-broccoli* and "yes." The conditions were predicted to differ, however, at transfer. Here, divided-focused subjects attended to a cued word, and so should retrieve only those instances from memory that were associated with the cued target or nontarget (i.e., *steel* and "yes"). Because divided-divided subjects would attend to both pair members at transfer, they should

retrieve instances associated with both words, and so face conflict on re-paired trials. Therefore, the divided–focused condition was predicted to show less cost than the divided–divided condition.

Old, compatible, and incompatible pairs. We measured subjects' sensitivity to the disruption of pairings in two ways: with the cost measure, and with a specific analysis of the effect of response incompatibility in the re-paired trials. We divided re-paired pairs into two categories: compatible and incompatible. In the former, both words in a pair had been in pairs in training that had been associated with the same response (i.e., "yes" or "no"). For example, assume that subjects saw the pairs *steel–broccoli* and *zinc–lettuce* during training. Both pairs would have been associated with the response "yes," because both contained a member of the target category, metals. If subjects saw the re-paired pair *steel–lettuce* at transfer, this would be considered a compatible pair, because both parts of the pair came from pairs that previously had been associated with the same response (i.e., "yes"). In contrast, incompatible pairs were re-paired pairs in which half of the pair had been in a pair associated with a "yes" response during training, whereas the other half had been in a pair associated with a "no" response. If, for example, subjects had learned the pairs *steel–broccoli* ("yes" response) and *dresser–radish* ("no" response) during training, the re-paired pair *steel–radish* would be an incompatible pair. We predicted that subjects would be fastest on old pairs, because they could take advantage of all associated instances in memory, and slowest on incompatible pairs, because they would be slowed by associations to conflicting responses. Predictions for the compatible pairs varied by condition. We predicted that focused–focused subjects would show no difference between compatible and incompatible pairs because they were attending to only one of the items in each pair (at both training and transfer) and so would not be affected by conflicting associations. In contrast, we predicted that divided–divided subjects would be slower on incompatible than on compatible pairs, because in attending to both pair members at training and transfer, they might retrieve conflicting associations to the words.

The focused–divided and divided–focused conditions presented subjects with both attentional manipulations over training and transfer, so it was not clear which attentional manipulation would be responsible for a given pattern of performance. In the re-pairing analysis, therefore, we examined just the focused–focused and divided–divided conditions, both of which held attentional manipulation constant.

METHOD

Subjects

One hundred fifteen subjects from the University of Illinois at Urbana-Champaign academic community participated in this study. Subjects were tested for color blindness with plates from the Ishihara color blindness test (Ishihara, 1987). Only subjects who responded correctly to the presented plates were included in the study. Five subjects were dropped for using the wrong hand in the exper-

iment or for not following transfer block instructions. Three subjects were dropped because of experimenter error. Ten subjects were removed because of an excessively high error rate, which was defined as an average error rate of over 10% per block over all blocks, for either the target-present or target-absent trials. The data of 96 subjects were included in the experimental analyses (24 per condition). Logan and Etherton (1994) had 32 subjects per condition. Subjects were paid for their participation.

Apparatus and Stimuli

The stimuli were the same as those used in Logan and Etherton (1994). Subjects saw 16 words from each of four categories from the Battig and Montague (1969) norms, for a total of 64 words. The categories were *metals*, *countries*, *vegetables*, and *furniture*. The categories were matched on four measures: frequency of mention (Battig & Montague, 1969), prototypicality (Uyeda & Mandler, 1980), word frequency (Kučera & Francis, 1967), and word length in letters. There were no significant differences on these measures between the categories, except for a statistically significant difference in word frequency for the highest and lowest frequency categories. This difference was not a problem; as Balota and Chumbley (1984) have shown, word frequency is not an important variable in category verification tasks.

Subjects were presented pairs of words on Amdek Model 722 color monitors driven by IBM PC XT and AT computers, and on IBM Model 8513 color monitors driven by IBM PS/2 computers. On each trial, one word was presented above the other. Words were presented, left-aligned, in the middle of the screen. Initial letters appeared in column 33 of rows 12 and 13 of the standard IBM 80 × 24 text screen. Since the first letter of country names had to be capitalized, we capitalized the first letter of all words. Remaining letters were in lowercase. From a distance of 60 cm, single words subtended .48° of visual angle in height and a minimum of .76°, and a maximum of 2.29° in length. Each pair of words subtended about 1.14° of visual angle vertically.

Each trial began with the presentation of two short lines that appeared immediately above and below where the word pair would appear, and that functioned as a fixation and warning display. The display was exposed for 500 msec and then was extinguished and immediately replaced by a word pair. The word pair was presented for 1,000 msec. The screen then went blank for 2,000 msec. Subjects responded to the presence or absence of a word from the target category using the "z" and "/" keys. Subjects' responses were recorded during the 3,000-msec window commencing with the display of the word pair and preceding the display of the next fixation and warning display.

Subjects were given a surprise recall test after they had completed the computer portion of the experiment. Each subject was given a lined sheet of paper and asked to list, in order of recall, all the words they remembered from the experiment.

Procedure

During training, subjects saw 32 pairs of items 16 times. These pairs were grouped into 16 blocks, with each pair occurring once within each block, and with the pairs in a random order within each block. Word pairs remained the same for all training trials, and each word appeared in only one pair. At transfer, each subject saw the 32 pairs he/she had seen earlier (old pairs), and 32 re-paired pairings of these words (re-paired pairs). There were two types of re-paired pairs: compatible (both words had associations to the same response, "yes" or "no"), and incompatible (one word had an association to "yes," and the other to "no"). Approximately half of each subject's re-paired pairs fell in each of these two categories.

Half of the trials in a block contained a word from the target category (target present) and half did not (target absent). The target-present trials contained a word from the subject's target category (e.g., *metals*) and a word from one of the two distractor categories (e.g., *vegetables*, *furniture*). The target-absent trials contained a word

from the subject's nontarget category (e.g., *countries*) and a word from one of the two distractor categories (e.g., *vegetables*, *furniture*). The four categories were used equally often as target, nontarget, and distractor categories. They were assigned their relative roles using a balanced Latin square design.

A given word pair never contained both a target and a nontarget word. Also, half of the targets and nontargets were paired with words from one distractor category, and the other half were paired with words from the other distractor category. These design features were important because they allowed us to pair targets and nontargets equally often with the two distractor categories, and therefore to keep subjects from accurately guessing the identity of the target or nontarget word's category on the basis of a distractor.

Targets and nontargets appeared equally often as the bottom and top words in word pairs. However, once a word had been presented in a certain position (top or bottom), it remained in that position throughout training and transfer. The results of changing position appear to be small, at least relative to the effects of changing pairing (Etherton, 1992). Half of the subjects responded to the target using their right index finger and to the nontarget using their left index finger; half did the opposite.

Subjects received both written and oral instructions. They were told to search two-word displays for a member of a given target category. Subjects in the focused-attention condition were told that they would see a green word and white word presented at the same time, and that they should concentrate on the green word and disregard the white word. They were told that the white word would never be a member of their target category, and that they should respond "yes" by pressing the appropriate key if the green word were a member of their target category, and "no" if it were not. Subjects in the divided-attention condition were told that they would see two words at a time on the computer screen, and that they should press the appropriate key if one of them were from their target category, or the other key if neither were from it. After subjects completed the training blocks, they received a message on their computer screens instructing them to notify the experimenter that they were ready for their next task. If they were continuing in the same attentional demand condition for the transfer block (i.e., focused-focused and divided-divided), they were told to continue doing what they had been doing—that their task had not changed. If subjects switched attentional condition (i.e., focused-divided and divided-focused), they were given the appropriate set of instructions, as explained above.

Design

Training RTs and accuracies were analyzed in a 2 (attention manipulation: focused or divided) \times 2 (target present or absent) \times 16 (block) analysis of variance (ANOVA). Transfer RTs and accuracies were analyzed in a 4 (condition) \times 2 (target same or different) \times 2 (target present or absent) ANOVA. The proportions of items recalled were analyzed in a 2 (condition: focused-focused or divided-divided) \times 3 (item type: target, nontarget, or distractor) ANOVA. For the recall analysis, we calculated the proportion of presented words that each subject recalled from the target, nontarget, and distractor categories. The distractor categories were collapsed into one category as there was no theoretically significant difference between them.

RESULTS

Training

Attentional manipulation. Two measures were used to assess whether or not the attentional manipulation was successful. First, we predicted that focused-condition subjects would be faster than divided-condition subjects in the first block, and they were (673 vs. 734 msec, respec-

tively). Planned comparisons indicated that the focused condition was faster than the divided condition in the first block [$F(1,1410) = 16.12$, $MS_e = 35,838.21$, $p < .001$]. Second, we predicted that focused subjects would recall fewer distractors than would divided subjects, and they did (Table 3). Because subjects in the focused-divided and divided-focused conditions experienced both attentional manipulations, it was not possible to determine which manipulation was responsible for their recall of items. Therefore, the recall data analysis was limited to the focused-focused and divided-divided conditions, both of which held the attentional manipulation constant.

The recall results showed that focused-focused and divided-divided subjects attended to different items. As predicted, focused-focused subjects recalled fewer distractors than did divided-divided subjects, but they also recalled more (cued) nontargets than did divided-divided subjects, even though there was no significant difference in the number of items recalled by subjects in each condition, or in their recall of targets. Subjects in both conditions showed the highest recall level for targets, but focused-focused subjects showed better recall for (cued) nontargets than for distractors, whereas divided-divided subjects showed no difference in their recall of nontargets and distractors, suggesting that they paid equal attention to both categories of items.

An analysis of the recall data showed no main effect of condition [$F(1,46) < 1$, $MS_e = 456.30$], but a significant effect of item type [$F(2,92) = 78.41$, $MS_e = 11,936.71$, $p < .0001$], and a significant interaction of condition and item type [$F(2,92) = 7.65$, $MS_e = 1,163.88$, $p < .001$]. A series of planned comparisons was used to examine this interaction. Although there was no significant difference between the groups on their recall of target items [$F(1,92) < 1$, $MS_e = 152.23$], focused-focused subjects recalled more nontargets than did divided-divided subjects [$F(1,92) = 15.44$, $MS_e = 152.23$, $p < .001$], whereas the latter recalled more distractors than did the former [$F(1,92) = 14.64$, $MS_e = 152.23$, $p < .001$].

We carried out a set of planned comparisons within each condition. Focused-focused condition subjects showed better recall of targets than of nontargets [$F(1,92) = 22.60$, $MS_e = 152.23$, $p < .001$], and better recall of nontargets than of distractors [$F(1,92) = 22.98$, $MS_e = 152.23$, $p < .001$]. Divided-divided condition subjects also showed better recall of targets than of nontargets [$F(1,92) = 65.84$, $MS_e = 152.23$, $p < .001$], but there was no significant difference in their recall of nontargets and distrac-

Table 3
Proportion of Items Recalled per Subject by
Attentional Manipulation and Item Type

Training-Transfer Condition	Item Type		
	Target	Nontarget	Distractor
Focused-focused*	64.6	47.7	30.6
Focused-divided	65.9	45.6	36.8
Divided-focused	58.6	36.5	32.0
Divided-divided*	66.7	37.8	40.2

*Condition on which analyses were performed.

tors [$F(1,92) < 1$, $MS_e = 152.23$]. These results suggest that the attentional manipulation was successful.

Evidence of automaticity. Mean RT \times block was examined across conditions to see if subjects showed increasingly automatic performance. Figure 1 shows that average RTs for both the focused and divided conditions declined across trials: Load effects decreased with practice, suggesting automatization (Logan & Etherton, 1994; Schneider & Shiffrin, 1977). This decrease in load effects was reflected in an interaction between attentional manipulation and block. The greatest differences between the attentional manipulations appeared in the early blocks, when the focused condition was faster than the divided condition. These differences disappeared by the last blocks, when load effects decreased with practice, and RT therefore decreased across blocks.

Average RT across conditions dropped from 776 msec in Block 1 to 627 msec in Block 16, with the largest gains in the early trials, as would be expected if subjects' performance had become automatized. RT data suggested that subjects developed increasingly automatic performance over the 16 training blocks, with the greatest gains in performance in the first few blocks, corresponding to a power function speed-up. This speed-up during training was predicted by the instance theory and is a widely accepted hallmark of automaticity (Anderson, 1982; J. D. Cohen et al., 1990; Logan, 1988; Newell & Rosenbloom, 1981). Fits of a power function model to the RT data produced an r^2 of between .89 and .98 for the two attentional manipulations (focused and divided, respectively). The root-mean squared deviations between predicted and ac-

tual values were 8.88 (focused) and 6.17 msec (divided). The model's parameters for focused training were a (asymptote), 632.07 msec; b (difference between initial and asymptotic trials), 105.28 msec; and c (learning rate), 2.57 msec. The comparable parameters for divided training were a , 604.24 msec; b , 208.78 msec; and c , .93 msec.

The results of an ANOVA supported these analyses. There was a main effect of block [$F(15,1410) = 65.13$, $MS_e = 279,633.00$, $p < .0001$], but no main effect of the attentional manipulation [$F(1,94) < 1$, $MS_e = 104,883.63$]. Like Logan and Etherton (1994), however, we found a significant interaction of attentional manipulation and block [$F(15,1410) = 8.35$, $MS_e = 35,838.21$, $p < .0001$]. Planned comparisons indicated that the focused condition was faster than the divided condition in the first block [$F(1,1410) = 16.12$, $MS_e = 35,838.21$, $p < .001$], but that there was no significant difference between them in the 16th block [$F(1,1410) < 1$, $MS_e = 35,838.21$].

Subjects were sensitive to target presence in training, with target-present pairs being responded to more quickly than target-absent pairs [$F(1,94) = 16.92$, $MS_e = 177,147.00$, $p < .0001$], and a significant interaction of target presence and block [$F(15,1410) = 2.92$, $MS_e = 5,132.37$, $p < .0001$].

Transfer

Attentional manipulation. Table 4 shows mean RTs by condition and target type (i.e., old or re-paired). The attentional manipulation worked at transfer: Subjects in the focused-attention conditions (622 msec average) were 60 msec faster at transfer than were subjects in the divided-

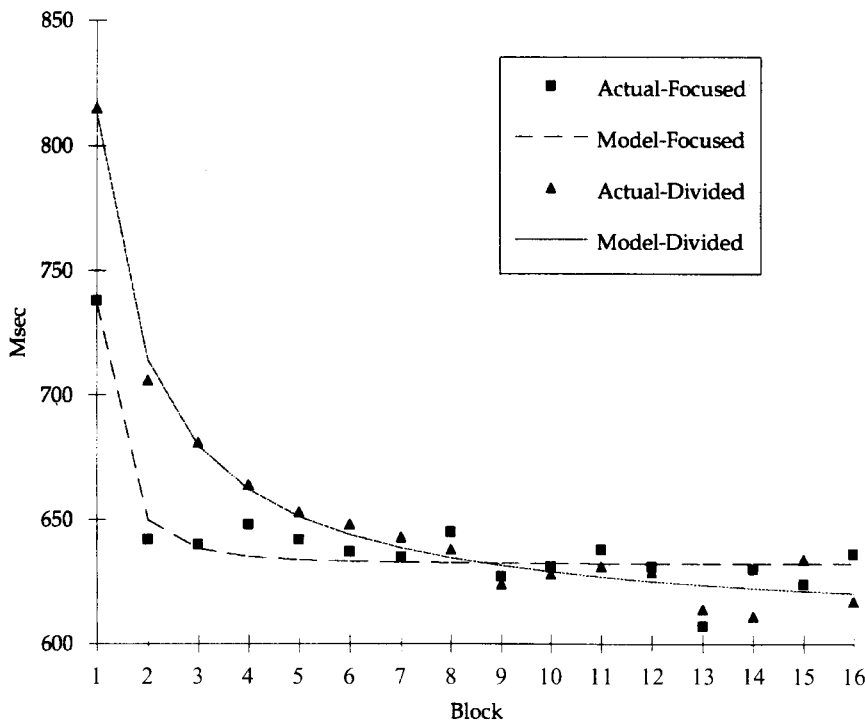


Figure 1. Training reaction times (in milliseconds) by block.

attention conditions (682 msec average). Furthermore, transfer results for the focused–focused and divided–divided conditions replicated the findings of Logan and Etherton (1994): Subjects in the former condition showed less cost than those in the latter condition, suggesting that focused-attention subjects could encode and retrieve associations related to just one of the words in each pair, whereas divided-attention subjects encoded and retrieved associations to both words.

A planned comparison revealed that subjects in the focused-attention condition were faster than those in the divided-attention condition at transfer [$F(1,92) = 273.219$, $MS_e = 2,557.32$, $p < .0001$]. Other planned comparisons revealed that the focused-focused condition showed less cost than did the divided-divided condition both in RTs [$F(1,92) = 64.28$, $MS_e = 2,557.32$, $p < .001$] and in error rates [$F(1,92) = 7.35$, $MS_e = 31.24$, $p < .01$]. These results support the claim that the attentional manipulation worked at transfer.

Obligatory encoding prediction. According to the obligatory encoding assumption, the cost of re-pairing should be greater in the divided–divided condition than in the focused–divided condition. This prediction was confirmed: A planned comparison comparing the cost differences in mean RTs for the focused–divided (25-msec) and divided–divided (66-msec) conditions revealed a significant difference in cost between these conditions [$F(1,92) = 30.95$, $MS_e = 2,557.32$, $p < .001$]. There was no significant difference in error rates between these conditions [$F(1,92) < 1$, $MS_e = 31.24$].

Obligatory retrieval prediction. Both obligatory retrieval predictions were confirmed: Subjects in the focused conditions at transfer showed less cost than those in the divided conditions at transfer. Cost was less for subjects in the focused–focused than for those in the focused–divided condition, and it was less for subjects in the divided–focused condition than for those in the divided–divided condition. In both comparisons, the opportunity to focus on a cued item at retrieval, as well as the opportunity to retrieve from memory instances only (or predominantly) associated with that word, helped performance. These results support a key assumption of the instance theory.

A planned comparison revealed a significant difference in cost between the focused–focused and focused–divided conditions [$F(1,92) = 6.03$, $MS_e = 2,557.32$, $p < .025$]. The same pattern of cost was apparent in error rates, with focused–focused subjects showing less cost

than focused–divided subjects [$F(1,92) = 6.03$, $MS_e = 31.24$, $p < .025$]. These data did not show a speed–accuracy tradeoff: The faster subjects were also the more accurate.

The other retrieval comparison, between the divided–focused and divided–divided conditions, also showed support for the retrieval assumption. Cost was less in the former condition than in the latter [$F(1,92) = 65.94$, $MS_e = 2,557.32$, $p < .001$]. The same pattern of cost was apparent in error rates, with divided–focused subjects showing less cost than divided–divided subjects [$F(1,92) = 11.62$, $p < .005$, $MS_e = 31.24$]. These data did not show a speed–accuracy tradeoff: The faster subjects were also the more accurate ones.

Subjects were sensitive to target presence at transfer, with target-present pairs being responded to more quickly than target-absent pairs [$F(1,92) = 25.73$, $MS_e = 2,557.32$, $p < .001$]. An interaction of target presence and condition approached but did not meet significance [$F(3,92) = 2.42$, $p < .07$].

Old, compatible, and incompatible trials. In addition to comparing cost across the conditions, we measured subjects' sensitivity to the disruption of pairings by examining the effect of response incompatibility in the re-paired trials. In this analysis, we examined the effect of changing pairs at transfer in more detail. We compared three types of pairs: old (pair identical to training pair), compatible (re-paired pair, both parts of which were associated with the same response, i.e., "yes" or "no"), and incompatible (re-paired pair, each part of which was associated with a different response, i.e., one "yes" and one "no"). See Table 5 for a summary of the compatibility data.

According to the obligatory encoding and retrieval predictions, focused condition subjects should show less of a compatibility effect than divided condition subjects, because they could focus on the cued item and disregard the distractor item. As predicted, changing pairing had no effect in the focused–focused condition: Old pairs (633 msec) were no faster than re-paired pairs (642 msec), and compatible pairs (637 msec) were no faster than incompatible pairs (647 msec). Also as predicted, re-pairing did have an effect for divided–divided pairs: Old pairs (647 msec) were significantly faster than re-paired pairs (714 msec), and compatible pairs (684 msec) were significantly faster than incompatible pairs (743 msec).

As predicted, planned comparisons for the focused–focused condition showed that old pairs were no faster than re-paired pairs [$F(2,23) < 1$, $MS_e = 2,734.91$], and

Table 4
Mean Reaction Times (in Milliseconds) and Error Rates, $p(e)$, by Condition

Transfer Pair	Condition (Training–Transfer)							
	Focused–Focused		Focused–Divided		Divided–Focused		Divided–Divided	
	<i>M</i>	<i>p</i> (<i>e</i>)	<i>M</i>	<i>p</i> (<i>e</i>)	<i>M</i>	<i>p</i> (<i>e</i>)	<i>M</i>	<i>p</i> (<i>e</i>)
Re-paired	643	5.4	695	6.5	608	3.8	715	6.7
Old	635	3.8	670	2.8	601	2.8	649	2.9
Cost	7	1.5	25	3.7	6	.98	66	3.7

Table 5
Mean Reaction Times (in Milliseconds) and Accuracy Scores, $p(c)$,
by Attentional Manipulation and Re-Paired Pair Type

Re-Paired Pair Type	Presence of Target Category Item	Condition							
		Focused– Focused		Focused– Divided		Divided– Divided		Divided– Focused	
		<i>M</i>	<i>p(c)</i>	<i>M</i>	<i>p(c)</i>	<i>M</i>	<i>p(c)</i>	<i>M</i>	<i>p(c)</i>
Old	present	627	97.0	650	97.9	633	95.5	583	96.4
Old	absent	639	96.2	682	96.5	660	98.4	620	98.0
Compatible	present	638	90.5	674	89.3	674	91.5	617	94.4
Compatible	absent	636	94.2	691	92.3	694	98.7	608	97.1
Incompatible	present	651	93.3	667	90.1	705	89.8	590	94.4
Incompatible	absent	642	93.0	757	94.0	782	94.0	626	97.3

that compatible pairs were no faster than incompatible pairs [$F(1,23) < 1$, $MS_e = 2,734.91$]. Accuracy data showed significantly higher accuracy for old pairs than for re-paired pairs [$F(2,23) = 6.04$, $MS_e = 38.88$, $p < .01$], but no difference between the re-paired pairs [$F(1,23) < 1$, $MS_e = 38.88$]. These data supported the obligatory encoding and retrieval assumptions.

Planned comparisons for the divided–divided condition showed that old pairs were significantly faster than re-paired pairs [$F(2,22) = 26.87$, $MS_e = 2,548.06$, $p < .0001$], and that compatible pairs were significantly faster than incompatible pairs [$F(1,22) = 16.07$, $MS_e = 2,548.06$, $p < .001$]. Post hoc comparisons showed that the three pair types were all significantly different from one another [$t(22, .05)$, $LSD = 30.87$]. Accuracy was higher for old pairs than for re-paired pairs [$F(2,23) = 4.99$, $MS_e = 36.51$, $p < .025$], but there was not a significant difference between the re-paired pairs [$F(1,23) = 3.31$, $MS_e = 36.51$, $p < .10$]. As predicted, there was a significant RT difference between compatible and incompatible pairs for the divided–divided condition. These data supported the obligatory encoding and retrieval assumptions.

The focused–divided and divided–focused conditions presented subjects with both attentional manipulations over the course of training and transfer, and so interpreting their re-pairing data was not straightforward: Both attentional manipulations contributed to their performance. Subjects in these conditions had the benefit of focused attention at either training or transfer to help them focus on the cued item and disregard the distractor. However, they also either encoded the distractors during training (divided–focused) or were slowed at transfer by any distractors they might have encoded during training (focused–divided). Therefore, our analysis dealt with the focused–focused and divided–divided pairs, for which we could make clear predictions.

DISCUSSION

In the present experiment, we examined the relationship of attention and automaticity; thus we needed to determine that the attention manipulation worked, and that subjects showed automatic performance. Three mea-

asures indicated that the attention manipulation worked; that is, that focused subjects attended primarily to the cued word, and divided subjects attended to both words. First, in the first training block, the focused condition was 78 msec faster than the divided condition, which would be expected if subjects were able to direct their attention to just one word. Second, the transfer block also showed a benefit for focused over divided attention: Subjects in the focused transfer conditions were on average 60 msec faster than were subjects in the divided transfer conditions (622 vs. 682 msec, respectively). Third, the recall data showed that subjects in the focused–focused condition recalled fewer distractors than did those in the divided–divided condition, suggesting that the attentional manipulation was successful in keeping the former subjects focused on the cued item.

Two measures indicated that we achieved some level of automaticity. First, the difference between divided and focused conditions decreased with practice, and second, subjects became faster across training blocks, as predicted by the power function speed-up model (Logan, 1992). Both measures are standard tests of automaticity (e.g., Logan, 1988, 1992; Newell & Rosenbloom, 1981; Schneider & Shiffrin, 1977).

Having demonstrated that our attentional manipulations were effective, and that we achieved automaticity, we can now address the main question that motivated this study—that of whether attention operates at encoding, retrieval, or both encoding and retrieval in automatized performance.

Obligatory Encoding Assumption

Support for the obligatory encoding assumption comes from the finding that subjects were differentially affected by re-paired pairs at transfer, depending on whether or not they were cued to focus on one word in each pair during training. In the present analysis, the focused–divided and divided–divided conditions were compared. Subjects with cues during training (focused) showed less cost at transfer than did subjects without such cues (divided), despite the fact that both were in the same transfer condition (divided). If attention played no role at encoding, subjects should have encoded the same information during training, regardless of attentional manipulation, and

performed comparably on the transfer task. The cost disparity between these conditions at transfer suggests that subjects in these conditions encoded different types or amounts of information during training.

It is evident that attention—not simply visual presentation—was significant here, because the word pairs were presented in such a manner that both words fell in the subjects' primary visual fields. Thus, both words should have been perceived and possibly have activated long-term memory. Apparently, they were either encoded in memory, in contrast with unattended stimuli, or they were encoded better than stimuli to which subjects did not attend.

The fact that the focused group showed some cost suggests that either some unattended material was encoded or that the attentional manipulation was not entirely successful in limiting focused subjects' attention to the cued item. If the former were true, it would indicate that the weak version of the obligatory encoding assumption (that primarily attended, but also some unattended material, is encoded) is correct. However, as noted earlier, it is extremely difficult to rule out the strong version of the obligatory encoding assumption (that only attended material is encoded), because of the difficulty in fully controlling subjects' attention. The strong version could be ruled out only if, after it was demonstrated that subjects paid complete attention to a particular item, there were still an effect for unattended items. In the present experiment, the item to be attended was spatially proximal to the distractor item, and so it was not possible to guarantee that subjects would not store the distractor in memory. Also, to make sure that subjects had sufficient time to make a judgment and encode each pair, each pair was presented for 1,000 msec. This time period may have given subjects time to look at, and store, both items. Although the data indicate that focused–focused subjects attended less to distractors than did divided–divided subjects, the fact that focused subjects recalled *any* distractors indicates that there was not complete control over their attention. Given the difficulty of controlling subjects' attention, this experiment did not discriminate between the strong and weak versions of this assumption, but it did find support for the assumption in general.

Obligatory Retrieval Assumption

The experimental data also supported the instance theory's obligatory retrieval assumption in that subjects were differentially affected by re-paired pairs, depending on whether or not they were cued to attend to one word in each word pair at transfer. In this analysis, we compared the focused–focused condition with the focused–divided condition, and the divided–focused condition with the divided–divided condition. Subjects in each comparison set were in the same condition for training, so they likely had the same representations of the word pairs in memory. If attention played no role at retrieval, subjects should have retrieved the same information at transfer and performed the same on old and re-paired pairs. Instead, transfer data showed greater cost at transfer for divided than for focused conditions. This disparity, as well

as the fact that the focused transfer condition cost was negligible in each comparison, suggests that focused transfer subjects were able to disregard the distractor, attend to the cued word, and use it as a retrieval cue.

Conclusions

The results of this experiment suggest that, during automatization, attention operates at both encoding and retrieval. As assumed in the instance theory, attention limits what will enter memory: What one pays attention to will enter, and what one does not will either not enter or not enter in as strong a manner. Also, as assumed in the instance theory, attention plays a role in retrieval from memory: What one pays attention to acts as a retrieval cue that draws associations out of memory. Logan and Ether-ton (1994) found that attention selects the parts of the display that govern automatic performance, but they did not specify whether this selection process occurred at encoding or at retrieval. The present experiment indicated that attention operates at both points.

The finding that attention operates at both encoding and retrieval supports two of the three basic assumptions of the instance theory of automaticity (Logan, 1988, 1990, 1992), the obligatory encoding assumption and the obligatory retrieval assumption. This support provides further evidence that automaticity is a memory-based phenomenon.

These findings about the relationship of attention and memory have practical implications for memory research. Namely, memory researchers should consider monitoring or controlling their subjects' attention to material more precisely, since differing levels of attention across conditions may be confounded with the variable of interest. Increasing sophisticated techniques for manipulating, measuring, and controlling attention have developed over the last 40 years (see Kramer, Coles, & Logan, 1995, for reviews of this research). Attention researchers may, in turn, wish to consider the relationship of memory and attention, especially the relationship of memory to automatic performance.

REFERENCES

- ANDERSON, J. R. (1982). Acquisition of cognitive skill. *Psychological Review*, *89*, 369-406.
- BALOTA, D. A., & CHUMBLEY, J. I. (1984). Are lexical decisions a good measure of lexical access? *Journal of Experimental Psychology: Human Perception & Performance*, *10*, 340-357.
- 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*, *80*, 1-46.
- BRYAN, W. L., & HARTER, N. (1899). Studies of the telegraphic language: The acquisition of a hierarchy of habits. *Psychological Review*, *6*, 345-375.
- COHEN, A., IVRY, R. I., & KEELE, S. W. (1990). Attention and structure in sequence learning. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *16*, 17-30.
- COHEN, J. D., DUNBAR, K., & MCCLELLAND, J. L. (1990). On the control of automatic processes: A parallel distributed processing account of the Stroop effect. *Psychological Review*, *97*, 332-361.
- CRAIK, F. I. M., & LOCKHART, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning & Verbal Behavior*, *11*, 671-676.

- ETHERTON, J. L. (1992). *Encoding of position and color during the acquisition of automaticity*. Unpublished master's thesis, University of Illinois, Urbana-Champaign.
- HOLLENDER, D. (1986). Semantic activation without conscious identification in dichotic listening, parafoveal vision, and visual masking. *Brain & Behavioral Sciences*, **9**, 1-23.
- ISHIHARA, S. (1987). *Ishihara's tests for colour-blindness*. Tokyo: Kanehara.
- KIRSNER, K., & DUNN, J. (1985). A perceptual record: A common factor in repetition priming and attribute retention. In M. I. Posner & O. S. Marin (Eds.), *Attention and performance XI* (pp. 547-565). Hillsdale, NJ: Erlbaum.
- KRAMER, A. F., COLES, M. G. H., & LOGAN, G. D. (EDS.) (1995). *Converging operations in the study of visual selective attention*. Washington, DC: American Psychological Association.
- KUČERA, H., & FRANCIS, W. N. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown University Press.
- LOGAN, G. D. (1988). Toward an instance theory of automatization. *Psychological Review*, **95**, 492-527.
- LOGAN, G. D. (1990). Repetition priming and automaticity: Common underlying mechanisms? *Cognitive Psychology*, **22**, 1-35.
- LOGAN, G. D. (1992). Shapes of reaction-time distributions and shapes of learning curves: A test of the instance theory of automaticity. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **18**, 883-914.
- LOGAN, G. D., & ETHERTON, J. L. (1994). What is learned during automatization? The role of attention in constructing an instance. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **20**, 1022-1050.
- MORRIS, C. D., BRANSFORD, J. D., & FRANKS, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning & Verbal Behavior*, **16**, 519-533.
- NEWELL, A., & ROSENBLUM, P. S. (1981). Mechanisms of skill acquisition and the law of practice. In J. R. Anderson (Ed.), *Cognitive skills and their acquisition* (pp. 1-55). Hillsdale, NJ: Erlbaum.
- NISSEN, M. J., & BULLEMER, P. (1987). Attentional requirements of learning: Evidence from performance measures. *Cognitive Psychology*, **19**, 1-32.
- ROEDIGER, H. L., III (1990). Implicit memory: Retention without remembering. *American Psychologist*, **45**, 1043-1056.
- SCHNEIDER, W., & SHIFFRIN, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review*, **84**, 1-66.
- STADLER, M. A. (1995). The role of attention in implicit serial learning. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **21**, 674-685.
- UYEDA, K. M., & MANDLER, G. (1980). Prototypicality norms for 28 semantic categories. *Behavior Research Methods & Instrumentation*, **12**, 587-595.

(Manuscript received June 16, 1995;
revision accepted for publication January 18, 1996.)