On the Development of Procedural Knowledge

Daniel B. Willingham
Harvard University

Peter Bullemer
University of Minnesota

Mary Jo Nissen
University of Minnesota

Amnesic patients demonstrate by their performance on a serial reaction time task that they learned a repeating spatial sequence despite their lack of awareness of the repetition (Nissen & Bullemer, 1987). In the experiments reported here, we investigated this form of procedural learning in normal subjects. A subgroup of subjects showed substantial procedural learning of the sequence in the absence of explicit declarative knowledge of it. Their ability to generate the sequence was effectively at chance and showed no savings in training. Additional amounts of training increased both procedural and declarative knowledge of the sequence. Development of knowledge in one system seems not to depend on knowledge in the other. Procedural learning in this situation is neither solely perceptual nor solely motor. The learning shows minimal transfer to a situation employing the same motor sequence.

This article deals with the distinction between procedural and declarative memory systems. Procedural memory is thought to support the acquisition and retention of skilled performance and is indexed by tasks in which memory is expressed implicitly by changes in performance as a result of prior experience. In contrast, declarative memory is thought to support the learning and retention of facts and the recollection of prior events. It is indexed by memory tests such as recall and recognition that require the explicit remembering of a prior episode.

This distinction has received some attention from philosophers (Ryle, 1949) and computer scientists (Winograd, 1975), but only recently have psychologists begun to explore it systematically (Anderson, 1983, 1987; Cohen & Squire, 1980; Squire, 1986, 1987; Tulving, 1985). Despite the obsession of psychology with learning during the first half of this century, the hypothesis that there are two systems of human learning, one available and the other unavailable to awareness, was not seriously examined. The issue tended to be dichotomized so that one either believed in the former system or the latter, but not both. The absence of a theory encompassing both systems negated the motivation to compare and contrast them (see Mishkin & Petri, 1984, for a discussion of this problem). Finally, the relative neglect of learning in general by cognitive psychology has kept the distinction between procedural and declarative learning from being explored.

Now that cognitive psychologists have turned their attention to issues related to learning, the differentiation of procedural and declarative knowledge has become central in theoretical work in the area of human memory and cognition. In Anderson's (1983) ACT* system, information in a declarative memory system is represented by a propositional network, whereas information in a separate procedural memory system is represented as condition-action rules. Tulving (1985) has distinguished between procedural knowledge on one hand and episodic and semantic knowledge on the other, proposing that these types of knowledge are associated with different types of consciousness.

Neuropsychology has provided much of the initiative in the current empirical examination of procedural and declarative memory. Milner (1962) and Corkin (1968) reported that severely amnesic patients could learn motor skills without the ability to recollect the episodes in which they learned them. Cohen and Squire (1980) extended this finding to a task with relatively minimal motor involvement (mirror reading), showing that amnesic patients learned and retained the ability to read mirror-reversed words efficiently, yet were severely impaired in recognizing those words. These developments in neuropsychology with respect to the procedural-declarative distinction are important in two ways. First, they show that the distinction has a neurological basis, with lesions of medial temporal or diencephalic structures impairing declarative but not procedural learning (Squire, 1986). Second, they suggest new methods for studying these systems in relative isolation.

Using a task developed by Nissen and Bullemer (1987), we sought to investigate the properties of procedural and declarative learning. On each trial in this task, a light appears in one of four positions arranged horizontally. Subjects have four response buttons and are instructed to press the button directly below the position of the light. In a repeating-sequence condition, these four lights appear in a repeating 10-trial sequence. Reaction time (RT) shows a rapid decrease with
training on the repeating sequence. We know that this reduction primarily reflects acquisition of the specific repeating sequence and not learning of more general task characteristics because (a) if subjects are subsequently switched to a random sequence, RT increases substantially, and (b) training on a random sequence yields minimal reduction in RT. Nissen and Bullemer (1987) found that the decrease in RT during four blocks of 100 trials each on the repeating sequence was 119 ms, whereas the decrease during the same amount of training on a random sequence was only 19 ms.

This task, using a chronometric measure, provides an implicit assessment of learning, without the requirement for conscious remembering of the sequence. It is a performance measure rather than an introspective measure of learning. We believe that performance on the task reflects procedural learning in that it taps the implicit knowledge of how to do something and does not require explicit factual knowledge (though explicit knowledge of the sequence, if acquired, may facilitate performance).

Previous research showed that amnesic patients with Korsakoff's syndrome (Nissen & Bullemer, 1987), most patients with memory disorders resulting from Alzheimer's disease (Knopman & Nissen, 1987), and healthy young adults injected with scopolamine, an anticholinergic drug that produces a temporary amnesia (Nissen, Knopman, & Schacter, 1987), demonstrated by their performance that they learned the new associations embedded in the repeating sequence. Additionally, amnesic Korsakoff patients showed normal retention of the sequence across an interval of 1 week (Nissen, Willingham, & Hartman, in press). This preserved learning and retention of the sequence occurred despite the patients' reports that they did not notice a repeating sequence.

In the set of experiments reported here, we used this task to investigate the characteristics of procedural learning in normal subjects. The first and second experiments addressed the question of whether procedural knowledge may be acquired independently of declarative knowledge or whether one type of knowledge must be acquired first. In the final experiment we asked what is learned when procedural knowledge is acquired; specifically, it addressed whether procedural learning on this task was primarily perceptual or primarily motor and assessed the specificity of procedural knowledge.

Experiment 1

According to Anderson's ACT* theory, "all knowledge ... starts out in declarative form and must be converted to procedural (production) form" (Anderson, 1987, p. 196). Our results from amnesic patients seem to violate this principle, inasmuch as the patients learned the sequence procedurally in the absence of explicit declarative knowledge of the sequence. This conclusion, however, rests upon acceptance of patients' verbal reports that they did not notice a sequence. Such reports might be particularly suspect coming from patients with Korsakoff's syndrome, one symptom of which is often confabulation, and Alzheimer's disease, which can be accompanied by difficulties in comprehension. It is important to determine whether in normal subjects procedural learning can occur in the absence of explicit declarative knowledge of the sequence. That is, can the procedural learning system be isolated in normal subjects, as well as in patients with organic amnesia?

We trained subjects on the reaction time task and then asked them whether they had noticed a repeating sequence. In order to obtain a more quantitative assessment of their explicit declarative knowledge of the sequence, we then administered a test that required subjects to generate the sequence in a cued recall procedure. This "generate" task was formally similar to the reaction time task; the same stimuli and responses were used in both. However, instead of pressing the button directly below the stimulus that appeared, subjects were instructed to press the button corresponding to where they thought the next stimulus would appear. Initial accuracy and rate of improvement of accuracy on this task were the variables of interest.

The primary goal of this experiment was to determine whether there were subjects who demonstrated procedural learning of the sequence in the absence of explicit declarative knowledge of the sequence. A second goal stemmed from our anticipation that some subjects would learn the sequence declaratively as well as procedurally. We sought to determine how declarative knowledge of the sequence manifests itself in the performance of the reaction time task by those subjects.

Method

Stimuli and apparatus. Stimuli were generated on a video monitor controlled by a microcomputer. The single stimulus appearing on each trial was an asterisk 0.35 cm in diameter that was centered at one of four locations, all of them 5.5 cm from the bottom of the monitor screen but separated horizontally by 2.9 cm. At a viewing distance of approximately 58 cm, the four locations were separated by 2.87° of visual angle. The stimuli were clearly suprathreshold in luminance, and the four locations were easily discriminable. The stimulus on each trial remained present until the correct response was made, at which time that stimulus was extinguished and the next one appeared after a delay of 500 ms.

Responses were made by pressing one of four keys arranged in a row on a board placed below and in front of the video monitor. The keys had center-to-center distances of 4 cm.

Procedure. Subjects were seated facing the video monitor and response board in a moderately lit room. They rested the middle and index fingers of both hands on the four keys. Subjects initiated each block of 100 trials with a key press. Successive blocks of trials were separated by a short rest period of 1.5 to 2 min.

Subjects first completed four blocks of trials in a reaction time task. They were instructed to press the key that was below the location in which the asterisk appeared and to respond as fast as possible without making errors. No feedback was given regarding response latency or accuracy. In all blocks of trials, the location of the stimulus followed a particular 10-trial sequence. Designating the four locations as A, B, C, and D from left to right, the sequence was as follows: D-B-C-A-C-B-D-C-B-A. Each block of trials comprised 10 repetitions of this 10-trial sequence, but the end of one 10-trial sequence and the beginning of the next was not marked in any way. Thus, in the absence of knowledge of the sequence itself, each block would seem to be a continuous series of 100 trials. The presence or absence of a repeating sequence was not mentioned to subjects before or during this task. Subjects were told they were performing a reaction time task and were not given any other information about the purpose of the experiment.
After the fourth block of trials, subjects were asked a series of three questions about the pattern in the task. First, they were asked a general question, such as, "Did you notice anything about the task?" Next they were asked a more specific question, "Did you ever notice any pattern or repeating sequence?" If subjects indicated they had, they were asked to indicate what it was. All subjects then completed two blocks of trials in the generate task. They were told that the procedure would change in such a way that instead of pressing the key below the stimulus, they were to press the key corresponding to where they thought the next stimulus would appear. Subjects were also told that we were more concerned with which response was made than with the speed of responses in this task. Subjects were not told that there would be a pattern, and if they asked if there was a pattern or what purpose of the task was, they were simply told to do their best. The location of the asterisk followed the same 100-trial sequence used previously. The stimulus remained present until the subject made the correct response, which in this task was to push the key corresponding the next location in the sequence.

A group of control subjects performed the generate task only, without having any prior training on the reaction time task.

Subjects. All subjects were members of the University of Minnesota community and received either $5 or research credit in an introductory psychology course for participating. The group of 60 experimental subjects included 29 women and 31 men with a mean age of 20.4 years. The control group who performed only the generate task comprised 15 individuals (6 women and 9 men) with a mean age of 20.7 years.

Results

Verbal reports. The verbal reports obtained after the fourth block of the reaction time task provided two types of information: whether subjects said that they had noticed a sequence and, for those who said that they had, their accuracy in indicating what the sequence was. Subjects typically attempted to designate the sequence by pointing to locations rather than by describing the sequence verbally. A total of 12 subjects said that they had not noticed a sequence. Of the remaining 48 who said that they had, 12 were able to indicate what the entire sequence was, 29 correctly specified at least 4 but fewer than 10 consecutive positions, and 7 failed to identify more than 3 consecutive positions correctly. Assuming that subjects realized following four blocks of training that the same position was never used on successive trials, the probability of correctly guessing a sequence of three positions was .28. Thus, we suspected that the 7 subjects who were unable to identify more than three consecutive positions might have said that they noticed a sequence because of the demand characteristics of the situation. When we compared this group with the 12 subjects who said they had not noticed a sequence, we found no significant effect of group or interactions with group in analyses of variance of results from the reaction time task or the generate task. The reduction in RT from Block 1 to Block 4 was 90 ms for subjects who said they had noticed a sequence and 7 subjects who said they had but were unable to indicate more than three consecutive positions in it. Those with some explicit knowledge (n = 29) included subjects who said that they had noticed a sequence and successfully identified between four and nine consecutive positions. Those with full explicit knowledge (n = 12) said they had noticed a sequence and specified all of it.

Reaction time task. For each subject, the median reaction time (RT) of correct responses in each set of 10 trials within a block was determined, and the mean of these 10 medians was computed for each block. Figure 1 shows the group means and standard errors of these values. All three groups in the current experiment showed a substantial reduction in RT with training. The decrease in RT from Block 1 to Block 4 was 94 ms for the no-knowledge group, 118 ms for the some-knowledge group, and 205 ms for the full-knowledge group. These data were examined by a two-way analysis of variance (ANOVA), with group as a between-subjects factor and block as a within-subjects factor. The analysis identified significant main effects of group, F(2, 57) = 6.32, MS_E = 16,497.63, p < .005, and block, F(3, 171) = 180.94, MS_E = 1,029.11, p < .001, and a significant interaction between group and block, F(6, 171) = 10.18, MS_E = 1,029.11, p < .001. Paired comparisons between groups in each block indicated that the no-knowledge and some-knowledge groups did not differ significantly in any block. (A criterion significance level of .05 was used in all analyses reported in this article.) The full-knowledge group responded significantly faster than the some-knowledge group in Blocks 3 and 4, t(39) = 2.61, p < .01, and t(39) = 4.0, p < .001, respectively, but not in Blocks 1 or 2. The full-knowledge group responded significantly faster than the no-knowledge group in Blocks 2, 3, and 4, t(29) > 2.46, p < .01 in all cases, but not in Block 1.

The accuracy data from these three groups are shown in Table 1. Accuracy was high throughout, and it increased with practice. Although the no-knowledge group tended to have

![Figure 1. Mean of median reaction time in milliseconds in Blocks 1-4 of Experiment 1. (Bars represent standard errors. Open circles = no explicit knowledge; filled circles = some explicit knowledge; open triangles = full explicit knowledge.)](image-url)
the lowest level of accuracy and the full-knowledge group the highest, this effect was not significant, possibly due to scale attenuation. An analysis of these data revealed a significant effect of block, \( F(3, 171) = 5.25, MS_e = 5.37, p < .005 \), but neither the main effect of group nor the interaction between group and block was significant.

An analysis was performed on the improvement in reaction time that each group showed from Block 1 to Block 4. Reaction time difference scores were calculated by subtracting Block 4 RTs from Block 1 RTs. Also included in this analysis are data from a random condition included in Experiment 1 of Nissen and Bullemer (1987). Subjects in this random condition performed the reaction time task under exactly the same conditions as subjects in the present experiment, the only difference being that in the random group the asterisks appeared in a random sequence—there was no repeating sequence. The difference scores were 19.1 ms for the random group, 94.2 ms for the no-knowledge group, 118.0 ms for the some-knowledge group, and 204.3 ms for the full-knowledge group.

A one-way ANOVA was conducted on the difference scores of reaction time, with group as the factor. The analysis showed a significant effect of group, \( F(3, 68) = 21.52, MS_e = 3,331.00, p < .001 \). Follow-up comparisons indicated that the random group improved less than each of the other groups, \( F(1, 68) > 12.46, p < .001 \) in all cases. As a precaution against the possibility that the variances of the groups were not homogeneous, nonparametric tests were conducted on the difference scores. A Kruskal-Wallis test was performed on the four groups, yielding \( H = 34.56 (p < .001) \). The two critical groups, random and no explicit knowledge, were tested in a separate Mann-Whitney analysis, yielding \( U = 213 (p < .001) \). These results agree qualitatively with the results from the parametric test. The lack of differences among groups may due to the fact that accuracy for all groups was near ceiling; the accuracy for all groups on all blocks was 93% or greater.

**Anticipatory responses.** The exceptionally fast responses of the full-knowledge group in Blocks 3 and 4 suggested that these subjects were anticipating the onset of the stimulus and were initiating a response before its appearance. The possibility of anticipatory responses was enhanced by the use of a constant response-to-stimulus interval.

If a choice reaction time requires detection of the stimulus, choice of response, and execution of the response, anticipations would reflect trials on which the subject did not need either to choose the response or, because of the constant response-to-stimulus interval, to detect the stimulus. From experience, the subject could predict the next asterisk position. Simple visual reaction times average between 185 and 200 ms, with a variance typically under 10 ms (Luce, 1986). In addition, the response execution portion of simple reaction time has been estimated to be at least 100 ms (Wood, 1977). Therefore, anticipatory responses were defined as those with latencies of less than 100 ms because such responses are much faster than what one would expect to obtain in a simple reaction time task and are unlikely to reflect any process other than response execution.

Table 2 shows the number of such responses by each group in each block of trials. Anticipatory responses were initially infrequent, occurring on approximately 1% of trials in Block 1, and their number remained relatively low for the no-knowledge and some-knowledge groups. In contrast, anticipatory responses increased dramatically with practice in the full-knowledge group; these subjects made anticipatory responses on nearly half of the trials in Block 4. An analysis of variance of these data indicated significant main effects of group, \( F(2, 57) = 13.0, MS_e = 562.56, p < .001 \), and block, \( F(3, 171) = 43.63, MS_e = 95.86, p < .001 \), and a significant interaction of group and block, \( F(6, 171) = 16.95, MS_e = 95.86, p < .001 \).

Reaction time data were reanalyzed after the exclusion of all trials with anticipatory responses. The results, shown in Figure 2, show similar patterns of improvement with training for all three groups. An analysis of variance of these data revealed a significant effect of block, \( F(3, 171) = 120.89, MS_e = 751.35, p < .001 \). However, the main effect of group, \( F(2, 57) = 1.35, MS_e = 11,837.05 \), was not significant. Thus, when anticipatory responses were excluded, subjects with no, some, and full explicit knowledge of the sequence demonstrated no significant differences in rate of learning.

**Generate task.** In the generate task, the repeating sequence from the reaction time task was used again, but instead of pushing the button below the asterisk, subjects were required to predict the next asterisk position. Thus subjects could demonstrate explicit knowledge of the sequence acquired during the preceding reaction time task, and because there was feedback about accuracy in this task (the next asterisk appeared only when its correct position was predicted), sub-

<table>
<thead>
<tr>
<th>Group</th>
<th>Mean Number of Anticipatory Responses in Experiment 1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Trial block</td>
</tr>
<tr>
<td></td>
<td>2</td>
</tr>
<tr>
<td>No knowledge</td>
<td>0.3</td>
</tr>
<tr>
<td>Some knowledge</td>
<td>0.6</td>
</tr>
<tr>
<td>Full knowledge</td>
<td>1.2</td>
</tr>
</tbody>
</table>
DEVELOPMENT OF PROCEDURAL KNOWLEDGE

A subgroup of subjects in this experiment demonstrated substantial procedural learning of the sequence in the absence of explicit declarative knowledge. Their response times decreased nearly 100 ms with training on the repeating sequence, yet their ability to generate and to learn to generate the sequence did not differ from subjects who had no prior experience with the sequence. These findings thus confirm the earlier results from patients with organic amnesia on this task, and they show that procedural and declarative learning can be dissociated in individuals with normal memory.

The results are also relevant to the claim that procedural knowledge is necessarily derived from preexisting declarative representations of the same information (Anderson, 1987). Although the two-memory systems may interact in that way in the acquisition of proof skills in geometry or of text-editing...
skill, it appears from this experiment that procedural learning is capable of proceeding in the absence of prior declarative knowledge. Within the framework of ACT*, it could be argued that successive sets of a few items were held temporarily within working memory and that proceduralization developed from those representations despite the fact that they did not yield a lasting explicit declarative representation of the sequence.

The strength of the findings from this experiment rests in part on the fact that the pair of tasks used to tap these two memory systems have a great deal of formal similarity. The reaction time and generate tasks employ the same stimuli and responses, and neither requires a verbal response. One might claim, however, that having to generate a sequence is more difficult than pressing buttons directly below lights. It should be noted in this regard that the comparison made in this experiment was not between absolute performance on the reaction time task and absolute performance on the generate task. Instead, the comparison was between evidence of learning in the reaction time task (i.e., amount of improvement in performance due to sequential learning) and evidence of learning on the generate task (i.e., accuracy on the generate task). There seems to be no a priori basis for claiming that it is more difficult to demonstrate learning in one of these ways than the other. A similar concern relates to the possibility that subjects whom we considered to have no explicit knowledge of the sequence did in fact have such knowledge but that the generate task was not sensitive enough to detect it. However, the generate task allows assessment of cued recall on the basis of the first set of 10 trials and of savings in learning on the basis of the improvement across 10-trial sets. It is worth noting that Nelson (1978) compared recall, recognition, and savings in learning and found savings in learning to be the most sensitive test of the three. Nevertheless, our subjects with no explicit knowledge demonstrated no savings in learning on the generate task as compared with the control group. One might argue that subjects would have shown transfer on a recognition test, even though it is not as sensitive, but performance on a recognition test would depend critically on the similarity of the target and distractors.

This experiment is also relevant to the question of how procedural and declarative knowledge interact in task performance. One possibility is that their effects are effectively superimposed. Explicit declarative knowledge of the sequence, which supports performance in the generate task, also allows the generation of responses in anticipation of the identification of the stimulus in the reaction time task. The occurrence of anticipatory responses in subjects with full explicit knowledge of the sequence makes it appear that their rate of learning exceeds that of other subjects. However, when the effect of their explicit knowledge is removed or reduced by excluding anticipatory responses, the rate of procedural learning appears to be the same for subjects with and without explicit knowledge.

Experiment 2

It is a common belief that the acquisition of skills takes longer than the learning of facts. On one hand, this claim seems consistent with our personal experiences with the relative difficulty of learning someone's name versus learning to play chess or racquetball well. On the other hand, comparing the learning of a single declarative fact with the learning of a complex skill comprising many components seems inappropriate. Comparing the time course of procedural and declarative learning requires parallel methods for assessing representations of the same experience in both systems, which the pair of tasks used here are intended to provide.

Our earlier studies indicated that evidence for procedural learning of the repeating sequence emerges by the end of the first block of 100 trials (Nissen & Bullemer, 1987). That is the time at which results from subjects receiving the repeating sequence and those receiving a random sequence diverge significantly. In contrast, subjects have reported not being aware of the repetition until the third or fourth block of trials. And, of course, results from Experiment 1 indicate that some subjects do not acquire explicit knowledge of the sequence within four blocks.

These data rule out the possibility that procedural knowledge is always preceded by declarative knowledge. Experiment 1 does not, however, determine whether procedural knowledge precedes declarative knowledge or whether the two develop in parallel in this task. Experiment 1 provided an assessment of both knowledge systems after four blocks of training. Experiment 2 was designed to show how knowledge develops during each block of training. The goal of this experiment was to investigate the temporal relation between the emergence of procedural and declarative knowledge of the sequence with training on the reaction time task.

We gave different groups of subjects from zero to six blocks of the reaction time task, obtained verbal reports of awareness of the sequence, and then determined their performance on the generate task. This design allowed us to assess, through the reduction in subjects' response times, the amount of procedural learning that occurred with each level of training. In addition, by interrupting training after different numbers of blocks and by determining subjects' performance on the generate task, we could also track the emergence of explicit declarative knowledge of the sequence with varying amounts of training on the reaction time task. In this experiment we thus examined incidental learning in both memory systems. It is also worth noting that instead of relying on self-selection to obtain different levels of explicit knowledge, we manipulated level of explicit knowledge by varying the amount of experience subjects had with the pattern.

Method

Procedure. The stimuli, apparatus, and most aspects of the procedure were the same as in Experiment 1. Subjects completed a series of blocks of the reaction time task with the repeating sequence of stimuli and responses. Following the last block of the reaction time task, they were asked if they had noticed a repeating sequence. Finally, they completed two blocks of the generate task.

Six groups of subjects differed in the number of blocks of trials they were given in the reaction time task; the groups were given from one to six blocks in that task before being asked about the sequence and being given the generate task. Data from the group of control subjects from Experiment 1, who performed the generate task with
no prior training on the reaction time task, were also included in analyses.

Subjects. The 72 subjects included 44 women and 28 men between the ages of 18 and 30 years. The mean age was 21.3 years. Subjects were members of the University of Minnesota community and received either research credit or $5 for participating. They were assigned randomly to Groups 1-6. Each group included 12 subjects. The control group, described in Experiment 1, included 15 subjects. Data from 4 of the subjects in Group 4 (the 4 individuals in that group who reported they had not noticed a sequence) were included in the analyses of Experiment 1.

Results

Verbal reports. The number of subjects in each group who reported after the reaction time task that they had noticed a sequence was five in Group 1, eight in Group 2, ten in Group 3, eight in Group 4, eleven in Group 5, and twelve in Group 6. Thus, fewer than half of the subjects reported that they had noticed it after only one block of trials, and all reported that they had noticed it after six blocks. Because subjects were not asked to describe the sequence, it is possible that some of the affirmative responses in each group reflected demand characteristics, as discussed in connection with Experiment 1.

Reaction time task. Means of median reaction times appear for each group in Figure 4. One-way analyses of variance conducted on results from each of Blocks 1–5, with group as a between-subjects factor, indicated no significant effect of group in any block (p > .20 in all cases). However, an analysis of RT in the last block performed by each group revealed a significant effect of group, F(5, 66) = 4.82, MSQ = 8,190.47, p < .01, reflecting a progressive reduction in RT with additional blocks of training.

The mean percent correct was computed for each group in each block. These values ranged from 94.2 to 97.7. There were no significant differences among groups in any block, including the last block performed by each group.

Generate task. Percentages of correct responses on each set of 10 trials in the generate task are shown in Figure 5.

Results from the control group, which had no prior training, are included. Results from the first block show, first, that the performance of Group 1 and the control group is very similar. It is also apparent that Group 6 maintains the highest level of accuracy. The remaining groups (2-5) demonstrate similar, intermediate levels of performance. The accuracy of all groups increases with practice; this increase appears to be greater for Group 1 and the control group than for the others. An analysis of variance of these data identified significant main effects of group, F(6, 80) = 2.43, MSQ = 3,582.95, p < .05, and set, F(9, 720) = 27.44, MSQ = 176.20, p < .001, and an interaction of group and set, F(54, 720) = 1.56, MSQ = 176.20, p < .01.

Results from the second generate block showed a gradual increase in accuracy across sets of trials, with Group 6 responding most accurately. Analysis of these data revealed significant main effects of group, F(6, 80) = 2.56, MSQ = 2,244.86, p < .05, and set, F(9, 720) = 17.44, MSQ = 133.68, p < .001. The interaction of group and set was not significant.

Criterion analysis. The results from this experiment are consonant with parallel development of procedural and declarative knowledge, but it is still possible that procedural knowledge always precedes declarative. It may be that as the number of blocks of trials is increased, a greater percentage of subjects have acquired sufficient procedural knowledge to
allow them to go on to gain declarative knowledge. Thus in Group 1 there may be only a few subjects who have gained enough procedural knowledge to allow them to acquire some declarative knowledge. In Group 2, subjects had more training on the procedural task, and so a greater number would have gained enough procedural knowledge to go on to gain some declarative knowledge. When the data are combined across subjects, it would appear that procedural and declarative knowledge are acquired in parallel.

To further investigate the temporal order of learning, one must assess the development of procedural and declarative learning within each subject. To that end, a learning criterion was set for each group for the reaction time task and the generate task. If procedural learning always precedes declarative learning, no subject should reach criterion on the generate task without showing learning on the procedural task. If procedural and declarative learning proceed in parallel, subjects may show either type of learning first.

The idea in setting the criterion was that subjects given some opportunity to learn the sequence should perform better than subjects who did not have that opportunity. Therefore, the performance of subjects in the present experiment was compared with the performance of untrained subjects, with untrained subject performance being considered a baseline. It was expected that those subjects who had learned would show better performance than would the untrained subjects, and a criterion could be set on the basis of the standard deviation of the untrained subjects' performance. If a subject in the present experiment is at the far end of the distribution of scores for untrained subjects, it is reasonable to conclude that this subject is unlikely to belong to the population of subjects who had not learned the sequence.

For the reaction time task, baseline performance was taken from the random condition of Experiment 1 from Nissen and Bullemer (1987). The baseline was defined as the difference between the reaction time for the first block and the last block of trials. For example, for Group 2 in this experiment, the baseline was the mean RT difference between the first and second blocks in the random condition of Nissen and Bullemer (1987)—that is, 5 ms. The standard deviation of those difference scores was 27 ms. Therefore, a lenient criterion would be the mean difference from the random condition, plus one standard deviation, or 32 ms. By this criterion, any subject whose Block 1–Block 2 difference score was 32 ms or more was classified as having learned procedurally. A more stringent criterion would be to set the criterion as two standard deviations beyond the mean, or 59 ms. A still stricter criterion would be to classify as learners only those subjects whose difference scores were three standard deviations beyond the mean (86 ms). All three criteria were evaluated.

Subjects in the random condition of Experiment 1 from Nissen and Bullemer (1987) improved 5 ms from Block 1 to Block 2, but they improved 19 ms from Block 1 to Block 4. Therefore, a difference criterion had to be calculated for each group in the present experiment: The criteria for Group 2 were calculated from a baseline difference score of 5 ms, whereas the criteria for Group 4 were calculated from a baseline difference score of 19 ms.

For the generate task, the goal was to assess the declarative learning that had taken place during the reaction time task. Therefore, rather than using a difference score, the criterion was based on the absolute level of performance during the first 20 trials of the generate task. Only the first 20 trials were used so as to get a measure of declarative knowledge gained during the reaction time task. Baseline performance for the generate task was the first 20 trials taken from the control group of the present experiment, which performed the generate task without any prior training on the reaction time task. The criteria for the generate task were calculated in a manner analogous to that used for RT; the baseline mean was added to some number of standard deviations calculated from the baseline group.

The results are shown in Table 3. The pattern of results is the same at each criterion. As the subjects are given more experience with the reaction time task, more subjects learn about the pattern both procedurally and declaratively; thus the number of subjects showing both kinds of learning rises with increasing practice with the task, and the number of subjects showing learning in neither system drops with increasing practice. Furthermore, there are a number of subjects in every group who learn the pattern only procedurally or declaratively; also, there are a number of such subjects at every criterion level. It does not appear that one type of knowledge is a prerequisite for the other to develop in this task—either type of knowledge may be acquired in the absence of the other. It does appear that, in this task, procedural knowledge may be more likely to be acquired first; at every criterion level, of the subjects who acquired only one type of knowledge, a greater number acquired procedural rather than declarative knowledge.

It is apparent that declarative learning can take place in the absence of procedural learning, just as procedural learning can take place in the absence of declarative learning. This finding holds for each criterion of learning used here and therefore is not the result of a particular criterion. The data

Table 3
Number of Subjects Reaching Learning Criterion on Procedural and Declarative Tasks in Experiment 2

<table>
<thead>
<tr>
<th>Learning</th>
<th>Group</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Criterion 1: One SD beyond baseline mean</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neither</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Procedural only</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>Declarative only</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Both</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>Total</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>Criterion 2: Two SDs beyond baseline mean</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neither</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>Procedural only</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Declarative only</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Both</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Total</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>Criterion 3: Three SDs beyond baseline mean</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neither</td>
<td>9</td>
<td>8</td>
</tr>
<tr>
<td>Procedural only</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Declarative only</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Both</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Total</td>
<td>12</td>
<td>12</td>
</tr>
</tbody>
</table>
also replicate the results of Experiment 1: Some subjects demonstrate procedural knowledge in the absence of declarative knowledge.

Discussion

The results of this experiment indicate that increased training on the reaction time task produced greater procedural learning of the sequence and more explicit declarative knowledge of the sequence. It is difficult to compare the rate of incidental learning in the two memory systems quantitatively because of the problem of translating a decrement in response latency to an increment in response accuracy. Nevertheless, the gradual development of both procedural and declarative knowledge of the sequence is apparent in these results from the first block of trials on.

The criterion measure provides another assessment of the progression of procedural and declarative memory. Experiment 1 demonstrated that it is possible to acquire procedural information in the absence of declarative information. The present experiment confirms that finding and also demonstrates that, in the reaction time task, it is also possible to acquire declarative information in the absence of procedural information. It should be noted that this pattern of results may not extend to other tasks. It may be that for some tasks procedural typically precedes declarative or that declarative typically precedes procedural. The important point, however, is that the two types of knowledge are separable.

Experiment 3

Experiment 3 addressed two questions. First, what do amnesic patients and normal subjects learn when they learn the sequence procedurally? One possibility is that they learn where the next stimulus will appear and thus which spatial position to attend to next. Schneider, Dumais, and Shiffrin (1984) have suggested that the basis for learning effects in divided attention paradigms may be a training of attention. Such a training of attention may be the heart of learning the pattern in the reaction time task. Lewicki, Czyzewska, and Hoffman (1987) found that subjects learned information about a complex pattern of stimuli in a search task. It is likely that subjects were learning where to search for the target and that their improvement in reaction times was a result of shifting attention in anticipation of the target location despite their lack of awareness of the pattern of attention movement.

Alternatively, subjects learning the pattern in the reaction time task we have used may learn which button to press next—that is, they may learn a sequence of motor responses. Clarification of what is learned is of obvious relevance to understanding the cognitive and neural structures and processes underlying procedural learning.

In order to examine these alternative hypotheses, we tested normal subjects in three conditions, each a variation on the reaction time task used before. In all of the conditions, however, subjects responded to the color rather than the location of the stimulus. All subjects first completed a pretraining phase to allow them to learn the correct color-to-response mapping to a criterion. In the subsequent training task, one group of subjects participated in a perceptual sequence condition, in which there was a repeating sequence of stimulus positions. This condition thus allowed the possibility of a form of perceptual learning: learning where the next stimulus would appear. However, because the sequence of colors was random and because response selection was governed by stimulus color, there was no motor sequence to be learned in this condition. A second group participated in a response sequence condition, in which there was a repeating sequence of responses (because the colors followed a repeating sequence), but the sequence of stimulus positions was random. Thus, these subjects had the opportunity to learn a motor sequence. For a group of control subjects, both the sequence of stimulus positions and responses were random. We wanted to determine whether the perceptual sequence condition or the response sequence condition (or both) would demonstrate an advantage with training relative to the control condition.

The second question this experiment addressed involves the specificity of procedural learning. It has been claimed that the learning demonstrated by amnesic patients is hyperspecific (Schacter, 1985) in that it is "relatively inflexible, rigidly organized, and only narrowly accessible" (Glisky, Schacter, & Tulving, 1986, p. 325). Learning by amnesic patients thus appears more specific than learning by nonamnesic subjects. That appearance, however, may result from the contribution of the declarative memory system to the performance of normal subjects. Thus, we asked whether the procedural memory system in normal subjects, when considered in relative isolation from the declarative system, demonstrates task specificity.

Following training in either the perceptual sequence, response sequence, or control condition, subjects performed a transfer task that was identical to the task used in Experiments 1 and 2: Subjects pressed a button corresponding to the location of the stimulus, which was always the same color. The sequence of stimulus locations was the same as that used in the perceptual sequence condition, and the sequence of responses was the same as that in the response sequence condition. We sought to determine whether subjects who had demonstrated learning of the sequence during the earlier training task performed better on this transfer task than control subjects, who had not been trained on any sequence. Additionally, we compared the amount of transfer shown by subjects who acquired some explicit declarative knowledge of the sequence during the training task with the amount of transfer shown by subjects who did not. This comparison was intended to address the role of explicit declarative knowledge in the transfer of skill.

Method

Stimuli and apparatus. The stimulus appearing on each trial during the pretraining phase was a colored rectangle centered on a video monitor. The rectangle was 1 cm high and 1.5 cm wide, and its color was blue, green, orange, or violet. These colors were generated by an Apple microcomputer on an RGB monitor and were easily discriminable.

In the training task, the stimulus was an X 1.9 cm in diameter that was centered at one of four locations on the video screen, all of them 6.5 cm from the bottom of the monitor screen but separated horizon-
they were. Then they were asked whether they had noticed any training task. If they said they had, they were asked to indicate what whether they had noticed any patterns or sequences during the same pattern as in the training task. The fifth block employed a (and, thus, response locations) in the first four blocks followed the location in which the X appeared. The sequence of stimulus locations would be white instead of colored and that they should respond to was as follows: V-G-O-B-O-G-V-Q-G-B. With the color-t(}-response response sequence condition, the sequence of stimulus locations was Each block included 10 repetitions of this 100trial sequence. In the was random, but the location of the stimulus followed a particular allowed. In the perceptual sequence condition, the sequence of colors was random, except for the constraint that the same color was not used on successive trials. Subjects continued with this task until they completed three successive blocks with a mean RT of 600 ms or less and an accuracy of 90% correct or better. All subjects, however, had to complete a minimum of six blocks of trials. Subjects who did not reach the performance criterion within 20 blocks did not participate in the two subsequent phases of the experiment. Ten subjects failed to reach this criterion.

The second session began with two blocks of pretraining levels (i.e., with colored rectangles centered on the monitor). Then subjects completed four blocks on the training task: A colored X appeared in one of four locations on the monitor, and subjects again responded to the color of the X by using the same response assignments as in the pretraining phase.

Subjects were assigned arbitrarily to one of three conditions in the training task. In the control condition, both the sequence of colors and the sequence of stimulus locations was random, with the constraint that immediate repetitions of color or location were not allowed. In the perceptual sequence condition, the sequence of colors was random, but the location of the stimulus followed a particular 10-trial sequence (the same as that used in Experiments 1 and 2). Each block included 10 repetitions of this 10-trial sequence. In the response sequence condition, the sequence of stimulus locations was random, but the color of the stimulus followed a 10-trial sequence. Designating each color by the first letter of its name, the sequence was as follows: V-G-O-B-O-G-V-O-G-B. With the color-to-response mapping that was used, this sequence produced the same response sequence as in Experiments 1 and 2.

Upon completion of the training task, all subjects completed five blocks of trials in the transfer task. They were told that the stimuli would be white instead of colored and that they should respond to the location of the stimulus by pressing the key that was below the location in which the X appeared. The sequence of stimulus locations (and, thus, response locations) in the first four blocks followed the same pattern as in the training task. The fifth block employed a random sequence of locations.

After the transfer task was completed, subjects were asked, first, whether they had noticed any patterns or sequences during the training task. If they said they had, they were asked to indicate what they were. Then they were asked whether they had noticed any patterns or sequences during the transfer task, and, if so, to indicate what they were.

**Results**

**Verbal reports.** When subjects were asked whether they had noticed a repeating sequence during the training task, 2 of the 15 subjects in the control condition, in which there was no repeating sequence, said that they had. Of the 12 subjects in the perceptual sequence condition, 3 said they had noticed a sequence. However, those 3 described a sequence of colors instead of a sequence of stimulus positions, when in fact the sequence of colors presented in that condition was random.

Of the 61 subjects in the response sequence condition, 30 said that they had not noticed a sequence, and 17 said that they had but could not identify more than three consecutive items. Because of the possibility that demand characteristics might have induced subjects to say that they noticed a sequence and because the performance of these two subgroups on the training task did not differ—a finding similar to that of Experiment 1—we grouped these subjects together in subsequent analyses and considered that they had no explicit knowledge of the sequence. The remaining 14 subjects in the response sequence condition said that they had noticed a sequence and were able to identify at least four consecutive positions correctly (5 identified the entire sequence). This group was considered to have some explicit knowledge of the sequence.

**Training task.** As in Experiment 1, for each subject the median RT of correct responses in each set of 10 trials was determined, and the mean of these 10 medians was computed for each block. Group means and standard errors appear in Figure 6. Mean response times of the four groups in the first block of this task are very similar, but the effect of training differs among groups. Response times of the control group and the perceptual sequence group decreased only 26 ms and 5 ms, respectively, from the first block to the
Mean Percent Correct in Training Task of Experiment 3

Table 4
Mean Percent Correct in Training Task of Experiment 3

<table>
<thead>
<tr>
<th>Trial block</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>87.2</td>
<td>88.9</td>
<td>88.4</td>
<td>91.3</td>
</tr>
<tr>
<td>Perceptual sequence</td>
<td>91.1</td>
<td>92.4</td>
<td>93.0</td>
<td>93.2</td>
</tr>
<tr>
<td>Response sequence (no EK)</td>
<td>90.2</td>
<td>91.4</td>
<td>91.1</td>
<td>91.5</td>
</tr>
<tr>
<td>Response sequence (with EK)</td>
<td>92.8</td>
<td>94.6</td>
<td>93.9</td>
<td>95.0</td>
</tr>
</tbody>
</table>

Note: EK = explicit knowledge

Although accuracy was generally high, there were differences among groups, with the control group having the lowest accuracy and the response sequence group with explicit knowledge having the highest, and all groups showed improvement with training. These impressions were confirmed by an analysis of variance demonstrating significant main effects of group, $F(3, 84) = 5.76$, $MS_e = 53.70$, $p < .002$, and block, $F(3, 252) = 8.72$, $MS_e = 7.93$, $p < .001$. The interaction of group and block was not significant, $F(9, 252) = 0.69$, $MS_e = 7.93$.

Transfer task. Reaction time data from the transfer task, in which subjects responded according to stimulus location, appear in Figure 7. All groups showed a decrease in RT during the first four blocks, in which stimuli and responses occurred in a repeating sequence, and a substantial increase in RT in the fifth block, which used a random sequence. An analysis of these data indicated that the effect of block was significant, $F(4, 336) = 209.50$, $MS_e = 970.35$, $p < .001$, but the effect of group was not, $F(3, 84) = 1.16$, $MS_e = 10,833.79$. There was a significant interaction of group and block, $F(12, 336) = 5.79$, $MS_e = 970.35$, $p < .001$.

It would appear that this interaction relates to the fact that the response sequence group with explicit knowledge of the sequence responded faster than the other groups during the first four blocks but that their responses were slowest when the random sequence was used. Separate analyses of variance were performed on results from Blocks 1–4 and on results from Block 5. Data from Blocks 1–4 revealed a significant effect of block, $F(3, 252) = 49.1$, $p < .001$, but the effect of group just missed significance, $F(3, 84) = 2.58$, $p = .058$, and the interaction of group and block was not significant, $F(9, 252) = 0.98$. In contrast, the effect of group in Block 5 was significant, $F(3, 84) = 2.74$, $p < .05$. Paired comparisons revealed that in Block 5, the response sequence group with explicit knowledge responded significantly more slowly than the control condition, $t(27) = 2.48$, $p < .01$. Other groups did not differ.

Mean accuracy decreased in Block 5 relative to Blocks 1–4, as shown in Table 5. An analysis of the accuracy data indicated a significant effect of block, $F(4, 336) = 29.23$, $MS_e = 7.93$.

Figure 6. Mean of median reaction time in milliseconds in the training task of Experiment 3. (Filled circles = perceptual sequence condition; open circles = control condition; open triangles = response sequence condition with no explicit knowledge; filled triangles = response sequence condition with explicit knowledge.)
Table 5

Mean Percent Correct in Transfer Task of Experiment 3

<table>
<thead>
<tr>
<th>Group</th>
<th>Trial block</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2</td>
</tr>
<tr>
<td>Control</td>
<td>94.4</td>
</tr>
<tr>
<td>Perceptual sequence</td>
<td>95.4</td>
</tr>
<tr>
<td>Response sequence (no EK)</td>
<td>95.3</td>
</tr>
<tr>
<td>Response sequence (with EK)</td>
<td>97.3</td>
</tr>
</tbody>
</table>

Between group and block. Positive transfer of learning from the training task to the transfer task: Subjects who learned during the training phase did not respond significantly faster on the transfer task than those who did not learn. It is possible, however, that subjects who learned in the training phase did benefit, but once the transfer phase began, all subjects learned quickly, and so the advantage that transferred was hidden by the learning that occurred within the transfer task. In other words, subjects who learned during the training phase may have started out faster, but the other subjects may have quickly caught up. To test that possibility, the first 20 trials from Block 1 of the transfer task were analyzed separately. An analysis of these data showed a significant effect of group, $F(1, 80) = 7.66, MS_e = 9,778.83, p < .01$, and block, $F(3, 240) = 28.22, MS_e = 626.47, p < .001$, and a significant interaction of explicit knowledge and block, $F(3, 240) = 6.72, MS_e = 626.47, p < .001$. These results thus confirm those from Experiment 1. A parallel analysis of the accuracy data revealed no significant effects.

Discussion

Results from this experiment support neither of the two proposed hypotheses regarding the basis of learning the repeating sequence. First, subjects in the perceptual sequence condition demonstrated no advantage relative to the control group in the training phase. This finding indicates that they did not learn the perceptual sequence of spatial locations in which the stimuli appeared. If they had, that knowledge should have benefitted their performance because spatial expectancy has been shown to affect tasks ranging from the detection of a visual stimulus (Posner, Nissen, & Ogden, 1978) to form discrimination (Posner, Snyder, & Davidson, 1980).

In contrast to results from the perceptual sequence condition, subjects in the response sequence condition demonstrated robust sequential learning during the training task, disregarding whether they noticed a repeating sequence. This result argues for the second hypothesis: that the procedural learning shown by normal subjects (and, by extension, amnesic patients) basically involves learning a sequence of motor responses. The transfer data do not unambiguously support this hypothesis, however. If the learning were entirely encapsulated within the motor system, then the response sequence subjects should have responded faster than the other groups on the transfer task, which employed the same motor sequence. On the whole, the data do not argue for such transfer: The analysis of the first four blocks in the transfer task provides no evidence of transfer, nor does the analysis of the first block of trials alone. If the mean from the first 20 trials is taken as the dependent measure, there is some evidence for transfer, but it is difficult to be confident of these findings. First, they are based on data from relatively few trials. Second, response sequence subjects with explicit knowledge showed no transfer, whereas response sequence subjects with no explicit knowledge did show transfer, even though the latter group showed less learning during the training phase. In summary, though the transfer data are not unambiguous, we think that, overall, very little if any advantage is gained from the training task.

We conclude, by default, that what subjects learn in these tasks is central to and incorporates both perceptual and motor information. What is learned may be thought of as a series of condition–action statements mapping stimuli onto responses. In the training task, the condition of each rule corresponded to a color, and the action corresponded to a response location. The perceptual sequence condition did not induce sequential learning because the location of the stimulus was not represented in either the condition or the action of the production.
rule. In contrast, the response sequence condition provided the opportunity for sequential learning because the dimensions on which repetition occurred (stimulus color and response location) were represented in the production rules controlling performance. One model that could accommodate these findings is that of Hunt and Lansman (1986), who postulate in their production activation model that long-term memory consists of a network of interconnected productions. The repetitious use of a sequence of productions would strengthen the connections between them, and performance would be facilitated by the spread of activation between productions.

The results of this experiment point to some specificity of procedural knowledge. This specificity derives, we believe, from the fact that the learning is neither solely perceptual nor solely motor but represents instead the mapping rules governing performance. Performance in the transfer task of Experiment 3 demonstrated minimal benefit from prior sequential learning because the production rules required for performing in the transfer task differed from those developed during training. Knowledge embedded within one set of productions will not transfer to a task requiring different productions.

With regard to the hypothesis that explicit declarative knowledge may facilitate the transfer of skilled performance, we have the following correlational evidence. First, subjects who acquired explicit declarative knowledge of the sequence during training tended to have the fastest response times in the repeating sequence blocks of the transfer task, at least after the earliest part of the first block. Second, those subjects responded more slowly than the other groups when stimuli began appearing in a random sequence in the transfer task, suggesting greater interference when expectancies were violated. Finally, that same group of subjects (who had acquired declarative knowledge of the sequence) were more likely than the other groups to develop explicit knowledge of the sequence in the transfer task. Although all of these trends are consistent with the idea that explicit declarative knowledge is less task-specific and more domain-general than procedural knowledge, these particular findings do not have strong statistical support, and each is open to other interpretations. We believe, however, that this hypothesis is worthy of future investigation.

**General Discussion**

These experiments have demonstrated the usefulness of dissociating procedural and declarative memory systems in normal subjects in providing information about the workings of the systems. Results from the first experiment indicated that some normal subjects, like amnesic patients, learned a sequence of new associations procedurally in the absence of explicit declarative knowledge of the sequence. With the feedback that is inherent in the generate task, however, these subjects ultimately learned the sequence, unlike amnesic Korsakoff patients who have been given these tasks and whose performance on the generate task remained at chance.

An interesting question these findings raise for future research involves determining other differences in memory and cognition between normal subjects who acquire declarative knowledge of the sequence and those who do not. At the least, our findings suggest that variability among subjects in the rate of learning declaratively exceeds that of learning procedurally.

We were not surprised to find in Experiment 1 that declarative knowledge of the sequence can affect performance. Others have reported that the performance of nominally implicit memory tasks can benefit from explicit remembering (Nissen et al., 1987; Squire, Shimamura, & Graf, 1987). Our findings, as well as our interpretation of them, are reminiscent of those of Posner and Snyder (1975). In their task, subjects saw a priming item and then a pair of target items to which they responded. Posner and Snyder manipulated the probability that the prime would match the target, thereby manipulating the amount of attention subjects would allocate to the prime as a predictor of the target. In Posner and Snyder's cost–benefit analyses, they found that a low-validity prime that matches the subsequent target produces some facilitation in the processing of the target. They attributed this facilitation to a process of pathway activation that is outside of conscious awareness. If the prime is highly predictive of the target, the amount of facilitation is increased because of an active, conscious, attentional effect that is superimposed on the pathway activation. Similarly, we would argue that the benefit derived from procedural learning shown by our subjects can be likened to a pathway activation effect. If and when subjects develop explicit declarative knowledge of the sequence, they can use this knowledge to form attentional expectancies regarding the next item in the sequence. This effect, we believe, is superimposed on the effects of procedural learning and increases the overall improvement with training. Thus, declarative knowledge does not improve procedural knowledge per se but may facilitate the behavior used to measure procedural knowledge. As Mishkin and Appenzeller (1987, p. 89) put it, "behavior could be a blend of automatic responses to stimuli and actions guided by knowledge and expectations."

The results from Experiment 2 argue that although declarative knowledge may influence performance on a procedural task, procedural and declarative knowledge may be acquired separately; one need not have knowledge of one type in order to build the other type of knowledge. In our task, knowledge of the sequence seemed to be acquired in parallel by the two systems. Although the two systems are capable of learning independently, the order in which knowledge accumulates in the two and the contribution of each system to a task almost certainly depends on the particular task used. One suggestion for predicting the contributions of the two systems comes from Fitts's (1964) analysis of perceptual–motor skills. He predicted greater involvement of an initial verbal-mediation phase of motor skill learning when stimulus–response compatibility was low.

Experiment 3 sheds further light on the development of procedural learning and its representation. Procedural learning in the reaction time task seems to develop through the button-press responses, rather than through the guiding of attention to the stimulus locations on the screen. But char-

---

1 Because one position was not invariably followed by another particular position in the sequence, a simple series of connections between immediately successive items would be inadequate (Lashley, 1951). Either higher order units or connections of decreasing strength from one item to multiple successive items would be required.
acterizing the representation of what is learned in the reaction time task as a series of associated motor responses is unlikely to be accurate, because once the skill is obtained there is little evidence for transfer to a different task that requires the identical motor responses. When the stimuli that specify these motor responses are changed from those used in training, the knowledge representation that facilitated motor performance in the training task cannot be used. This finding is consistent with other work demonstrating the specificity of procedural learning (Catrambone & Holyoak, 1987; Kolers & Perkins, 1975; Masson, 1986) and with the observation that learning by amnesic patients is unusually specific (Gisly et al., 1986; Schacter, 1985).

In summary, we have attempted to dissociate empirically the effects of procedural and declarative learning, to assess the contribution of declarative knowledge to performance, and to investigate the temporal dependence between systems in knowledge acquisition. In this work, we have sought to link research on learning in normal subjects to parallel studies conducted on patients with organic amnesia in order to allow phenomena found in normal subjects to be considered within a neuropsychological framework and in order to further elucidate the nature of preserved learning in amnesia.

References


Luce, R. D. (1986). *Response times: Their role in inferring elementary mental organization*. New York: Oxford University Press.


Received October 6, 1987
Revision received May 1, 1989
Accepted May 2, 1989.