

Conscious Knowledge and Changes in Performance in Sequence Learning: Evidence Against Dissociation

Pierre Perruchet and Michel-Ange Amorim
Université René-Descartes, Paris, France

Two experiments examined the relation between explicit knowledge and motor performance on the serial reaction time task developed by Nissen and Bullemer (1987). Tests of free recall and recognition of sequence components revealed that reliable explicit knowledge was acquired after an amount of practice that was hardly sufficient to improve mean motor performance. In addition, reaction time improvement was limited to the ending trials of the 3- and 4-trial sequence components that Ss recalled or recognized. These results were replicated in Experiment 3, in which Ss were trained under attentional distraction in the task developed by Cohen, Ivry, and Keele (1990). Overall, these findings undermine the most direct experimental support for the widespread view that conscious knowledge and performance in sequence-learning tasks tap 2 independent knowledge bases in normal Ss.

Nissen and Bullemer's study (1987) and several subsequent studies patterned after the same paradigm (Knopman & Nissen, 1991; Nissen, Willingham, & Hartman, 1989; Willingham, Nissen, & Bullemer, 1989; and with some procedural changes, Cohen, Ivry, & Keele, 1990) claim that people's behavior on serial reaction time (SRT) tasks provides evidence for a dissociation between two forms of knowledge. In these studies, a target stimulus appeared on successive trials at one of the three or four possible positions. Subjects were asked to react to the appearance of the target by pressing a key that spatially matched the location of the target on a keyboard. Unknown to subjects, the same sequence of trials (typically 10) was repeated throughout the sessions. Under these conditions, experimental subjects exhibited reliable improvement in performance as compared with control subjects presented with randomly generated series. This improvement in performance was attributed to the subjects' growing knowledge of the repeating sequence or, more precisely (at least in earlier phases of training), to their knowledge of some chunks of a few trials composing the sequence.

In order to assess whether this knowledge was available to conscious awareness, subjects were submitted to a specially designed *generate task* at the end of training. In the generate task, subjects were instructed to press the key corresponding to the location on which they thought the next target would appear, instead of pressing the key below the present target. The stimulus was displayed until the subject made the correct

prediction for the next trial, or in some experiments (Cohen et al., 1990), the stimulus was displayed at its correct location after the subject's response, whether correct or not. The results showed that experimental and control subjects may differ in reaction time (RT) performance without, or at least before, a reliable between-groups difference can be detected in the proportion of correct predictions. This apparent dissociation between RT and the explicit generation of the repeating sequence lends weight to the claim that performance and conscious knowledge reflect independent databases, referred to as procedural and declarative knowledge.

However, a closer examination of the generate task indicates that the evidence put forward for the lack of explicit knowledge in subjects presented with the repeating sequence is far less compelling than is currently acknowledged.

The Inadequacy of the Generate Task

First, it is worth noting that the instructions given to the subjects before the generate task do not mention reproducing the prior sequences. This lack of information on the relation between study and test phases clearly classifies the generate task as an implicit memory test (Schacter, 1987). Using implicit memorylike instructions in a test devised to measure explicit knowledge is highly paradoxical. As an analogy to the verbal list learning situations, this procedure is equivalent to assessing the explicit knowledge of the words making up the study list through a word-stem completion instead of a cued-recall test.

Even if we assume that subjects really guess that they have to reproduce the prior sequences, the generate task remains particularly insensitive to detecting explicit knowledge because of the correction procedure. On the one hand, introducing trial-by-trial correction makes the task somewhat difficult for the subjects. As pointed out earlier, Nissen and her associates posited that subjects learn isolated chunks of few elements anywhere in the whole sequence, at least during the

The study was supported by the Centre National de la Recherche Scientifique and Université René-Descartes, Paris, France. Experiments 2 and 3 were conducted by Michel-Ange Amorim in partial fulfillment of a doctoral program. We thank Yves-André Fery for his assistance in data collection and J. Baker, A. Cohen, A. Reber, R. Zacks, and one anonymous reviewer for their extensive and helpful comments on an earlier version of the manuscript.

Correspondence concerning this article should be addressed to Pierre Perruchet, Laboratoire de Psychologie Différentielle, Université René-Descartes, 28 rue Serpente, 75006, Paris, France.

early phase of training. To recognize the beginning of any chunk in order to correctly anticipate its ending components, subjects must store the earlier correct locations of the target in working memory and override intrusions from their erroneous predictions. It is likely that at least some subjects experience difficulty in handling this unfamiliar procedure. Note that complete failure to follow instructions are occasionally reported (e.g., Cohen et al., 1990, p. 22; Nissen & Bullemer, 1987, p. 12). On the other hand, the procedure dramatically restricts the number of trials on which subjects' knowledge can be assessed. Because the trial-by-trial correction itself provides information on the correct sequence, only the first few (typically 10) trials of the generate task were used to assess the knowledge acquired during the training phase. This limit clearly had damaging consequences on the reliability of measures. In fact, analysis of subsequent trials was also performed, with the rationale that a better rate of learning of the structured group over the random group would provide a saving score. Although a saving score is indeed a sensitive measure of learning (Nelson, 1978), it is worth noting that the saving score is a measure of implicit knowledge (e.g., Roediger, 1990; Slamecka, 1985). Focusing on saving scores further reinforces the ambiguous status of the generate task with regard to the implicit-explicit memory distinction.

Overview of the Present Study

The experiments here were designed to provide a better description of the explicit knowledge acquired in situations derived from the Nissen and Bullemer (1987) paradigm. Two main changes were made in the generate task. First, our instructions to subjects stressed the relation between study and test phases to make the task a genuine explicit memory test. Second, we suppressed feedback on response accuracy. In the resulting *free generation task*, we instructed subjects to generate a sequence of trials that looks like the sequence they encountered in the training phase, without any form of correction. Thus, subjects may produce the components of the sequence they have memorized without being constrained to replace these components in their correct position within the whole sequence. Note that this way of collecting expression of explicit knowledge is the best match for the manner in which subjects spontaneously express what they have learned; for instance, Willingham et al. (1989) mentioned that subjects who were asked for verbal reports "typically attempted to designate the sequence by pointing to locations rather than by describing the sequence verbally" (p. 1049). An additional advantage of suppressing correction is that the effect of the training conditions may be tracked for all the generated sequences rather than solely for the earliest responses.

In addition to the free generation task, we devised a new recognition procedure of sequence components to assess explicit knowledge in Experiments 2 and 3, for reasons that are made clear later.

We hypothesized that conscious knowledge as measured by appropriate instruments closely parallels RT performance. The parallelism between conscious knowledge and performance was investigated at two levels. The first level pertains to the changes in both variables over the training sessions. Our

contention is that, contrary to the commonly held position, results from SRT tasks provide no evidence for a temporal primacy of performance improvement over explicit knowledge emergence. Parallelism was also investigated on a trial-by-trial basis. We hypothesized that portions of the sequence in which an improvement in RT is observed would coincide with those that are available to subjects' explicit remembering. To the best of our knowledge, this aspect of data has not been systematically investigated in previous studies.

Experiment 1

Summarizing earlier studies patterned after the Nissen and Bullemer (1987) paradigm, Willingham et al. (1989) stated that "evidence for procedural learning of the repeating sequence emerges by the end of the first block of 100 trials. . . . In contrast, subjects have reported not being aware of the repetition until the third or fourth block of trials" (p. 1052). The training procedure used in the following experiment replicated that of Nissen and Bullemer (1987), with the exception that subjects here were submitted to only two blocks of trials. One group of subjects received the 10-trial repeating sequence during the training phase. Another group of subjects received pseudorandom sequences. We assessed the explicit knowledge acquired during the training phase by comparing the performance of both groups on a subsequent free-generation task.

Method

Subjects. Twenty-four students at the Université René-Descartes served as subjects. They were 15 women and 9 men with a mean age of 25 years.

Materials. The target stimuli were asterisks, which were displayed on a cathode-ray tube (CRT) screen under conditions that exactly replicated those of Nissen and Bullemer (1987). Other aspects of the material were highly similar to the Nissen and Bullemer experiment, with two minor changes. First, the four possible positions of the asterisk on the screen were indicated by a permanent mark at the very bottom of the monitor screen. Second, subjects responded by pressing one of the four keys on the bottom row (instead of the top row) of the microcomputer keyboard. This change was introduced to make hand position more comfortable for the subjects. The *V*, *N*, semicolon, and equals sign keys on an AZERTY keyboard, which has a center-to-center distance of 3.8 cm, were used; they were covered with blue patches. The keyboard was positioned so that the keys were aligned approximately with the target locations.

Procedure. The subjects were seated in front of the microcomputer keyboard and the CRT screen. They were instructed to press as fast as possible the key that was below the position where the asterisk had appeared. After the correct key was pressed, the target was removed and, after a 500-ms delay, the next stimulus appeared. All subjects completed two blocks of 100 trials, separated by a break of about 1 min.

Half of the subjects were presented with the repeating 10-trial sequence used in all the Nissen and associates' experiments (D-B-C-A-C-B-D-C-B-A, in which A, B, C, and D designate the four locations from left to right). Each block of trials comprised 10 repetitions of this sequence, but beginnings and endings of the sequences were not marked.

The other half of subjects were tested in a random condition. The location of the target was determined randomly, except that there

was at least one and at most six trials before the target appeared in the same location. The random series were different across blocks and across subjects.

Immediately after this RT task, all of the subjects were submitted to the free-generation task. We asked subjects to generate a series of trials that looked like the series they saw in the preceding phases. We did not inform them of the presence or absence of a repeating sequence at any time during the experiment.

After subjects pressed a key, an asterisk appeared on the screen in the location matching the location of the key on the keyboard. This asterisk remained on the screen until the subjects pressed a key again. This phase ended after 100 self-generated trials.

Data analysis. Reaction times longer than 1,000 ms were discarded and replaced by the overall mean score of the subject. They represented 1.21% of RTs for all subjects and trials.

Results

Reaction time task. We performed an analysis of RTs to (a) determine how early in the course of training performance improvement may be evidenced (this information was relevant to the design of Experiment 2) and (b) explore the parts of the sequence that were learned better or earlier.

To serve the first objective, we averaged RTs of correct responses for each of the 10 sets of 10 trials making up the first block of training trials. As shown in Figure 1, RTs decreased for subjects submitted to the repeating sequence and for subjects submitted to random sequences but reached lower values for the former than for the latter. An analysis of variance (ANOVA) performed on these data with groups as a between-subjects factor and set of trials as a repeated measures factor revealed a significant effect for set, $F(9, 198) = 27.41$, $p < .0001$, but no main effect for group, $F(1, 22) = 1.57$, and

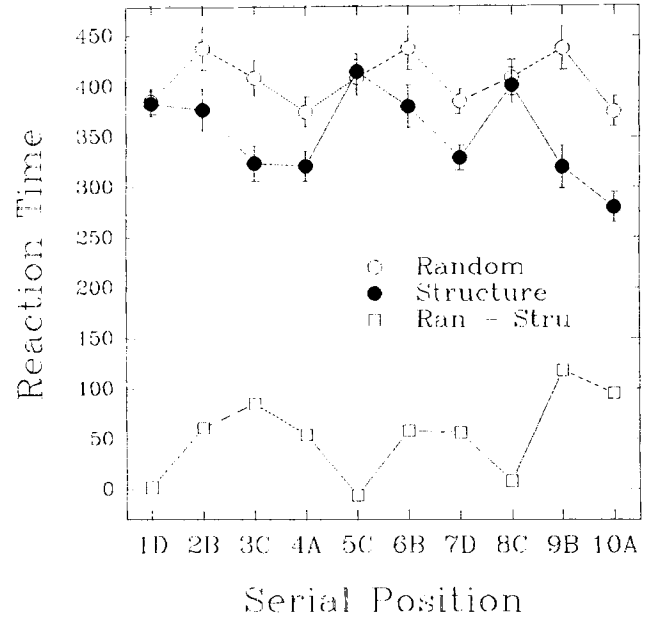


Figure 2. Mean reaction times (RTs) on each serial position of the 10-trial sequence within Experiment 1, Block 2, for the subjects trained with structured sequences ($n = 12$). (For the subjects trained with random sequences, the values represent the mean RTs collected on trials in which the targets were in the location indicated on the x-axis, whatever their serial position. Error bars represent standard errors of the mean. Algebraic differences between each pair of values are also plotted. Positive values represent the RT improvement due to the repeating sequence.)

a marginally significant interaction between groups and trials, $F(9, 198) = 1.78$, $p = .073$. No firm conclusions emerged from paired comparisons between groups on each set of trials as regards the time course of learning. The groups did not differ in RT on Sets 1-5 ($F_s < 1$). However, this result cannot be taken as evidence for absence of learning, given the low power of the statistical tests: Paired comparisons on each set of trials can only detect differences larger than 50 ms (52 to 68 ms, for Sets 2-5), a high threshold value in RT experiments. The F values for Sets 6-10 were higher, but only the difference for Set 6 approached the conventional significance level, $F(1, 22) = 4.19$, $p = .052$.

A reliable difference between structured and random groups was reached in the second block of 100 trials, $F(1, 22) = 6.14$, $p = .021$. We performed an analysis of the level of learning for different components of the sequence on this block by computing the average RT of correct responses for each element in the sequence. Figure 2 shows the mean RTs for each serial position. As pointed out by Nissen and Bullemer (1987, Experiment 4), the resulting measure may be confounded by perceptual and motor effects. A better evaluation calls for comparison with a reference value. Our reference here was the mean RT of subjects from the random group in each of the four possible positions. These reference values differed significantly, $F(3, 33) = 9.45$, $p < .0001$, which confirmed the need for a correction procedure. The signed difference between these reference values and observed RTs

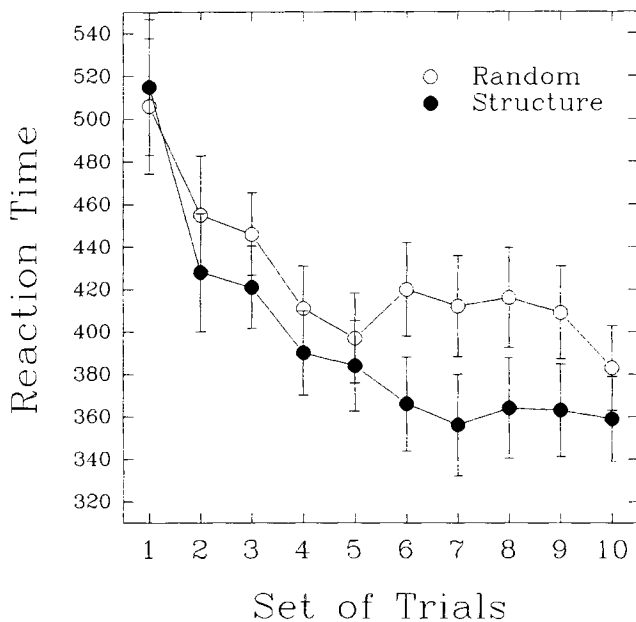


Figure 1. Mean reaction times in each set of 10 trials within Experiment 1, Block 1. (Results are shown for the subjects trained with structured [$n = 12$] and random [$n = 12$] sequences. Error bars represent standard errors of the mean.)

is also reported in Figure 2. Inspection of the data, confirmed by paired comparison, reveals that there was no effect for Serial Positions 1, 5, and 8. Reaction times were reliably shorter than the reference values for other positions ($ps < .05$), except for Position 6, for which the difference was only marginally significant, $F(1, 22) = 3.45, p = .077$. The larger differences were attained for Serial Positions 9 and 10, for which the advantage of the repeating sequence group exceeded 100 ms. These findings are in close correspondence with those reported by Nissen and Bullemer (1987, Experiment 4). The only minor discrepancy between the two sets of data concerns Serial Position 2, for which a difference was obtained only in the present experiment.

One last point warrants consideration. The same pairwise association, namely C-B, occurred at two locations in the sequence. If RTs depended on first-order dependency rules, RTs on Trial B for structure subjects should have been the same in both locations. This was not the case: Mean values differed significantly, $F(1, 11) = 6.47, p = .027$, suggesting that higher order sequential dependency rules need to be taken into account.

Free-generation task. Results on the generation task are displayed in Table 1 for both groups of subjects. For the sake of comprehension, take for example the entries corresponding to the ninth serial position (rows 17 and 18) and the three-trial sequences (column 6). The upper line shows that there were 70 sequences D-C-B in the production from the structured group and that each of the 12 subjects produced at least

1 such sequence. In contrast, the random group (bottom line) generated only 25 sequences, and 9 out of the 12 subjects produced at least 1 sequence. Statistical comparisons were performed on the mean number of generated sequences by t tests when at least 5 subjects in each group produced at least 1 sequence. In the other cases, a Fisher exact test was performed on the number of subjects generating at least 1 sequence. cursory inspection of Table 1 shows that there were an impressive number of reliable differences between both groups.

Consider the first column of numbers, which shows the frequency of placement of the target in each of the four possible positions, without taking into account any prior trials. The subjects from the structured group located more targets in B and C and fewer targets in A and D than other subjects (only the differences for B and D reached significance). These differences reflected a property of the repeating sequence, which had more trials in B (30%) and C (30%) and fewer trials in A (20%) and D (20%) than random sequences (25% in each position). The second column of numbers displays the frequencies of generated two-trial sequences. There were few differences between structured and random conditions with regard to the observed two-trial sequences. Indeed, 9 out of the 12 (4×3) possible transitions were represented in structured conditions. The main difference occurred for the C-B transition, which was displayed twice in each sequence, that is, in 20% of trials, whereas it occurred 100/12, that is, 8.33% in random conditions. This departure was indeed

Table 1
Results on Generation Task in Experiment 1

Serial position	Last trial		Length of sequences							
	Spatial location	Group	1	2	3	4	5	6	7	8
1	D	S	262* (12)	99 (11)	37 (9)	33* (9)	25 (7)*	5 (4)*	1 (1)	0 (0)
		R	315 (12)	112 (12)	22 (10)	12 (8)	4 (2)	0 (0)	0 (0)	0 (0)
2	B	S	306* (12)	85 (12)	44 (10)	22* (8)	22* (8)	17 (6)*	5 (4)*	1 (1)
		R	262 (12)	104 (12)	48 (11)	7 (6)	5 (5)	1 (1)	0 (0)	0 (0)
3	C	S	342 (12)	111 (12)	54 (11)	29 (9)	14 (7)	14 (7)	11 (6)*	5 (4)*
		R	316 (12)	93 (11)	61 (9)	36 (9)	4 (3)	3 (3)	0 (0)	0 (0)
4	A	S	267 (12)	102 (12)	64 (10)	43 (9)	24 (8)	13 (7)*	13 (7)*	10 (5)*
		R	288 (12)	109 (12)	55 (11)	35 (9)	19 (9)	0 (0)	0 (0)	0 (0)
5	C	S	342 (12)	108 (12)	22 (8)	17 (7)	12 (6)	3 (3)	1 (1)	1 (1)
		R	316 (12)	119 (12)	23 (9)	10 (6)	7 (4)	4 (4)	0 (0)	0 (0)
6	B	S	306* (12)	160* (12)	70 (11)	14 (6)	10 (5)	7 (4)	3 (3)	1 (1)
		R	262 (12)	106 (12)	68 (11)	14 (8)	6 (5)	5 (4)	2 (2)	0 (0)
7	D	S	262* (12)	87 (11)	56 (10)	41 (10)	8 (5)	7 (4)	6 (4)	3 (3)
		R	315 (12)	99 (12)	67 (12)	49 (11)	12 (7)	6 (5)	5 (4)	2 (2)
8	C	S	342 (12)	111 (12)	34 (11)	13 (9)	10 (8)	2 (2)	2 (2)	2 (2)
		R	316 (12)	88 (12)	29 (11)	15 (8)	10 (7)	3 (3)	2 (2)	2 (2)
9	B	S	306* (12)	160* (12)	70* (12)	18* (10)	6 (4)	4 (2)	1 (1)	1 (1)
		R	262 (12)	106 (12)	25 (9)	7 (6)	5 (4)	4 (3)	2 (2)	2 (2)
10	A	S	267 (12)	100* (12)	71* (12)	55* (11)	14 (10)*	4 (4)*	2 (2)	0 (0)
		R	288 (12)	63 (12)	24 (11)	8 (5)	1 (1)	0 (0)	0 (0)	0 (0)

Note. Each cell corresponds to a "correct" sequence (i.e., a series of one or several consecutive elements representing a portion of the repeating sequence) whose composition is defined by the ending trial (rows) and the length (columns). The entries represent the total number of sequences generated by subjects in the structured (S) and random (R) groups. The number of subjects, out of 12 for each group, generating at least one sequence appears in parentheses. Statistical comparisons between S and R groups were performed on the number of generated sequences by t tests when at least 5 subjects in each group produced at least one sequence. In the other case, a Fisher exact test was performed on the number of subjects generating at least one sequence.

* $p < .05$, in both cases, but the Fisher exact tests were one-tailed (t tests were two-tailed in the tables as well as elsewhere in the article).

reflected in subjects' generated sequence: The structured group produced 13.3% whereas the random group produced only 8.8% of sequence C-B.

Consider now the sequences of three and more elements. The pattern of reliable differences between structured and random conditions was clear-cut. The subjects trained under structured conditions generated a more or less extended portion of the sequence D-C-B-A-D-B-C-A. Five out of the 12 structure subjects generated the whole sequence at least once. None of the subjects trained under the random condition generated this sequence, which had an a priori probability of $1/(4 \times 3^7)$, that is, .0001 to occur in random generation. The four-trial part of the sequence that revealed the strongest difference between both groups of subjects was D-C-B-A; subject-by-subject analysis indicated that 9 out of the 12 subjects in the structured group and only 1 subject in the random group generated several D-C-B-A sequences. This pattern of results is not surprising: The D-C-B-A component, which forms a simple descending right-to-left sequence across the keyboard, is highly salient, and its intrinsic salience may be reinforced by positional effects, insofar as each of the two 100-trial blocks ended with this subsequence.

RT improvement and generated sequences. Is the knowledge revealed through the free-generation task consistent with the RT pattern? Comparison of both sets of data reveals striking congruencies. The fact that the greatest improvement in RT was observed for Serial Positions 9 and 10 (B-A) was congruent with the highest rate of generation of the four-trial sequence D-C-B-A. Likewise, the lack of difference in RTs for Serial Positions 5 and 8 is consonant with the results of the generation task: There was no difference between the two groups of subjects for these trials as shown in Table 1, whatever the length of the prior sequence taken into account.

In order to obtain a quantitative assessment of the fit between RTs and generation performance, we computed correlations between the two sets of measures over the 10 serial positions. Each pair of values taken into account for the correlations included (a) the mean RT for a given position and (b) the number of n -trial sequences ending on this position. By varying n from 2 to 8, this method yielded information on the number of prior trials that were able to modulate RT on a given trial. The Pearson product correlations were $-.794$ and $-.756$ ($ps < .015$) when the database comprised the three-trial and four-trial sequences, respectively, and were nonsignificant ($ps > .10$) in the other cases. This means that RTs were shorter for trials that ended the most frequently generated three- and four-trial sequences.

A trial-by-trial comparison revealed a few discrepancies. These mainly concerned serial Positions 1 and 7, in both of which the target was in Location D. In Position 1, analysis of RTs indicated no difference, whereas generated sequences did, at least when sequences of more than three trials were taken into account. The reverse dissociation was observed in serial Position 7: Reliable RT modifications occurred without any corresponding trend in free generation.

This latter dissociation warrants closer examination, insofar as it is the most damaging to the general stance of this article. This dissociation may be due to an intrinsic characteristic of the free-generation task: Producing well-learned sequences

necessarily impedes the generation of other possible sequences. As shown in Table 1, structure subjects generated 160 C-B associations and completed the C-B sequences by D in 56 occurrences, that is, in a proportion of .35. Considered in isolation, this value suggests that subjects did not learn about the sequence C-B-D (indeed, if one takes for granted that there is no repetition on the same location, there are three possible locations, each with a chance probability of .33). However, subjects generated an impressive number of sequence C-B-A, namely, 71. Examination of the C-B sequences that were not followed by A ($160 - 71$, i.e., 89) shows that D was preferred to other possible completions (B or C) on a proportion of trials of .63 ($56/89$). This a posteriori analysis suggests that subjects indeed acquired explicit knowledge of the sequence C-B-D but that direct expression of this knowledge was impeded by the high rate of generation of the sequence C-B-A.

Discussion

Our first hypothesis was that subjects would acquire conscious knowledge of substantial portions of the repeating sequence after only two 100-trial blocks of training. Performance in the free-generation task confirms this hypothesis. These results run counter to the contention that no reliable explicit knowledge emerges during such a brief training phase (e.g., Willingham et al., 1989, p. 1052) and support the assumption that there is a common data base for RT improvement and introspective knowledge.

Fine-grained analysis of the relation between RT modifications and explicit knowledge of the sequence reveals that both variables were in close correspondence: The shortest RTs corresponded to the last trials of the three-trial or four-trial chunks that frequently appeared in the free-generation task. The few apparent dissociations could be attributed to intrinsic characteristics of the free-generation task: Generating a high frequency of well-learned sequences could prevent other parts of acquired knowledge from being expressed.

Experiment 2

Although the impressive explicit knowledge revealed in Experiment 1 occurred after only 20 repetitions of the sequence, it is worth noting that the difference in motor performance between structured and random conditions was even more precocious. Taken together with the Nissen and Bullemer (1987, Figure 2) results, the data plotted in Figure 1 suggest that RTs on repeating sequence began to differ from RTs on random sequences by the sixth or seventh set of 10 trials. Experiment 2 replicated Experiment 1, with the exception that in this second experiment the training phase was stopped after six repetitions of the sequence; that is, at the point at which, presumably, evidence for learning from motor performance was just emerging. We anticipated that explicit knowledge could be evidenced after this very limited amount of training.

In addition to using the free-generation task as in Experiment 1, we also assessed explicit knowledge through a recognition procedure of the four-trial chunks composing the re-

peating sequence. The main objective of this new procedure was to eliminate the abovementioned restriction of the generation tasks. In a recognition procedure, knowledge of one sequence can be assessed independently of the knowledge of any other sequences. An additional advantage of a recognition procedure is to provide a more sensitive index of explicit knowledge in comparison with the generation task, the latter being essentially a free-recall test.

Method

Subjects. Thirty-six new subjects from the same pool as for Experiment 1 served as subjects. They were 25 women and 11 men with a mean age of 25 years.

Materials and procedure. The apparatus and stimuli were the same as those in Experiment 1.

Twenty-four subjects were submitted to the RT task with the repeating sequence. There was a single block of six successive 10-trial sequences. Half of the subjects were then submitted to the free-generation task, without any change with regard to Experiment 1.

The other half of the subjects were shown 20 four-trial sequences. These sequences comprised the 10 different four-trial sequences that can be formed with the repeating sequence (D-B-C-A, B-C-A-C, . . . and A-D-B-C). The other four-trial sequences were generated at random with the following restrictions: There should be no successive trials in the same location, and each sequence must differ from the other ones. Old and new sequences were intermixed and displayed in random order. Randomization (for the composition of the new sequences, the order of presentation, etc.) was performed separately for each subject. We instructed subjects to respond to each of the four trials of the sequence as they had done in the RT phase (i.e., by pressing as fast as possible the key that was below the target). We then asked the subjects to indicate on a 4-point scale if they believed that the just-displayed four-trial sequence was part of the series they had seen in the previous phase. The scale was displayed on the screen with adequate labels for its endpoints: *unrecognized* at the left end and *recognized* at the right end. Subjects were asked to respond by pressing one of the same keys they used for the RT task. They were told that about half of the four-trial sequences were part of the series previously displayed.

Another group of 12 subjects was run with random sequences and a subsequent free-generation task as in Experiment 1, except that here the RT phase was reduced to a single block of 60 trials.

Data analysis. Reaction times longer than 1,000 ms were discarded and replaced by the overall mean score of the subject. They represented 4.58% of all RTs.

Results

Reaction time task. Performances of the two experimental groups in the RT task were closely parallel: Neither their mean difference nor their interaction with other investigated factors (rank of trial or serial position) reached significance (F 's < 1). Therefore the data from these groups were pooled for subsequent RT analyses.

Figure 3 shows the mean RT of correct responses of subjects submitted to structured and random sequences for each of the six sets of 10 trials. An ANOVA performed with group as a between-subjects factor and set of trials as a repeated measures factor indicated a significant effect for set, $F(5, 65) = 63.42, p < .0001$, and no main effect for group ($F < 1$). Paired comparisons between groups on each set of trials showed no

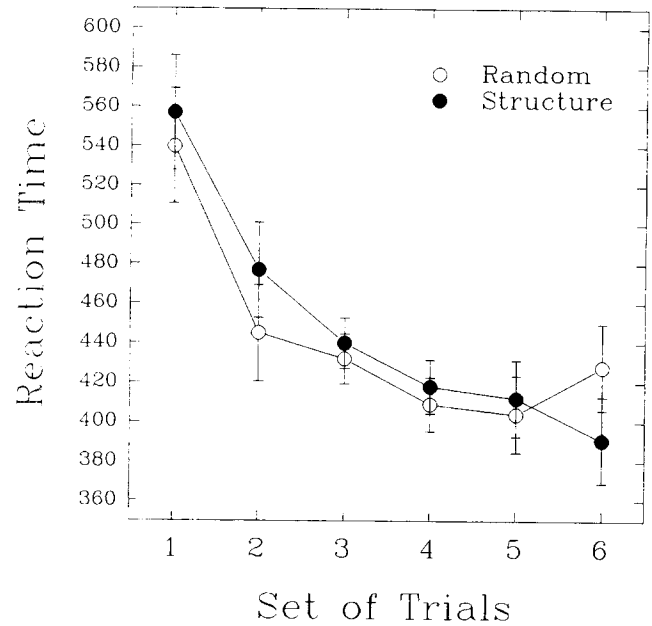


Figure 3. Mean reaction times in each set of 10 trials for the subjects trained with structured ($n = 24$) and random ($n = 12$) sequences in Experiment 2. (Error bars represent standard errors of the mean.)

significant differences, even for the last block, $F(1, 33) = 2.00, p = .17$. Evidence for learning was suggested by a reliable interaction between groups and sets, $F(5, 165) = 2.44, p = .037$. However, this evidence is far from compelling, insofar as the interaction seems to reflect the increase in RT between Sets 5 and 6 in the random group as much as the further drop in RT in the structure group.

Analysis of the serial position effect was performed on Sets 2–6. Figure 4 shows the mean RTs for each serial position, along with their reference values computed as in Experiment 1. Paired comparisons revealed only a marginally significant advantage for RT from the structured sequence for the last serial position, $F(1, 33) = 3.15, p = .085$. However, the whole pattern warrants consideration, insofar as it is highly comparable to the results reported by Nissen and Bullemer (1987, Figure 9, performance of control subjects in Block 1). An intriguing correspondence is that RTs were longer in the structured sequence than in the random sequence for a few serial positions; largest differences were attained for Trials 2 and 5 in both sets of data.

Free-generation task. Table 2 displays the results for the generation task on Experiment 2 in the same way that Table 1 did for Experiment 1. The overall pattern of both tables is similar, although differences between structured and random groups were far less pronounced in Experiment 2. Note, however, that the structured group generated the D-C-B-A sequence more than twice as often as the random group (57 vs. 27 occurrences, respectively). Only differences pertaining to this sequence or part of it reached significance in Experiment 2.

Recognition task. The responses on the 4-point scale were scored from 1 (*unrecognized*) to 4 (*recognized*). As reported

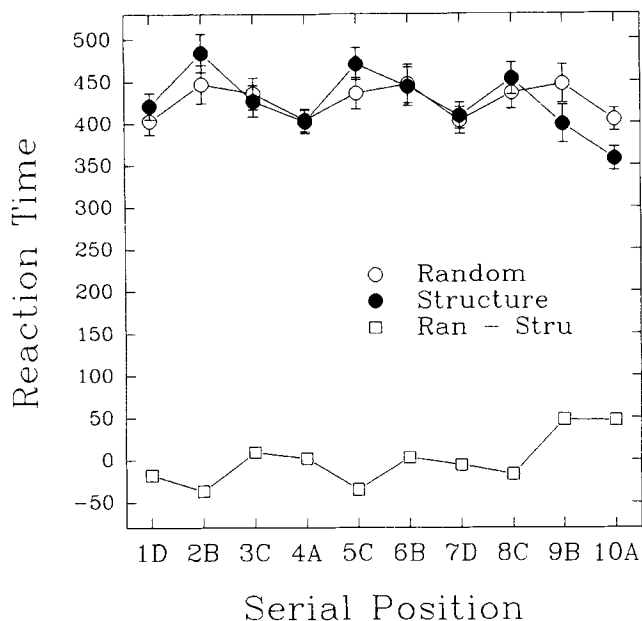


Figure 4. Mean reaction times (RTs) to each serial position of the 10-trial sequence (on Sets 2-6 of Experiment 2) for the subjects ($n = 24$) trained with structured sequences. (For the subjects [$n = 12$] trained with random sequences, the values represent the mean RTs collected on trials in which the targets were in the location indicated on the x -axis, whatever their serial position. Error bars represent standard errors of the mean. Algebraic differences between each pair of values are also plotted.)

in Table 3, the mean score on old sequences (i.e., 4-trial sequences composing the 10-trial repeating sequence) exceeded the mean score on new sequences by about 0.31 scale unit. This difference just failed to reach significance at the conventional threshold, $F(1, 11) = 4.63, p = .054$.

Sequence-by-sequence analysis revealed that four old sequences were significantly better recognized than new sequences. The best scores were found for sequences D-C-B-A and D-B-C-A. These sequences were recognized with certainty (i.e., scored 4) by 9 and 8 (out of 12) subjects, respectively. Another recognized sequence was A-C-B-D. The interest of this specific result stems from the fact that the D trial of this sequence was the only instance in which RT improvement was unaccompanied by clear explicit knowledge in Experiment 1. This apparent dissociation was tentatively explained by intrinsic limitations of the free-generation task. The present results lend considerable weight to this interpretation.

Some old sequences were scored lower than new sequences, although in a nonsignificant way. The lowest score was observed for the sequence B-A-D-B, which was recognized with certainty by only 1 out of the 12 subjects. The particular status of this sequence is discussed in the next section.

It could be argued that the recognition scores do not provide a genuine measure of explicit knowledge. The line of reasoning is that, to make their recognition judgments, subjects perceived the to-be-recognized sequences and responded to the targets in the same way as they did during the training task. Let us suppose that subjects performed better on sequences they had experienced previously than they performed

Table 2
Results on Generation Task in Experiment 2

Last trial			Length of sequences							
Serial position	Spatial location	Group	1	2	3	4	5	6	7	8
1	D	S	303 (12)	97 (12)	49 (12)	38 (10)	27 (9)	9 (4)	2 (1)	1 (1)
		R	303 (12)	105 (12)	32 (10)	25 (10)	10 (5)	2 (1)	0 (0)	0 (0)
2	B	S	308 (12)	71 (12)	35 (9)	17 (9)	12 (6)	11 (5)	3 (2)	0 (0)
		R	301 (12)	97 (12)	44 (10)	8 (5)	8 (5)	3 (2)	0 (0)	0 (0)
3	C	S	300 (12)	84 (12)	37 (10)	20 (7)	9 (6)	7 (4)	6 (3)	2 (2)
		R	297 (12)	113 (12)	52 (11)	25 (9)	5 (4)	5 (4)	2 (2)	0 (0)
4	A	S	275 (12)	67 (12)	25 (12)	22 (10)	10 (7)	5 (4)	4 (3)	4 (3)
		R	286 (12)	87 (12)	46 (11)	36 (11)	14 (9)	2 (2)	2 (2)	1 (1)
5	C	S	300 (12)	97 (12)	7 (5)	1 (1)	1 (1)	0 (0)	0 (0)	0 (0)
		R	297 (12)	103 (12)	13 (6)	10 (5)	9 (5)	2 (2)	0 (0)	0 (0)
6	B	S	308 (12)	153 (12)	61 (12)	5 (4)	1 (1)	1 (1)	0 (0)	0 (0)
		R	301 (12)	130 (11)	63 (11)	10 (5)	8 (4)	8 (4)	2 (2)	0 (0)
7	D	S	303 (12)	119 (12)	68 (12)	43 (12)	2 (2)	0 (0)	0 (0)	0 (0)
		R	303 (12)	117 (12)	58 (11)	38 (11)	7 (4)	6 (3)	6 (3)	1 (1)
8	C	S	300 (12)	118* (12)	46* (12)	16 (9)	11 (7)	1 (1)	0 (0)	0 (0)
		R	297 (12)	79 (11)	27 (11)	7 (5)	5 (3)	1 (1)	1 (1)	1 (1)
9	B	S	308 (12)	153 (12)	79* (12)	24* (10)	12 (5)	8 (4)	1 (1)	0 (0)
		R	301 (12)	130 (11)	44 (11)	10 (8)	3 (3)	2 (2)	0 (0)	0 (0)
10	A	S	275 (12)	97 (12)	73 (11)	57* (10)	16 (7)	8 (3)	5 (2)	0 (0)
		R	286 (12)	65 (11)	44 (11)	23 (7)	5 (3)	1 (1)	0 (0)	0 (0)

Note. Each cell corresponds to a "correct" sequence (i.e., a series of one or several consecutive elements representing a portion of the repeating sequence) whose composition is defined by the ending trial (rows) and the length (columns). The entries represent the total number of sequences generated by subjects in the structured (S) and random (R) groups. The number of subjects, out of 12 for each group, generating at least one sequence appears in parentheses. Statistical comparisons between S and R groups were performed on the number of generated sequences by t tests when at least 5 subjects in each group produced at least one sequence. In the other cases, a Fisher exact test was performed on the number of subjects generating at least one sequence.

* $p < .05$ in both cases, but the Fisher exact tests were one-tailed (t tests were two-tailed in the tables as well as elsewhere in the article).

Table 3
Scores on the 4-Point Recognition Scale in Experiment 2

Sequence	Recognition score	
	<i>M</i>	<i>SD</i>
D-B-C-A	3.17*	1.27
B-C-A-C	1.83	1.19
C-A-C-B	1.83	1.19
A-C-B-D	3.08*	1.24
C-B-D-C	2.42	1.38
B-D-C-B	2.67	1.30
D-C-B-A	3.33*	1.23
C-B-A-D	3.08*	1.31
B-A-D-B	1.50	1.00
A-D-B-C	2.42	1.24
Old sequences	2.53	0.67
New sequences	2.22	0.55

* $p < .05$ and represents the scores that significantly exceed the mean score for the new sequences.

on new sequences. Their recognition judgment could then be based on the analysis of their relative perceptual and motor fluency, in addition to or instead of the actual remembering of the sequences. Subjects would attribute their feeling of readiness to the only explanatory factor available to them: oldness. This account, which derives from theoretical frameworks positing a dual component for recognition judgment (e.g., Jacoby, 1983; Mandler, 1990), cannot easily be ruled out. However, it is worth noting that the RTs for old to-be-recognized sequences did not significantly differ from, and were in fact slightly longer than, RTs for new to-be-recognized sequences ($M = 546$ vs. $M = 541$, respectively; $F < 1$). This result runs counter to a version of this alternative account that construes rapidity of responding as an index of perceptual and motor fluency.

RTs and explicit knowledge. Data collected in this experiment provided us the opportunity to study the parallelism between one measure of motor performance and two measures of explicit knowledge (free recall and recognition of sequences) as a function of serial position. A global assessment of parallelism was obtained by computing correlations of the three measures over the 10 serial positions. The correlations between RTs and generated sequences were computed as in Experiment 1, and results were extremely similar. High correlations were obtained when the pairs of values were composed of (a) the mean RT in a given position and (b) the number of three- and four-trial generated sequences ending in this position (Pearson product correlations = $-.720$ and $-.815$, respectively). When other lengths of sequences were considered, correlations were low and unreliable ($r_s < .33$).

The correlation between RTs and recognition was computed on pairs of values comprising (a) the mean RT at a given position and (b) the mean score for recognition of the four-trial sequences ending at this position. The resulting value was $-.821$. Lastly, the correlation between generated and recognized four-trial sequences was $.814$.

It is worth noting, first, that all these correlations were significant ($p_s < .02$) and fairly high and, second, that pairwise comparisons revealed no reliable differences ($t < 1$; Mc-

Nemar, 1969, p. 158). In other words, there was no evidence that one measure of explicit knowledge correlated better with the other measure of explicit knowledge than with RT performance indices. It could be argued that correlations involving RT measures may have been inflated because the data were collected on the same subjects, whereas the correlation between generation and recognition involved different subjects. Therefore, the former correlations were computed again so that the members of each pair of values came from different subjects (e.g., the RTs collected in the generation group were used for computing the correlation between RT and recognition score). The new values were $-.70$ and $-.87$, respectively, for generation and recognition scores, which did not substantially differ from original values.

The more impressive correspondence among the three sets of measures concerns the last four-trial chunk of the sequence, namely D-C-B-A. This chunk was the most frequently generated and the best recognized in the free-generation task, and the shortest RTs were collected for its terminal trials (B and A).

Although this result yields insights on what is learned in the SRT tasks, inspecting performance on other serial positions is informative as well. As a case in point, consider Serial Position 2. Mean RT for this position was the longest, and the four-trial sequence ending at this position scored the lowest on the recognition test. Both RT and recognition scores were actually worse than their respective reference values (although not significantly). It is noteworthy that this sequence (B-A-D-B) forms a "bridge" between the ending trials and the beginning trials of the best-learned sequences, D-C-B-A and D-B-C-A. Parenthetically, this particular status may explain the frequency of generation of B-A-D-B as an automatic consequence of the generation of these other sequences. This result is suggestive of the fact that forming subjective units in a sequence could be detrimental to the formation of other possible units.

Discussion

The present experiment confirmed the results of Experiment 1 and Nissen and Bullemer (1987) with regard to performance improvement; a difference between RTs collected in repeating and in random sequences just begins to appear by the sixth set of 10 trials. The new, important point is that a clear explicit knowledge of the ending 4-trial sequence of the repeating 10-trial sequence was revealed in the free-generation task after this very limited amount of training. Moreover, a (presumably more sensitive) recognition procedure showed that subjects were able to recognize this and other portions of the sequence.

Analysis focusing on the serial position showed striking correspondences among RTs, free generation, and recognition. Correlations for the three measures were above $.70$. More important, the residual discrepancies can scarcely be put forward as support for a procedural-declarative dissociation, insofar as the alleged measures of procedural and declarative knowledge were not less closely associated than the two measures of declarative knowledge.

Experiment 3

In the sequence used by Nissen and her associates (in all of her studies) and in Experiments 1 and 2, the location of a target on a given trial never determined its location on the next trial. For instance, when the target was in Location B, the next location might be C, D, or A, according to the position of B in the whole sequence. Only second-order or third-order dependency rules were deterministic. In a recent article, Cohen et al. (1990) argued that processing these sequences could differ from processing sequences involving at least some unique pairwise associations. An example of the repeating sequences they used in several of their experiments is A-B-C-B-C. Location C may be followed by B or A, but other associations (namely A-B and B-C) are unique.

Cohen et al.'s (1990) main contention was that attention, although needed for sequences involving multiple-order dependency rules, could be useless for learning unique pairwise associations. This point is somewhat marginal with regard to the present study. However, if the presence of unique pairwise associations within the sequence changes the relations between learning and attention, the same factor may also alter the relations between learning and awareness. Therefore, further empirical investigations are needed to generalize our claim for a conscious access to the knowledge underlying performance improvement to situations involving unique pairwise associations. On this point, Cohen et al. reported that subjects have no conscious knowledge of sequences involving unique pairwise associations even after extensive training on their task. However, this part of their results may be challenged insofar as they assessed explicit knowledge of the repeating sequence by the same generate task as Nissen and Bullemer (1987).

Experiment 3 was designed to shed light on this point. The procedure replicated that of Cohen et al.'s (1990, Experiment 1) difficult tone task. Subjects were submitted to a key-pressing task similar to the one used in Experiments 1 and 2, except that here the repeating sequence comprised unique pairwise associations. In addition, subjects were submitted to a distraction task. A tone was presented after each key press. The tone was high pitched on a majority of trials and low pitched on the remaining trials. Subjects were asked to count the high-pitched tones and report their number. Other minor modifications included in Experiment 3 were designed to provide some methodological improvement over the Nissen and Bullemer (1987) procedure. The repeated sequence, which involved only three possible target locations, differed from subject to subject and, for a given subject, started on a different trial over the successive blocks. The latter change was intended to prevent part of the sequence from becoming salient as a consequence of a positional effect.

In the original Cohen et al. (1990, Experiments 1 and 2) study, the training phase comprised 10 blocks, each block being composed of 20 sets of the five-trial sequence. In the present experiment, the training phase was reduced to 4 blocks. This value was chosen because the Cohen et al. results show that RTs begin to differ between subjects trained on repeating and random sequences at about the 4th block of trials. Thus the present experiment allows for assessment of

conscious knowledge at the start of emergence of motor facilitation. Explicit knowledge was measured by the free recall procedure used in Experiments 1 and 2 and by a slightly modified version of the recognition procedure used in Experiment 2.

Method

Subjects. Thirty-six new subjects taken from the same pool as the previous experiments participated in the experiment. They were 22 women and 14 men with a mean age of 24 years.

Materials. The target was displayed as in the previous experiments, except that there were three possible locations instead of four. The three locations were separated horizontally by 4.5 cm. Responses were made by pressing one of the keys of the numeric keypad of the microcomputer. The keys for the digits 1, 2, and 3 were used.

The tones used for the distractor task were 40 ms in duration. Frequencies of low-pitched and high-pitched tones were 260 and 1040 Hz, respectively (i.e., the C tuning notes with a two-octave interval between them).

Procedure. The general procedure closely replicated that of Cohen et al. (1990). The subjects were seated in front of the microcomputer keyboard and were requested to rest the index, middle, and ring fingers of their right hand on the three response keys. They were instructed to press as fast as possible the key that was in spatial correspondence with the location of the target on the TV screen. After the correct key was pressed, the target was removed, and a tone occurred 40, 80, or 120 ms later. We asked the subjects to count the high-pitched tones and ignore the low-pitched ones, and our instructions emphasized accuracy. The proportion of high-pitched tones ranged from 50% to 75% of the trials. The next target appeared 200 ms after the onset of the key press. All subjects received two 30-trial blocks of practice in dual-task conditions with pseudorandom sequences. Then they completed four blocks of 100 trials. After each block, we asked subjects to report aloud the number of high-pitched tones they counted. The correct response was then displayed on the TV screen.

Subjects were randomly assigned to three different groups of 12 subjects each. For the first two groups, there was a repeating sequence in the training phase. The prototypical sequence was *A-B-C-B-C*. (Italics are used here to indicate that a letter no longer designates a fixed spatial location.) Six specific sequences were formed from the prototype: in keeping with the preceding notation, these were *A-B-C-B-C*, *A-C-B-C-B*, *B-A-C-A-C*, *B-C-A-C-A*, *C-A-B-A-B*, and *C-B-A-B-A*. Two subjects from each group were trained on each sequence. Each 100-trial block started at a randomly selected place in the sequence, and beginnings and endings of the sequence were not marked. The subjects in Group 1 were then submitted to the free-generation task, without any change with regard to the previous experiments.

The subjects in Group 2 were submitted to a recognition procedure similar to the one used in Experiment 2, with the exception that the 20 to-be-recognized sequences varied in length. The 10 old sequences comprised the five different four-trial sequences and the five different five-trial sequences that may be formed with the repeating sequence. For instance, if a subject was trained with *B-A-C-A-C*, the four-trial old sequences included *B-A-C-A*, *A-C-A-C*, *C-A-C-B*, and so on. The 10 new sequences (half four-trial and half five-trial in length) were generated at random with the restriction that the same location could not be used successively. They differed from one another and (at least partially) for each subject. The 20 resulting sequences were presented in random order. Subjects were asked to respond on a 4-point scale by pressing one of the four keys used in Experiment 2.

Subjects in Group 3 served as controls. First, they were trained with random sequences without immediate repetition of an element; then they were submitted to the free-generation task.

Data analysis. Performing the distraction task in this experiment made subjects' RTs to the target considerably longer than those in Experiments 1 and 2. As a consequence, the preceding 1,000-ms criterion for rejecting atypical RTs was increased to 1,500 ms. Any RTs exceeding this criterion were replaced by the overall mean score of the subject. They represented 2.27% of RTs.

Results

Reaction time task. Performances of the two experimental groups in the RT task were similar: Neither their mean difference nor their interaction with other factors (block or serial position) reached significance ($F_s < 1$). Data were pooled for subsequent RT analyses.

The mean RTs of correct responses of subjects submitted to structured and random sequences for each of the four 100-trial blocks are reported in Figure 5. There was a gradual decrease in RT over the trials, as confirmed by the significant effect of blocks, $F(3, 99) = 7.30, p = .0002$. The linear component of the comparison was also significant, $F(1, 33) = 11.58, p = .002$. However, RTs in the structured condition were in fact longer than RTs in the random condition, although nonsignificantly ($F < 1$). Performance of the two groups improved in parallel, as confirmed by the lack of interaction between groups and blocks ($F < 1$).

These results do not imply that the RTs of subjects in the structured condition remained unaffected by the repetition of the sequence. Prior data (Nissen & Bullemer, 1987; intra-Experiment 2) suggest that at the start of training, the RT for

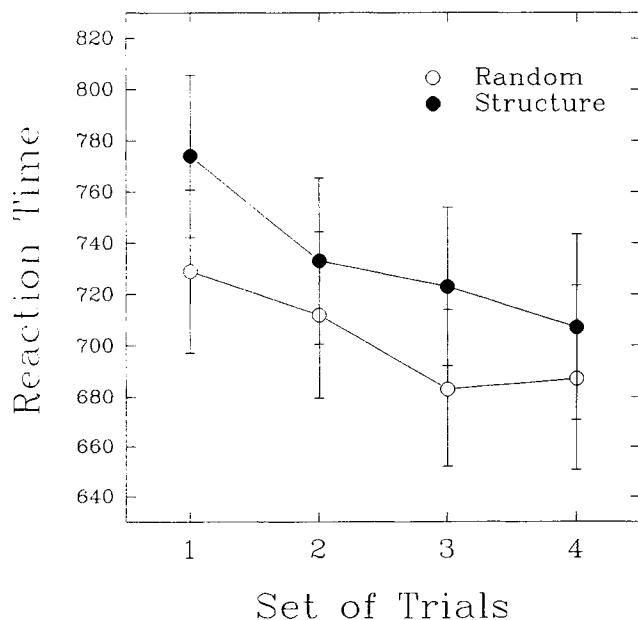


Figure 5. Mean reaction times in Blocks 1–4 for the subjects trained with structured ($n = 24$) and random ($n = 12$) sequences in Experiment 3. (Error bars represent standard errors of the mean.)

Table 4

Mean Reaction Times (RTs) to Serial Positions in the Five-Trial Sequence Within Block 4, for Subjects Trained With Structured Sequences in Experiment 3

RT	Serial position				
	A	B	C	B	C
<i>M</i>	716	700	684	718	719
<i>SD</i>	111	133	126	117	125

some serial position of the repeating sequence may be longer than the RT collected from random sequences, hence partially compensating for the improvement in performance observed on the other serial positions.

An analysis of RT as a function of the serial position was performed on the last block of 100 trials. The resulting mean values are reported in Table 4. Overall, they significantly differed, $F(4, 88) = 2.57, p = .043$. Cohen et al. (1990) claimed that learning unique associations and ambiguous structures call for different mechanisms. Because the repeating sequence contained both unique associations ($A-B$ and $B-C$) and ambiguities ($C-A$ and $C-B$), we hypothesized that the overall difference in RT over serial positions was attributable to this factor. Reaction times collected on the trials following A and B were indeed shorter than RTs collected on trials following C ($M = 701$ vs. $M = 717$), $F(1, 22) = 4.84, p = .038$. However, RTs on the two C positions, both of which consistently followed B , also differed, $F(1, 22) = 8.99, p = .007$. This latter result confirms those of Experiment 1 and suggests that RTs on a trial n is not only affected by the location of the target on trial $n-1$ but is also affected by sequential higher order dependencies.

Free-generation and recognition tasks. As shown in Table 5, there were impressive differences between sequences generated after training in structured and random conditions. All the significant differences were in favor of the former condition, except that structure subjects chose Location A less often than random subjects did; this result reflects the lower frequency of A in comparison with B or C in the repeating sequence. Subjects trained in the structured condition generated the whole sequence more than twice as often as subjects trained in the random condition. The most frequently generated five-trial sequence was $B-C-A-B-C$.

The recognition task also revealed extensive explicit knowledge (see Table 6). All of the subjects scored higher on the 4-point scale for old than for new sequences. The mean difference reached 0.71 scale units, $F(1, 11) = 44.12, p < .0001$. There were no significant differences between the four-trial and the five-trial sequences or within the five series of each length (all $F_s < 1$).

Unlike results in Experiment 2, RTs for the old to-be-recognized sequences were significantly shorter than RTs for the new sequences ($M = 633$ vs. $M = 713$ ms), $F(1, 11) = 8.27, p = .015$. This result is consonant with an interpretation positing that a feeling of perceptual or motor fluency may have led subjects to an attribution of oldness, in addition to, or instead of, the genuine remembering of the training se-

Table 5
Results on Generation Task in Experiment 3

Last trial			Length of sequences				
Serial position	Symbolic location	Group	1	2	3	4	5
1	A	S	341* (12)	178 (12)	129 (12)	55 (11)	42* (11)
		R	396 (12)	181 (12)	89 (12)	36 (12)	15 (12)
2	B	S	441* (12)	209 (12)	127 (12)	89 (11)	34 (10)
		R	396 (12)	181 (12)	89 (12)	45 (12)	17 (12)
3	C	S	405 (12)	275* (12)	153* (12)	86* (11)	64* (10)
		R	396 (12)	181 (12)	89 (12)	45 (12)	25 (9)
4	B	S	441* (12)	199 (12)	122* (12)	72* (12)	48* (10)
		R	396 (12)	181 (12)	75 (12)	37 (12)	18 (12)
5	C	S	405 (12)	275* (12)	107 (12)	64* (12)	37* (12)
		R	396 (12)	181 (12)	75 (12)	32 (12)	15 (12)

Note. Each cell corresponds to a "correct" sequence (i.e., a series of one or several consecutive elements representing a portion of the repeating sequence) whose composition is defined by the ending trial (rows) and the length (columns). The entries represent the total number of sequences generated by subjects in the structured (S) and random (R) groups. The number of subjects, out of 12 for each group, generating at least one sequence appears in parentheses. Statistical comparisons between S and R groups were performed on the number of generated sequences by *t* tests. Italics indicate that letters do not designate a fixed spatial location.
* *p* < .05.

quences. Note, however, that this result is also consistent with our general contention for a parallelism between performance and explicit knowledge.

Overall, the findings on free-generation and recognition tasks stand in sharp contrast with the data reported by Cohen et al. (1990), who found no evidence for explicit knowledge after far more extensive training. A plausible explanation for this discrepancy is that the amount of attention subjects allocated to the tapping task and to the tone-counting task differed between studies. Support for this hypothesis lies in the fact that our subjects made far more errors than those of Cohen et al. in the tone-counting task. In Cohen et al.'s Experiment 1, for instance, only 1 out of 48 subjects reported tone counts in error, on average, by more than 10%; in contrast, 10 out of 36 subjects exceeded this criterion in the present experiment. In addition, Cohen et al. eliminated these

subjects from data analysis to ensure that the final sample included only subjects who paid sufficient attention to the tone task, whereas we performed analyses on the whole sample.

However, additional analyses reveal that the differences in attentional allocation do not explain the discrepancies between Experiment 3 and Cohen et al.'s (1990) experiment. Note first that the just-drawn comparison of error percentages is flawed by the fact that the scores were averaged on 10 blocks in the Cohen et al. experiments and on 4 blocks in ours. The resulting values cannot be compared directly because errors in the tone task decreased drastically across blocks; in our experiment, the mean error percentage decreased from 11.94 in Block 1 to 6.48 in Block 4. Because Cohen et al. did not report their block-by-block results, we cannot draw valid comparisons between experiments in this regard.

What we can explore is whether a difference in tone-counting performance, if real, could explain the differences in the amount of explicit knowledge of the sequence revealed in the two sets of experiments. To this end, we performed separate analyses on subgroups on subjects scoring higher and lower than the median value of tone-counting errors. The findings fail to show any reliable difference in explicit knowledge between subgroups. In the generation task, the 6 subjects making the most errors in tone counting (mean percentage score = 13.44) generated 106 five-trial correct sequences, whereas the 6 subjects making the least errors (mean percentage score = 3.88) generated 119 such sequences; in fact, the trend was in the opposite direction from the expected one. Scores in the recognition task were highly similar: On the 4-point scale of recognition, old sequences were scored 3.10 by subjects making the most errors (11.65%) and 3.06 by subjects making the least errors (4.10%) on the tone task. These results rule out the hypothesis that the differences between our results and Cohen et al.'s (1990) are linked to attentional allocation variations, which purportedly affect the actual amount of

Table 6
Scores on 4-Point Recognition Scale in Experiment 3

Sequence	Recognition score	
	<i>M</i>	<i>SD</i>
A-B-C-B	2.92	1.31
B-C-B-C	3.08*	0.99
C-B-C-A	3.08	1.16
B-C-A-B	2.83	1.19
C-A-B-C	3.42*	0.79
A-B-C-B-C	2.92	1.00
B-C-B-C-A	3.42*	0.90
B-C-A-B-C	3.00	1.04
B-C-A-B-C	3.17	1.34
C-A-B-C-B	3.00*	1.04
Old sequences	3.08*	0.20
New sequences	2.37	0.31

Note. Italics indicate that letters do not designate a fixed spatial location.
* *p* < .05 and represents the scores that significantly exceed the mean score for the new sequences.

subjects' explicit knowledge of the sequence; differences in results stem more probably from the method of assessment of this knowledge.

RTs and explicit knowledge. The correlations between RTs and generated sequences were assessed as in the previous experiments, except that they were computed across 5 serial positions instead of 10. Results were strikingly similar: Correlations were only high for the sequences of three and four trials ($r = -.635$ and $-.938$, respectively). The correlations between RTs and recognition were computed separately for the to-be-recognized sequences of four and five trials. When the pairs of values comprised (a) the mean RT in a given position and (b) the mean score for recognition of the four-trial sequences ending on this position, the correlation coefficient was $-.986$. When the scores for recognition of the five-trial sequences were involved, correlation fell to $-.347$.

The correlation between generated and recognized four-trial sequences was surprisingly low: $.04$. The correlation between generated and recognized five-trial sequences was scarcely higher: $.27$. Thus it appears that correlations between RT and explicit knowledge were far higher than the correlations between the two measures of explicit knowledge themselves. Recall nevertheless that all these correlations were computed across only five pairs of values, so that the resulting values were particularly inaccurate.

Using a trial-by-trial analysis of the parallelism among the three sets of measures may be more informative on what subjects learned in the present situation than using a global assessment of correlations. As a case in point, consider the results reported earlier on RT data: In the sequence *A-B-C-B-C*, RT on the first *C* was shorter than RT on the second *C*, although both occurred after *B*. A fine-grained analysis of free-generation data revealed that the proportion of *C* occurring after *A-B* (153/209, i.e., $.732$) was larger than the proportion of *C* occurring after *C-B* (107/199, i.e., $.537$). The latter value was close to $.50$, which was the expected chance value when assuming that subjects knew that there was no repetition in the same location on consecutive trials. Thus, a joint examination of RTs and generation data suggests that subjects did not learn association *B-C* at all (despite the fact that this association is unique and occurred the most frequently) but rather learned the three-trial chunk *A-B-C*.

In order to tentatively generalize this line of analysis, consider the other serial location on which RT is relatively low, namely the first *B* of the sequence *A-B-C-B-C*. The proportion of *B* occurring after *C-A* (127/178, i.e., $.713$) exceeded the chance value. Further similar analyses for the remaining 3-trial chunks (namely, *B-C-B*, *C-B-B*, and *B-C-A*) revealed that the conditional probability of generating the third element of each chunk given the first two was close to chance ($.414$, $.537$, and $.470$, respectively). Thus the results from both RT and generation tasks may be conveniently accounted for by assuming that subjects learned two 3-trial chunks, *A-B-C* and *C-A-B*, and responded quickly to the last trial of each chunk. However, it remains unclear why subjects learned those chunks rather than the other ones (recall that italic letters do not designate fixed spatial locations, so that it was unlikely that factors such as perceptual and motor facilities came into play).

General Discussion

Summary of Results

The present experiments provide data evidencing a close parallelism between conscious knowledge of the repeating sequence and changes in RT on the very sequence-learning situations in which independence has repeatedly been claimed in recent years. Parallelism was first found in the emergence of changes in the early phase of training. Earlier evidence for temporal primacy of RT change over explicit knowledge (Cohen et al., 1990; Nissen & Bullemer, 1987; Willingham et al., 1989) was clearly disconfirmed. Reliable explicit knowledge of the sequences was revealed through recall-like (Experiments 1 to 3) and recognition (Experiments 2 and 3) tasks after an amount of training that is hardly sufficient to elicit a difference in motor performance between subjects trained under structured and random conditions. This result was obtained when training occurred without attentional distraction with a repeating sequence that included no unique pairwise associations (Experiments 1 and 2), as well as when training occurred under dual-task conditions with a sequence that included unique pairwise associations (Experiment 3).

Parallelism was also evidenced from analyses focusing on the serial position. The trials within the sequences on which RTs were the lowest were the ending elements of the best-recalled or best-recognized chunks of trials. Correlations over the serial positions between RT and explicit knowledge of three- and four-element chunks ranged from $.63$ to $.98$. Fine-grained analysis of the correspondence between explicit knowledge and RT provided an even more impressive idea of the goodness of fit, whatever the structure of the sequence (Experiments 1 and 2 vs. Experiment 3) or the amount of training (Experiment 1 vs. Experiment 2).

Overall, these findings undermine the most direct experimental support for the widespread opinion that performance and awareness tap two independent systems, referred to as procedural and declarative knowledge. However, at least three sets of results may be construed as arguments in defense of independent knowledge bases. These are (a) the few residual dissociations between RT and explicit knowledge that occur in our own data, (b) the related pieces of evidence provided by similar earlier SRT studies, and (c) the data from SRT studies that depart from the Nissen and associates paradigm. These three points are discussed in the following sections.

Although the present experiments were not designed to investigate questions pertaining to which components of the repeating sequence subjects learn in SRT tasks, they provide some interesting insights on this issue. These aspects, which are beyond our main concern, are discussed first.

Pairwise Associations or Longer Chunks?

A cautionary note is in order to prevent unwarranted interpretation concerning the data on the generation and recognition tasks. The fact that subjects trained with the structured sequence generated more correct sequences totalling up to eight elements when compared with subjects trained

with random sequences (in Experiment 1) or that they recognized five-trial sequences (in Experiment 3) does not indicate that they learned such long sequences. This performance may be accounted for by assuming that only small chunks have been memorized, because generation or recognition of long sequences may proceed from the linkage of elementary subunits by random juxtaposition.

If the maximum length of learned chunks cannot be ascertained from our data, their minimum length is a more manageable issue. The present experiments provide straightforward evidence that subjects learned more than simple pairwise associations. If only pairwise associations had been learned, the same results would be obtained when the same pair of consecutive locations occurred at two points in the cycle, which is contrary to the actual findings. Labeling the repeated association A-B, the RT to B in fact differed as a function of the location of A-B within the whole sequence, as well as within the 10-trial sequence used in Experiment 1 and within the five-trial sequence used in Experiment 3. The measures of explicit knowledge provided converging evidence. For instance, analysis of generated sequences in Experiment 3 showed that the probability of generating C after B depended on the location of the target on the trial preceding B. It is worth emphasizing that these results were obtained after a very limited amount of training.

These results run counter to the Cohen et al. (1990) position, which conferred a leading role to pairwise associations in sequence learning. Note that the data Cohen et al. presented in support of their view can be accounted for by assuming that the relevant subjective units are made up of more than two elements. For instance, a cursory comparison between the sequences including and not including unique pairwise associations reveals that the three-trial chunks making up the sequences differ in their degree of resemblance: They are more similar in the latter case than in the former case, hence prompting confusion errors and interference.

Interpreting the Residual Dissociations Between Awareness and Performance

Although impressive, the correspondence between conscious knowledge and RT exhibited in our experiments was not perfect. For instance, correlations over serial positions never reached unity. Are the residual dissociations a sufficient argument to revise our contention about parallelism? Two remarks are in order.

First, the clearest dissociations between explicit knowledge and RT improvement occurred to the detriment of the latter variable. Thus in Experiment 3, free-generation tasks and recognition tests revealed reliable explicit knowledge of the repeating sequence, whereas the mean RT of the structure subjects was not shorter than the mean RT of the random subjects. It is worth emphasizing that there is no symmetry between the two kinds of possible dissociations with regard to their theoretical implications. What is challenged here is the claim that RT improvement reveals knowledge on the structured sequence unavailable to conscious thought. We have no objection to the reverse contention and fully agree that some pieces of knowledge about the sequence may have no effects

on RT. For instance, we fail to see how knowing that sequences are composed of 10 elements rather than 8 or 12 may directly affect RT.

The second remark is that some amount of dissociation is inherent to experimental data, inasmuch as the measure of all postulated variables is affected by a number of specific factors, random variations, and errors of measurement. Defining the cutoff point between the amount of dissociation that may be reasonably attributed to these undesirable sources of variation and the amount of dissociation that testifies to separate underlying systems is not only a matter of subjective preference. One way of providing an empirical criterion consists of collecting multiple indicators for each of the postulated systems. The rationale is that closer associations should be found between indicators of the same system than between indicators of different systems. Although we did not intend to systematically implement this methodology in the present experiments, two indicators of explicit knowledge were available in Experiments 2 and 3 (free generation and recognition). It is important to note that their association was not closer than the associations between each of these indicators and motor performance. Overall, these considerations led us to discard the argument that the residual dissociations between RT and explicit knowledge in our own data provide some support to the claim for independent knowledge bases.

Demonstration of Independence in Earlier Similar SRT Studies

Our results are in sharp contrast to previously published reports. For instance, Willingham et al. (1989, Experiment 2) indicated no difference in generation accuracy between groups trained in structured and random conditions for 100 trials, whereas we found in Experiment 2 that subjects trained in the structured condition generated the ending four-trial chunk of the repeating sequence more than twice as often as subjects trained with random sequences after only 60 trials of training. Still more striking is the fact that although Cohen et al. (1990, Experiment 1) reported no difference in generation accuracy between groups trained in structured and random conditions over ten 100-trial blocks under attentional distraction, we found in Experiment 3 that subjects trained in structured conditions generated the whole repeating sequence more than twice as often as subjects trained with random sequences after only four blocks of training in the very same experimental setup.

These differences can be readily explained by comparing the ways explicit knowledge was assessed in both sets of studies. When compared with earlier studies, our subjects were explicitly instructed to generate a sequence that looked like the one they encountered in the training phase, and they were not provided any informative feedback. The present experiments were not designed to dissociate the effects of these two procedural changes. Cohen et al. (1990) reported briefly on a supplementary experiment in which subjects were instructed on the relation between study and test phase, from which they concluded that "lack of explicit instructions in the previous experiments apparently did not affect the results" (p. 27). If this were the case, the discrepancy between our

findings and earlier ones would be due to the suppression of the trial-by-trial correction procedure. (The reasons for assuming that introducing a correction procedure impedes expression of explicit knowledge are detailed in the introduction.) However, the conclusion that Cohen et al. drew from this supplementary experiment calls for caution, because subjects were exposed to two blocks of random events near the end of the training phase to assess learning from RT data. This procedure may have had unknown consequences on the subsequent measure of subjects' explicit knowledge of the repeating sequence. Thus, the respective roles of the two procedural changes we introduced in the generation task remain to be elucidated.

In addition to using group data, Nissen and Bullemer (1987) and Willingham et al. (1989) performed subject-by-subject analyses. They reported that some subjects exhibited (in their terminology) declarative knowledge without procedural knowledge and, more important for the main issue of this article, procedural knowledge without declarative knowledge. The authors claimed that these subjects are evidence for independent knowledge bases. Our results are sufficient to invalidate the empirical basis of the argument, inasmuch as they demonstrate the lack of sensitivity of the tests used in these studies. Moreover, we believe the rationale of their argument is basically flawed.

Formally, the method consists of dichotomizing the scores of procedural knowledge on the one hand and the scores of declarative knowledge on the other to assign the subjects to a fourfold contingency table. Then, major theoretical conclusions are drawn from the observation that some subjects fall into the discordant cells of the contingency table or, in other words, that the correlation is not perfect. The problem with this method is that a preliminary condition for obtaining a perfect correlation is perfect reliability of measures, a condition that is highly unrealistic in psychological investigations, whatever the variables. In the context of SRT tasks, the measures are especially prone to error variation. Improvement in RT is assessed on a few responses, because RT is measured on the training trials whose number is limited by obvious study constraints, and estimation of explicit knowledge must be short enough to impede forgetting and interference effects. As a consequence, the dissociation between performance and awareness observed in 1 subject is highly unreliable. Although we assume that this assessment will meet with no objections, we also feel that evidence is perceived as far more compelling when dissociations are reported for a subgroup of subjects, as in the Nissen and Bullemer (1987) and Willingham et al. (1989) studies. It is worth emphasizing that the confidence attributed to group data in the present case is unwarranted. Considering the cells of a contingency table in isolation is uninformative; the only valid conclusion that may be drawn would require the computation of a contingency coefficient (naturally involving the four cells of the table) and its evaluation with regard to the reliability coefficients of the two variables.

Demonstration of Independence in Other SRT Studies

Although most SRT studies are patterned after the Nissen and Bullemer (1987) paradigm, a few others depart notably

from it. The main difference is that the successive locations of the target follow complex deterministic rules, instead of being determined by the simple repetition of a fixed sequence. One of the most frequently cited studies in this field of research is Lewicki, Hill, and Bizot (1988), in which the location of the target is determined by multiple second-order dependency rules. To discover (and subsequently apply) these rules, subjects must have learned to partition the whole sequence into "logical" blocks of five trials, because the selection of specific rules depends on the serial position of the trials within these logical blocks. The findings indicate an improvement in motor performance, although, according to Lewicki et al., subjects were unable to mention anything even remotely similar to the manipulated pattern.

Perruchet, Gallego, and Savy (1990) proposed a far more parsimonious explanation for these data. They showed that the observed improvement in performance can be accounted for by subjects' sensitivity to the relative frequency of particular target transitions on the whole sequence of trials. Replication and reanalysis of the Lewicki et al. (1988) experiment indicated that one of the main pieces of knowledge that actually underlied performance improvement was that the target tended to move through all its possible locations before returning to a previously occupied one. Of major interest in the present context, this knowledge was available to subjects' consciousness, as revealed by a subsequent explicit prediction task.

These findings do much more than invalidate the demonstration in the Lewicki et al. (1988) target article. Through the detailed analysis of an example, the Perruchet et al. (1990) article points to a major bias likely to plague all studies aimed at investigating the explicit-implicit or declarative-procedural distinctions from performance in situations structured by highly complex rules. Subjects can tackle the situation successfully in a way that has no obvious relation to the composing rules. As a consequence, improvement in performance without any explicit knowledge of such rules does not imply that subjects engage in unconscious processes.

Conclusion

Recent influential studies claim that motor performance and awareness of the structured sequence in the early phase of sequence-learning tasks in normal subjects tap two independent databases, generally referred to as procedural and declarative knowledge (e.g., Cohen et al., 1990; Lewicki et al., 1988; Nissen & Bullemer, 1987; Willingham et al., 1989). The major conclusion of this article is that this claim is groundless. Neither studies in which a single sequence is simply repeated nor studies in which sequences are the end product of complex composition rules provide reliable empirical support for a dissociation. Of course, this article does not demonstrate that explicit knowledge of the sequence parallels changes in performance in all the possible situations—a contention that goes beyond any direct experimental inquiry—but it does show that the available studies offer no counterevidence.

At least three additional remarks are in order to prevent overgeneralization or misinterpretation. First, it must be emphasized that our concern encompasses only the early phase of learning. This specification is important insofar as much

of the intuitive, everyday evidence for a performance–awareness dissociation stems from familiar situations. Several experimental investigations confirm that performance and verbalizable knowledge may be dissociated with extensive training in SRT tasks. Thus for instance, although Cleeremans and McClelland (1991) reported that significant explicit knowledge of the composition of their sequences is revealed by a generation task after 20 training sessions of 3,100 trials each, other studies (Lewicki, Czyzewska, & Hoffman, 1987; Stadler, 1989) fail to show any evidence of explicit knowledge after a comparable amount of training. However, neither introspective reports nor experimental data on overlearned situations run counter to our main conclusion.

Second, the parallelism between performance and conscious knowledge in the early phase of training does not necessarily imply a functional dependency relation between the two variables. Our data are obviously compatible with the idea that the acquisition of conscious knowledge exerts some causal influence on performance changes. However, another interpretation is that normal individuals have conscious access to the knowledge base underlying performance improvement, or in other words, that performance and conscious knowledge tap a common knowledge base. It could be argued that data from neurologically impaired patients fit better with the latter than with the former interpretive framework, but space limitations prevent us from discussing the point here.

Finally, our conclusion does not exclude the fact that learning in the situations involved here may be implicit in some respects. As stressed elsewhere (Perruchet & Pacteau, 1991), the term *implicit* when associated with learning is used in two different ways. Implicit may refer to the resultant memorial representation of the learning process. The main claim of this article is that learning, as investigated in typical SRT paradigms, is not implicit in this regard. However, implicit can also refer to the learning process itself. In this latter case, it is virtually synonymous with *automatic*, as used in the context of frequency coding (Hasher & Zacks, 1979). A process is said to be automatic when it requires few or no attentional resources and is out of intentional control. It is conceivable that an automatic mode of acquisition may generate knowledge that is available to conscious thought. Although this aspect was beyond our main concern, our study provides some empirical support for this hypothesis, insofar as subjects acquired explicit knowledge of a sequence presented under attentional distraction (in Experiment 3).

The scope of the present article was limited to sequence-learning tasks. It is worth noting that a critical reappraisal of the evidence for a dissociation between performance and explicit knowledge has been undertaken in germane areas of research, with a similar outcome. For instance, in artificial grammar learning situations (see Reber, 1989, for a review), several recent works (Brooks & Vokey, 1991; Dienes, Broadbent, & Berry, 1991; Dulany, Carlson, & Dewey, 1984, 1985; Perruchet, Gallego, & Pacteau, 1992; Perruchet & Pacteau, 1990, 1991) challenge earlier interpretations of experimental data framed in terms of unconscious abstraction of grammatical rules. Likewise, it appears that the dissociation between performance and conscious knowledge that studies on interactive control tasks (e.g., Berry & Broadbent, 1988) were intended to illustrate could be far less clear-cut than has been

previously acknowledged (Green & Shanks, 1991; Marescaux & Karnas, 1991; Sanderson, 1989, 1990; see also Shanks & Dickinson, 1991). Taken together with the present data, these studies challenge the widespread idea that changes in performance of a subject faced with a new structured situation reveal acquisition of some pieces of knowledge about the structure of this situation that are unavailable to conscious thought. These convergent research trends in the implicit learning area (see also Dulany, 1991; Greenwald, 1992) suggest that the relations between learning processes and consciousness are headed for thorough reappraisal in the near future.

References

- Berry, D. C., & Broadbent, D. E. (1988). Interactive tasks and the implicit–explicit distinction. *British Journal of Psychology*, *79*, 251–272.
- Brooks, L. R., & Vokey, J. R. (1991). Abstract analogies and abstracted grammars: Comments on Reber (1989) and Mathews et al. (1989). *Journal of Experimental Psychology: General*, *120*, 316–323.
- Cleeremans, A., & McClelland, J. L. (1991). Learning the structure of event sequences. *Journal of Experimental Psychology: General*, *120*, 235–253.
- Cohen, A., Ivry, R. I., & Keele, S. W. (1990). Attention and structure in sequence learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*, 17–30.
- Dienes, Z., Broadbent, D., & Berry, D. (1991). Implicit and explicit knowledge bases in artificial grammar learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*, 875–887.
- Dulany, D. E. (1991). Conscious representation and thought systems. In R. S. Wver & T. K. Srull (Eds.), *Advances in social cognition* (Vol. 4, pp. 97–120). Hillsdale, NJ: Erlbaum.
- Dulany, D. E., Carlson, A., & Dewey, G. I. (1984). A case of syntactical learning and judgment: How conscious and how abstract? *Journal of Experimental Psychology: General*, *113*, 541–555.
- Dulany, D. E., Carlson, A., & Dewey, G. I. (1985). On consciousness in syntactic learning and judgment: A reply to Reber, Allen, and Regan. *Journal of Experimental Psychology: General*, *114*, 25–32.
- Green, R. E. A., & Shanks, D. R. (1991). *On the existence of independent learning systems: An examination of some evidence*. Manuscript submitted for publication.
- Greenwald, A. G. (1992). New look 3: Unconscious cognition reclaimed. *American Psychologist*, *47*, 766–779.
- Hasher, L., & Zacks, R. T. (1979). Automatic and effortful processes in memory. *Journal of Experimental Psychology: General*, *108*, 356–388.
- Jacoby, L. L. (1983). Perceptual enhancement: Persistent effects of an experience. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *9*, 21–38.
- Knopman, D., & Nissen, M. J. (1991). Procedural learning is impaired in Huntington's disease: Evidence from the serial reaction time task. *Neuropsychologia*, *29*, 245–254.
- Lewicki, P., Czyzewska, M., & Hoffman, H. (1987). Unconscious acquisition of complex procedural knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *13*, 523–530.
- Lewicki, P., Hill, T., & Bizot, E. (1988). Acquisition of procedural knowledge about a pattern of stimuli that cannot be articulated. *Cognitive Psychology*, *20*, 24–37.
- Mandler, G. (1990). Your face looks familiar but I can't remember your name: A review of dual process theory. In W. E. Hockley &

- S. Lewandowsky (Eds.), *Relating theory and data: Essays on human memory in honor of Bennet B. Murdock*. Hillsdale, NJ: Erlbaum.
- Marescaux, P.-J., & Karnas, G. (1991). *The implicit versus explicit knowledge distinction revisited: When finding associations between verbalizable knowledge and some performance criteria*. (Report No. 4PR3GK of the Knowledge Acquisition and Use in Dynamic Task Environments ESPRIT project, Basic Research Action 3219)
- McNemar, Q. (1969). *Psychological statistics* (4th ed.). New York: Wiley.
- Nelson, T. (1978). Detecting small amounts of information in memory: Savings for nonrecognized items. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 4, 453-468.
- Nissen, M. J., & Bullemer, P. (1987). Attentional requirements of learning: Evidence from performance measures. *Cognitive Psychology*, 19, 1-32.
- Nissen, M. J., Willingham, D., & Hartman, M. (1989). Explicit and implicit remembering: When is learning preserved in amnesia? *Neuropsychologia*, 27, 341-352.
- Perruchet, P., Gallego, J., & Pacteau, C. (1992). A reinterpretation of some earlier evidence for abstractiveness of implicitly acquired knowledge. *Quarterly Journal of Experimental Psychology*, 44A, 193-210.
- Perruchet, P., Gallego, J., & Savy, I. (1990). A critical reappraisal of the evidence for unconscious abstraction of deterministic rules in complex experimental situations. *Cognitive Psychology*, 22, 493-516.
- Perruchet, P., & Pacteau, C. (1990). Synthetic grammar learning: Implicit rule abstraction or explicit fragmentary knowledge? *Journal of Experimental Psychology: General*, 119, 264-275.
- Perruchet, P., & Pacteau, C. (1991). The implicit acquisition of abstract knowledge about artificial grammar: Some methodological and conceptual issues. *Journal of Experimental Psychology: General*, 120, 112-116.
- Reber, A. S. (1989). Implicit learning and tacit knowledge. *Journal of Experimental Psychology: General*, 118, 219-235.
- Roediger, H. L. (1990). Implicit memory: Retention without remembering. *American Psychologist*, 45, 1043-1056.
- Sanderson, P. M. (1989). Verbalizable knowledge and skilled task performance: Association, dissociation, and mental models. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15, 729-747.
- Sanderson, P. M. (1990). *Implicit and explicit control of a dynamic task: Empirical and conceptual issues* (Tech. Rep. No. EPRL-90-02). University of Illinois at Urbana-Champaign, Engineering Psychology Research Laboratory.
- Schacter, D. L. (1987). Implicit memory: History and current status. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 13, 501-518.
- Shanks, D. R., & Dickinson, A. (1991). Instrumental judgment and performance under variations in action-outcome contingency and contiguity. *Memory & Cognition*, 19, 353-360.
- Slamecka, N. J. (1985). Ebbinghaus: Some rejoinders. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 11, 496-500.
- Stadler, M. A. (1989). On learning complex procedural knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15, 1061-1069.
- Willingham, D. B., Nissen, M. J., & Bullemer, P. (1989). On the development of procedural knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15, 1047-1060.

Received April 19, 1991

Revision received October 4, 1991

Accepted October 18, 1991 ■