Age Differences in Implicit Learning of Higher Order Dependencies in Serial Patterns

James H. Howard, Jr.  
The Catholic University of America  
Darlene V. Howard  
Georgetown University

3 experiments examined serial pattern learning in younger and older adults. Unlike the usual repeating pattern, the sequences alternated between events from a repeating pattern and those determined randomly. The results indicated that no one was able to describe the regularity, but with practice every individual in all 3 age groups (including old old) became faster, more accurate, or both, on pattern trials than on random trials. Although this indicates that adults of all ages are able to learn second-order statistical dependencies in a sequence, age-related deficits were obtained in the magnitude of pattern learning. There were also age differences in what was learned, with only younger people revealing sensitivity to higher order statistical dependencies in the sequence. In addition, whereas younger people revealed evidence of their pattern learning in a subsequent conceptually driven production test, young-old and old-old people did not.

Everyday experience demonstrates how remarkably sensitive people are to regularities in their environment. In some cases people are aware of these regularities and can describe them (e.g., the sequence of events required to start a car), but in many other cases they seem oblivious to them and find them difficult or impossible to describe even though there may be other evidence that the pattern is known (e.g., the syntactic rules of language). In the latter case, implicit learning is said to have occurred, and several theorists have proposed that such implicit learning is fundamentally different from explicit learning (e.g., Reber, 1993; Seger, 1994). For example, Schacter and Tulving (1994) have argued that implicit and explicit learning are based on neurologically distinct memory systems, and Reber and his colleagues (Reber, Allen, & Reber, in press) proposed that the implicit system is more basic than the explicit from both the developmental and evolutionary points of view.

Implicit learning has been studied by using a variety of laboratory tasks. In each case, people are presented with material that contains some subtle regularity or structure, and the acquisition of this knowledge is measured (see Reber et al., in press, for a review). Implicit learning is said to occur when participants demonstrate knowledge of the regularity by the speed, accuracy, or both, of their responses, but they reveal little awareness of what they have learned. Exactly how implicit learning should be defined and assessed is a matter of much debate (e.g., Shanks & St. John, 1994). We conclude that learning in the present task is implicit if people show evidence of pattern sensitivity on the implicit measure described below but are unable to describe anything about the pattern that would account for this performance.

In the present article we ask whether such implicit learning remains intact in old age. So far, the answer is unclear, though there is some evidence that implicit learning declines little, if at all, in either advanced age or in amnesia, conditions that result in large and persistent deficits in most forms of explicit learning (e.g., Seger, 1994). Dissociations of this kind are particularly important for the theoretical distinctions that have been made between implicit and explicit memory systems.

Much of this evidence for relative age-related sparing of implicit, as opposed to explicit, learning comes from studies using a variation of a serial reaction time task (SRT; Nissen & Bullemer, 1987) in which people encounter four boxes arranged horizontally across a computer screen. They are told that when a stimulus (e.g., an asterisk or dot) appears in one of the boxes, they should push the key under that box as quickly as possible. This causes the stimulus to disappear and another to appear, at which point they should push the key under that box, and so on. Typically, what the participants are not told is that the stimulus follows a predetermined pattern and that the goal of the experimenter is to see whether people become sensitive to the pattern as revealed in their response speed and accuracy. Such sensitivity is usually tested by using either a between-subjects or a within-subjects design. In the between-subjects case, the rate at which performance improves over trials is contrasted between a group in which the stimuli do follow a predictable pattern versus a group in which the stimuli occur randomly; faster improvement for the pattern than for the random group.

The data from Experiment 2a were included in a poster presentation at the Cognitive Aging Conference in Atlanta, Georgia, in April 1996. This research was supported by Grant R37 AG02751 from the National Institute on Aging.

We thank Amy Pettit, Sarah Oldham, Wilbert Watts, Kim Gallagher, and Brian Cho for their assistance in data collection. We are grateful to Peter Frensch and Michael Stadler for their helpful comments on a version of this article.

Correspondence concerning this article should be addressed to James H. Howard, Jr., Department of Psychology, The Catholic University of America, Washington, DC 20064, or to Darlene V. Howard, Department of Psychology, Georgetown University, Washington, DC 20057. Electronic mail may be sent via Internet to James H. Howard, Jr., at howard@cua.edu, or to Darlene V. Howard at howarvd@georgetown.edu.
indicates that pattern learning has occurred. In the within-subjects case, people respond to the patterned sequence for several blocks of trials and then (without being informed of it) are switched to a random sequence; to the extent that response speed, accuracy, or both are disrupted by removing the pattern, this provides evidence that the pattern has been learned.

In earlier research, using both the within-subject and the between-subjects procedures described above, we have found that younger and older adults show equal amounts of pattern learning on the implicit measure, for repeating sequences of both 10 and 16 items (D. V. Howard & J. H. Howard, 1989, 1992). French and Miner (1994) and Cherry and Stadler (1995) have also reported age similarity on the implicit measure of learning in the SRT task, though French and Miner showed that age deficits did emerge under dual-task conditions, and Cherry and Stadler found that older adults of low educational level (but not those of high levels) were deficient on the serial learning task. In contrast to these several studies showing age constancy, Harrington and Haaland (1992) found some evidence of age differences on the implicit learning measure for highly educated older people working under single-task conditions. The reasons for this difference are unclear but may be related to the fact that the Harrington and Haaland study involved complex hand-sequence motor learning, whereas the others required only button presses to simple visual events; we have shown that motor responding is not necessary for implicit pattern learning in the conventional task (J. H. Howard, Mutter, & Howard, 1992). So it is possible that it is the motor-learning component of the Harrington and Haaland task that is leading to the age deficits they observed.

In the experiments reported here we introduce a modification of the SRT task. Our present version differs from the original described above primarily in the nature of the sequence structure. In the original version, every stimulus follows the predetermined pattern, whereas in the present version (which we call the alternating SRT task), alternate stimuli follow a predetermined pattern, whereas the remaining stimuli are selected randomly, with each of the four events sampled uniformly. For example, if a given participant had been assigned the pattern 1432 (where 1 stands for the left-most position, and 4 for the right-most position), then that participant would encounter the following series, where r stands for a random selection of one of the four possible positions: 1r4r3r2r1r4r3r2r1r4, and so on.

Our alternating SRT task relates to everyday experience in that patterns seldom occur in isolation but frequently occur in a noisy background. In addition, this alternating SRT task has several advantages over the original. First, people are much less likely to become consciously aware of any regularity—in the sense of being able to describe it—and so the pattern learning is more clearly implicit. In the original SRT task, many people, and more younger than older ones, reported becoming consciously aware of the presence of a repeating pattern by the end of the training, and they were often able to describe part or all of it (D. V. Howard & J. H. Howard, 1992). Thus, in the original task both explicit and implicit learning occurred. In contrast, in pilot studies with younger people, we found that none of the more than 10 people we tested on the alternating task was able to report anything at all about the nature of the pattern. This was true even of research assistants who knew the results of earlier SRT tasks.

Second, alternating pattern with random trials makes it easier to examine the course of pattern learning in individual participants. This is the case because accuracy and response time can be calculated separately for pattern versus random trials for each testing unit (e.g., block or session). Hence, it is possible to determine precisely when pattern and random trials diverge for each individual. In the original, within-subjects version of the task (D. V. Howard & J. H. Howard, 1989, 1992; Nissen & Bullemer, 1987), in contrast, typically 400 or more pattern-only trials were presented before any random trial occurred. During these early trials, it is impossible to know whether the improvement in performance was due to pattern acquisition or to a more general response learning. These alternatives can only be distinguished on the subsequent random block, which reveals the extent to which removing the pattern disrupts performance. The between-subjects variation of the traditional SRT task suffers from similar limitations. In this case, although it is possible to determine the first block at which the pattern and random groups diverge, this between-subjects comparison provides no information on when individual participants acquired the pattern. The relatively large between-subjects variability in overall response time further complicates these comparisons.

A third advantage offered by the alternating pattern structure introduced in the present experiments is that it can yield additional insights into what people are actually learning when they do reveal pattern sensitivity. Recent research (e.g., Stadler, 1992) has shown that learning in the SRT task depends on the degree of statistical structure present in the event sequence—greater structure leads to more learning. In the conventional task the repeating short pattern introduces substantial structure compared with a random sequence in which the events are selected randomly. This structure exists at several levels. For example, zero-order structure exists when some pattern events occur more frequently than others (e.g., 423124321 ...). Here, people could learn to anticipate the more frequent events (e.g., 2 and 3). First-order structure exists when pairs of events occur with different frequency (e.g., the pair 32 is more frequent than other pairs in the above-indicated example). Triplets of events can also vary in relative frequency, indicating second-order structure and so on. In the conventional SRT task, including those specifically designed to investigate the role of statistical structure (e.g., Stadler, 1992), many levels of structure exist making it difficult to investigate learning at a particular level.

In contrast, the alternating sequence used in the present experiments required individuals to learn at least second-order pattern structure (i.e., runs of three events or triplets) because the sequences were unstructured below this level. To explain, in the present experiments each participant in an age group encountered a different pattern, but for everyone, alternate events were determined by a pattern that was a permutation of the Positions 1, 2, 3, and 4. On the remaining random trials, the positions from 1 to 4 were chosen randomly and uniformly. Thus, if people reveal pattern sensitivity in the form of better performance on pattern than on random trials, this cannot reflect simple learning of the frequency of individual items (zero-order learning) because the four events occur equally often on both random and...
pattern trials. Nor can better performance on pattern than random trials be due to learning the relative frequencies of individual pairs of events (first-order learning); these too occur equally often for all possible pairs of positions on both pattern and random trials.

If performance is better for pattern than for random trials, then the simplest regularity that people could be learning would be that some triplets (sequences of three or second-order learning) are more likely than others because this is the lowest level at which random and pattern trials differ. This becomes clear by examining a sample sequence for a participant whose pattern is 1432. This person would encounter sequences of the following form:

\[ \ldots 1r4r3r21r4r3r21r4 \ldots \]

Thus, for this person, sequences such as 1r4, 4r3, 3r2, and 2r1 (where r refers to any position) are frequent because they are the only triplets that can occur when the third item is from the pattern as well as when it is random. In contrast, sequences such as 4r1 or 4r4 will be infrequent because they can only occur when the third item in the triplet is random.

Further, if a given person reveals pattern sensitivity (in the form of better performance on pattern than random trials), then one can distinguish between two possibilities of what is learned. First, it is possible that participants simply learn which triplets of events are relatively likely to occur. We term this triplet-only learning. Second, it is possible that individuals learn more than the relative frequency of event triplets. Although there are a number of possibilities here, one is that they learn something about the higher level alternating structure of the sequence. In other words, they may learn that the sequence is structured on only alternate trials. Another possibility is that people learn that runs of four or greater are more or less likely to occur. This is possible because in the present procedure we imposed structure at yet higher orders (fourth, sixth, etc.). We term these alternatives collectively as higher-order learning because the present experiments were not designed to distinguish among these latter possibilities.

We can distinguish between triplet-only and higher order learning here by breaking each participant’s performance on random items into two classes: those random items that occurred as the last item in a low-frequency triplet versus those random items that occurred as the last item in a high-frequency triplet. This enables us to examine the high-frequency random triplets separately. If at a given point in training, a person has acquired only knowledge about triplets, then performance should be just as good on these high-frequency random triplet items as it is on pattern items; in both cases the person is responding to the third item of a high-frequency sequence. For such people, it would only be low-frequency random items that would lead to poorer performance. In contrast, if the person has acquired some higher order knowledge in addition to (or instead of) triplets, then performance to pattern events should be better than that to random events that are the third items of high-frequency triplets.

Using this alternating SRT task, then, in the following experiments we addressed three primary questions. First, are both younger and older people sensitive to the regularities in the sequences? If so, then over sessions, performance on random versus pattern trials should diverge for all age groups.

Second, are there age differences in sensitivity to the patterns?

Any such age differences would be revealed by Age x Trial type (random vs. pattern) interactions.

Third, what are people learning when they reveal pattern sensitivity, and are people of different ages learning the same thing? That is, are people only learning which sequences of triplets are relatively frequent, or are they learning higher order information? Triplet-only learning would be indicated if performance on the third event of high-frequency triplets is equally good on pattern and random trials. This would occur because in triplet-only learning people simply learn to anticipate high-frequency triplets. On the other hand, higher order learning would be indicated if performance is better for high-frequency triplets ending on pattern trials than for those ending on random trials. Such a difference cannot be explained by triplet learning alone. Furthermore, if younger people show evidence of such higher order learning, it will be of interest to see whether older people do as well.

Experiment 1

Method

Participants. There were 12 participants, 6 young and 6 old, each of whom participated in six 1-hr sessions. The younger people were recruited from the Georgetown University student body, and the older by an advertisement in the Health Section of the Washington Post. None of the participants had taken a course in cognitive psychology or had been involved in similar studies. All were paid for participating. Their characteristics are summarized in Table 1. The age groups did not differ significantly in their years of education completed, WAIS-R Vocabulary score, or self-rating of health, but as is typical, the younger group scored significantly higher than the older on the WAIS-R Digit Symbol test, \( t(10) = 4.44, p < .01 \) (Wechsler, 1981).

Design. The design was a 2 X 2 X 6 (Age x Trial Type x Session) mixed factorial, with age (young vs. old) as a between-subjects variable and trial type (pattern vs. random) and session (1–6) as within-subjects variables.

Stimuli and apparatus. Participants were seated in front of a personal Macintosh SE computer (Apple Computer, Inc., Cupertino, CA) with a 9-in. monitor. They were instructed to rest the index and middle fingers of each hand on the following four response keys, which were marked with blue stickers: z, x, , and l. The screen displayed four open circles evenly spaced across the middle of the screen so that they were aligned approximately with the response keys. On each trial, one of the circles was filled in with white and remained so until the participant pressed the key corresponding to this target. An incorrect response elicited a brief tone, and the circle remained filled-in until the correct key was pressed. After a delay of 120 ms, the next target appeared.

The following six patterns were used, one for each of the 6 people in each age group: 1234, 1243, 1324, 1342, 1423, 1432. Here, the numbers 1 to 4 correspond, respectively, to the circles from left to right on the screen.

Procedure. After completing a biographical questionnaire, participants were seated at the computer. People were not informed of any regularity. Instead, they were told, “In this study, we are trying to learn more about how practice affects motor performance. We want to find out just how much people are able to speed their responses when they are given extended practice on a simple reaction time task.” They were
Table 1
Mean Participant Characteristics

<table>
<thead>
<tr>
<th>Experiment Group</th>
<th>Old</th>
<th>Young</th>
<th>Young-old</th>
<th>Old-old</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Male</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>Age (in years)</td>
<td>71.17 (65-87)</td>
<td>20.33 (19-22)</td>
<td>21.00 (20-23)</td>
<td>77.67 (76-80)</td>
</tr>
<tr>
<td>Education (in years)</td>
<td>15.17 (13-17)</td>
<td>14.50 (14-15)</td>
<td>69.00 (65-73)</td>
<td>16.00 (12-20)</td>
</tr>
<tr>
<td>Self-rated health*</td>
<td>4.67 (4-5)</td>
<td>4.33 (3-5)</td>
<td>4.50 (3-5)</td>
<td>4.40 (4-5)</td>
</tr>
<tr>
<td>WAIS-R Vocabulary test</td>
<td>53.33 (36-64)</td>
<td>57.33 (47-66)</td>
<td>57.00 (29-68)</td>
<td>57.67 (46-70)</td>
</tr>
<tr>
<td>WAIS-R Digit Symbol test</td>
<td>53.00 (35-70)</td>
<td>82.67 (71-90)</td>
<td>50.70 (43-62)</td>
<td>46.83 (38-64)</td>
</tr>
<tr>
<td>Computation span test</td>
<td>3.00 (2-5)</td>
<td>5.00 (2-7)</td>
<td>2.00 (1-3)</td>
<td></td>
</tr>
</tbody>
</table>

Note. Ranges indicated in parentheses. WAIS-R = Wechsler Adult Intelligence Scale—Revised.
*Responses ranged from 1 (poor) to 5 (excellent).

Instructed to press the key corresponding to the target circle as quickly as possible while maintaining a "very high" level of accuracy.

After a short (10-trial) practice block, people were presented with 21 blocks of 90 trials. Each of these blocks began with 10 random trials followed by 80 trials in which the pattern trials alternated with random ones. After each block, the screen displayed their speed and accuracy for the most recent block, as well as their speed and accuracy for the block preceding that one so that they could monitor their own performance and attempt to improve. People were required to take at least a 30-s break between blocks to minimize fatigue. At the end of the final block for each day, participants completed a written end-of-session questionnaire containing the following questions:

1. Did you use any strategy to try to improve your performance (speed and accuracy)? If so, what was your strategy?
2. Do you think your strategy worked? Why or why not?
3. Any other comments?

Participants returned to the lab for five additional sessions, with no more than 2 days between sessions. On each day they completed 21 blocks as described above and then the end-of-session questionnaire. At the end of the sixth and final session, participants were given a postexperimental interview, which contained the following questions:

1. Do you have anything to report regarding the task?
2. Did you notice anything special about the task or the material?
3. Did you notice any regularity in the way the stimulus was moving on the screen? (If the participant answered yes, the experimenter probed for more specifics, and then asked the following.)
4. Did you attempt to take advantage of the regularities you noticed in order to anticipate subsequent targets? If so, did this help?
5. In fact, there was some regularity to the sequences you observed. What do you think it was? That is, try to describe any regularity you think might have been there. (The experimenter encouraged people to describe any regularities at all that they noticed, even if they were vague or unsure.)

Then the WAIS-R Digit Symbol and a written version of the WAIS-R Vocabulary tests (Wechsler, 1981) were administered in this order. Finally the participant completed a health screening questionnaire (from Christensen, Moyer, Arisman, & Kern, 1992).

Results and Discussion

Data reduction. The data from the first 8 trials of each block, which were always random, were not analyzed. For each subsequent set of 8 trials (4 random and 4 pattern), the median response time for correct responses was calculated separately.
for random and pattern trials, and the means of these medians were calculated to yield a mean response time for each block for the 40 random and 40 pattern trials. For most of the analyses (except as indicated), the means of these means were then calculated across the 21 blocks of each session to yield a mean session response time for each trial type for each participant, and these were submitted to the group analyses of variance (ANOVAs) referred to below.

Are people of both age groups sensitive to the regularity? The data in Figures 1 and 2 indicate that they are. Figure 1 shows the mean of the median response times for correct responses for pattern versus random trials for younger and older groups, and Figure 2 shows the proportion correct. It is clear that the random and pattern trials diverged across sessions for people of both ages, for both dependent measures. Over sessions, response times declined more for pattern than for random trials for both age groups. Accuracy declined little over sessions for pattern trials, but more for random ones. (Accuracy would out of necessity remain constant or decline over sessions, rather than increase, because performance was nearly perfect in Session 1, and people attempted to improve their speed. If, as we argue below, people become sensitive to which triplets are most likely, then anticipatory responses based on such sensitivity would hurt accuracy on random trials because the anticipations would often be wrong. Anticipations would be less likely to hurt accuracy on pattern trials because they would often be correct, leading to the pattern of divergence observed here.)

These observations were confirmed by ANOVAs, which revealed significant Session × Trial Type (random vs. pattern) interactions for response times, $F(5, 50) = 12.74, p < .0001$, $MSE = 22.0700$, and for proportion correct, $F(5, 50) = 14.47, p < .0001$, $MSE = 0.0001$.

The pattern versus random difference is not just a group effect but can be seen in individual data. We conducted separate Type (random vs. pattern) × Session (1–6) ANOVAs for each person for proportion correct and for mean response time, using blocks within each session to determine error variance. The interaction of Type × Session was usually not significant for these individual subject ANOVAs, presumably because of the reduced power. However, the main effect of trial type (random vs. pattern) was significant, $F(1, 240), p < .05$, for at least one of the dependent measures for every participant. For all 6 younger people and for 4 of the 6 older people, pattern trials were significantly better than random for both proportion correct and response time measures. For 1 older person the pattern–random difference was significant only for proportion correct, and for another older person the difference was significant only for response times.

It is clear, then, that people of both ages are sensitive to the regularity in the sequence, sensitivity that is revealed in both accuracy and in response times and that is apparent by at least one of these measures for every individual tested.

Are there age differences in sensitivity to the pattern? When the group data are examined by sessions, there is an age-related deficit in pattern sensitivity. Figures 1 and 2 and Table 2 suggest that the younger people are showing greater sensitivity than the older people, in the form of a larger difference between pattern and random trials (i.e., the overall trial type effect) and in the form of a greater increase in this trial type effect across sessions. As Table 2 shows, when all sessions were collapsed, the overall trial type effect for accuracy (i.e., accuracy on pattern minus that on random trials) was .07 for the younger people versus .01 for the older. The overall trial type effect for response time (i.e., response time on random minus pattern trials) was 22 ms for the younger and 17 ms for the older people. Further, as
Figures 1 and 2 and Table 2 show, for both dependent measures, the trial type effects were similar for the two age groups on Session 1, but by Session 6 are larger for the younger group.

This apparent age difference in the magnitude of the trial type effect was not statistically significant for the response times, but it was for accuracy. For response times, there was an overall age difference, \( F(1, 10) = 65.41, p < .0001, \text{MSE} = 13,549.3200 \), but no interactions involving age approached significance. In contrast, when accuracy was examined, there was a significant Trial Type \( \times \) Age interaction, \( F(1, 10) = 14.12, p < .01, \text{MSE} = 0.0025 \), and a significant Trial Type \( \times \) Age \( \times \) Session interaction, \( F(5, 50) = 8.20, p < .0001, \text{MS(error)} = .0001 \).

To get a better picture of the time course of learning, we examined group and individual data to determine the earliest session on which a significant trial type effect occurs. At the group level, as early as Session 1, both age groups were responding significantly faster on pattern than on random trials: \( t(5) = 3.10, p < .05 \) for the younger group, and \( t(5) = 2.81, p < .05 \) for the older group. Both age groups were also more accurate on pattern than on random trials in Session 1: \( t(5) = 3.15, p < .05 \) for the younger group, and \( t(5) = 3.02, p < .05 \) for the older group. Nonetheless, the Session 1 accuracy data revealed age-related deficits in the magnitude of the trial type effect, with this effect being larger for the younger group than for the older, \( t(10) = 2.28, p < .05 \).

At the individual level, for each person, we determined the first session in which that person showed significant pattern sensitivity, as indicated by a significant difference between pattern and random trials on the response time and on the accuracy measure \( (p < .05 \) on a matched \( t \) test by using blocks as observations, \( df = 20 \)). These data are shown in Table 3. Consistent with the group data, the individual accuracy data showed a trend toward an earlier divergence of pattern versus random trials for younger than for older people.

It is clear, then, that both age groups showed some pattern sensitivity as early as the first session, but, nonetheless, there were age-related deficits in pattern sensitivity. The magnitude of the trial type effect was greater for younger than for older people as early as Session 1, and the effect increased more over sessions for the younger group than for the older group, at least when accuracy was considered. However, interpretation of this age difference was complicated by the age difference in overall accuracy, which is apparent in Figure 2. As Table 2 shows, the mean overall proportion correct across all sessions was .92 for the younger group, but .98 for the older, a significant age difference, \( F(1, 10) = 14.42, p < .01, \text{MSE} = 0.0103 \). Thus, the older people were extremely cautious in that they made very few errors. We return to this issue in Experiments 2a and 2b.

Were people able to describe the regularity? The end-of-session questionnaires and the postexperimental interviews show that they were not. In their end-of-session responses, several people of both ages reported searching for a pattern, but none felt as though they had found one. When told at the end of the last session (postexperimental Question 5) that there was indeed a pattern, no one was able to describe it with any accuracy. The most common guess was that it had something to do with single repetitions happening frequently (e.g., the Position 1 occurring twice in a row); this guess is inaccurate in that repetitions are no more likely than nonrepetitions. The most accurate description came from 1 older participant who offered that there seemed to be an equal number of targets in each position; this observation was accurate but did not reveal any knowledge of the repeating pattern nor would it have enabled
Table 2

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Young</th>
<th>Old—old</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall accuracy</td>
<td>.92 (.04)</td>
<td>.96 (.02)</td>
</tr>
<tr>
<td>Trial type effect for accuracy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall</td>
<td>.07 (.04)*</td>
<td>.02 (.02)*</td>
</tr>
<tr>
<td>Session 1</td>
<td>.02 (.02)*</td>
<td>.01 (.00)*</td>
</tr>
<tr>
<td>Session 6</td>
<td>.10 (.06)*</td>
<td>.02 (.02)*</td>
</tr>
<tr>
<td>Overall RT</td>
<td>337 (25)</td>
<td>524 (88)</td>
</tr>
<tr>
<td>Trial type effect for RT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall</td>
<td>22 (8)*</td>
<td>10 (5)*</td>
</tr>
<tr>
<td>Session 1</td>
<td>7 (6)*</td>
<td>7 (3)*</td>
</tr>
<tr>
<td>Session 6</td>
<td>33 (18)*</td>
<td>14 (5)*</td>
</tr>
</tbody>
</table>

Note. Standard deviations indicated in parentheses. RT = response time.

* Accuracy on pattern minus accuracy on random trials. * RT on random minus RT on pattern trials.

Despite their inability to describe the pattern verbally, several people mentioned having sensed that some regularity was influencing their responses. For example, on the end-of-session questionnaire for Day 4, one younger person asked, “Is there ever a pattern that my conscious isn’t aware of? Sometimes my fingers would move faster than my mind actually registering what I would see.” Also, 1 older participant responded to Question 3 in the postexperimental interview by saying that there seemed to be a regularity, but she just could not work out how it was repeating; she speculated that her inability to pick it up might be due to her age.

Thus, although many people of both ages searched for a pattern, and some reported a vague sense that one was present, no one was able to describe it accurately. We conclude that learning in this task was implicit, by the criterion stated earlier in the article. That is, every participant revealed sensitivity to the pattern by better performance on pattern than on random trials, and yet no one was able to describe anything about the nature of the regularity that would account for this pattern versus random trial difference.

What is being learned and are there age differences therein? To determine whether people were engaging in higher order or triplet-only learning, we divided each participant’s random trials into two categories: (a) high-frequency triplets (i.e., those identical to triplets that end with pattern trials) versus (b) low-frequency triplets (i.e., those that do not correspond to triplets that end with a pattern trial). As outlined in the introduction, if people are only learning which triplets are frequent, then high-frequency random trials should be identical to pattern ones (because the frequencies of the triplets are identical to each other). In contrast, if people are engaging in higher order learning, then over sessions, pattern trials should become faster, more accurate, or both, than high-frequency random trials.

The mean response times across subjects for these two types
of high-frequency trials (pattern vs. random) are shown with open symbols in Figures 3a and 3b for the younger and older groups, respectively, and the corresponding accuracy data are shown in Figures 4a and 4b. These figures also show the data for the low-frequency random triplets by filled symbols, although these were not relevant for the present analyses in which we examined only high-frequency triplets.

Both dependent measures suggest a similar conclusion; performance on pattern trials was different from that on high-frequency random trials for the younger group, but not for the older group, a trend particularly apparent in the accuracy data. In keeping with these observations, an ANOVA of the high-frequency trials (pattern vs. random) × Age interaction revealed a significant Trial Type (pattern vs. random) × Age interaction, $F(1, 10) = 12.93, p < .01, MSE = 0.0006,$ and a marginally significant Trial Type × Age × Session interaction, $F(5, 50) = 2.34, p < .06, MSE = 0.0002.$ Subsequent separate ANOVAs on each age group indicated that there was a significant effect of trial type for the younger group, $F(5, 25) = 14.35, p < .02, MSE = 0.0011,$ but not for the older group, $F(5, 25) = 1.02, p > .10, MSE = 0.0000.$ The Session × Trial Type interaction did not approach significance for either age group alone.

Also in keeping with the above-indicated observations, the response time data revealed weaker evidence for age differences in higher order learning. An ANOVA on the mean response times for high-frequency triplets only revealed a main effect of trial type (pattern vs. random), $F(1, 10) = 5.15, p < .05, MSE = 77.0069,$ and a Trial Type × Block interaction, $F(5, 50) = 2.66, p < .05, MSE = 32.7936,$ but no significant interactions involving the combination of Age × Trial Type (all $p s > .10$). However, subsequent separate ANOVAs on each age group revealed a significant trial type effect for the younger group, $F(5, 25) = 7.88, p < .05, MSE = 77.7333,$ but not for the older group, $F(5, 25) = 0.15, p > .10, MSE = 76.2806.$

Thus there appear to be qualitative differences in what is being learned by the different age groups, with the younger group, but not the older group, showing significant higher order learning. As noted in considering overall pattern sensitivity above, however, it is difficult to interpret age interactions because of the age differences in overall accuracy.

### Experiment 2a

In Experiment 1, every individual of both ages showed implicit learning of the pattern, as revealed by a significant difference between pattern and random trials, in the absence of any ability to describe the nature of the regularity. However, small but statistically significant age-related deficits did appear in the magnitude of such implicit learning. Further, whereas the younger people showed evidence of higher order learning, the older people did not, seeming only to have learned which triplets were more likely.

In Experiment 2a we had three purposes. First, we wanted to find out if the age-related deficits observed in the first experiment would still be present if error rates were more similar across age groups. Therefore, we introduced a number of variations in procedure, described below, in an attempt to encourage the older people, who had been near perfect accuracy in Experiment 1, to make more errors.

Second, we included both a young-old and an old-old group to find out how implicit pattern learning might vary in old age. To our knowledge, no serial learning studies and only a few studies of implicit memory (e.g., Hultsch, Hertzog, Small, McDonald-Miszczak, & Dixon, 1992; Isingrini, Vazou, & Leroy, 1995) have included such a comparison. These studies have shown no decline of implicit memory, as measured by word stem completion and category exemplar generation, when young-old and old-old people are compared, even though differences between these older groups did emerge on explicit tests of memory for the same material. However, these studies focused on implicit memory for familiar words rather than on the learning of novel sequential patterns, and they showed either very small or no age-related deficits at all when older people were compared with younger ones. So whether there are detectable declines after the age of 65 on the present age-sensitive implicit learning task is an open question.

Third, the task used in Experiment 1 was clearly a data-driven one, in that the task was to respond to a stimulus once it had occurred. We wondered whether people would also demonstrate that they had learned the pattern if they were given a more conceptually driven test (i.e., one in which there are no external stimuli for responding). To find out, we included a production task at the end, in which people were
told to produce a typical sequence by pressing keys on the computer. In this task, the person's key press caused the corresponding circle to fill in. Unlike the generation task used in some earlier serial learning experiments (e.g., Nissen, Knopman, & Schaeter, 1987; D. V. Howard & J. H. Howard, 1989), there was no feedback as to accuracy. We reasoned that evidence of pattern learning during the earlier SRT task would be revealed if people produced more high- than low-frequency triplets during this production task.

We consider the production task to be conceptually driven in the narrow sense that any pattern knowledge revealed during this task could not be due to facilitation of response to an external stimulus (i.e., data driven) because no such stimulus is present. Further, we do not believe that any pattern knowledge revealed on this measure is necessarily explicit; if people produce more pattern-consistent than pattern-inconsistent triplets on this production task, but are unable to describe anything about the nature of the regularities they are producing, then by the criterion we are using here, such knowledge would be considered implicit.

**Figure 3.** Mean of median (Med) reaction times (RTs) for Experiment 1 as a function of testing session and trial type for the younger and older people. Data for random trials are sorted into high-frequency triplets (H; open squares) and low-frequency triplets (L; filled squares). Error bars of one standard error (within-subjects computed on block-level data) are plotted for each point but are only visible when greater in magnitude than the size of the symbol. P = pattern; R = random.
Method

Participants. There were 18 participants: 6 young, 6 young-old, and 6 old-old, each of whom participated in six-hr sessions. They were drawn from the same sources as the previous experiment, though none had participated in Experiment 1. Their characteristics are summarized in Table 1. As had been the case in Experiment 1, the age groups did not differ significantly in their years of education completed, WAIS-R Vocabulary score, or self-rating of health, but they did differ in their WAIS Digit Symbol score, $F(2, 15) = 13.58, p < .001, MSE = 87.8548$. On the Digit Symbol task, the young group had higher scores than the young-old group, $t(9) = 4.42, p < .002$, but the young-old and old-old groups did not differ from each other. The age groups also differed in a measure of computational span, which we added for this experiment, $F(2, 15) = 8.75, p < .01, MSE = 1.6000$. The young-old group had a shorter computational span than the young group, $t(10) = 2.33, p < .05$, and a marginally longer computational span than the old-old group, $t(10) = 1.94, p < .10$.
Design. The design was a $3 \times 2 \times 6$ (Age $\times$ Trial Type $\times$ Session) mixed factorial, with age (young vs. young-old vs. old-old) as a between-subjects variable and trial type (pattern vs. random) and session (1–6) as within-subjects variables.

Stimuli and apparatus. These were identical to Experiment 1, except that the computer for this experiment was a Macintosh 6100.

We also added a test of computational span, to provide a test of working memory, and for this we constructed 84 arithmetic problems, each consisting of two single digits from 1–9 to be either added or subtracted, and three possible answers including the correct one.

Procedure. The procedure was the same as Experiment 1, with the following exceptions. First, we made a number of changes in an attempt to encourage the older groups to make more errors. We removed the tone that had occurred following errors in Experiment 1 because several people had mentioned that they tried hard to avoid hearing it. We changed the instructions to state that people should not try to be perfectly accurate but instead should attempt to achieve an accuracy level of about 90%. Also participants were asked to report their current speed and accuracy scores to the experimenter at the end of each block (during the break). If these scores were very high (above 95%) or low (below 85%), participants were given additional reminders about the accuracy goal; if their accuracy was very high, they were told that they were well within the acceptable error range and that they should try to get faster, sacrificing accuracy to do so. If their accuracy was low, they were told to pay more attention to accuracy, slowing down a bit if necessary. Finally, the end-of-session questionnaire Question 1 was modified to remove the parenthetical reference to ‘speed and accuracy.’

To minimize fatigue and decrease the length of session 6, we distributed the supplemental tasks throughout the sessions. The WAIS Digit Symbol task was given at the end of Session 2, the WAIS vocabulary at the end of Session 3, the computational span task at the end of Session 4, and the health screening questionnaire at the end of Session 5. Session 6 ended with a postexperimental interview (identical to Experiment 1) followed by four blocks of 80 trials each of a production task. For this task, participants were told that sometimes people have learned about regularities even though they may not be able to describe them in words. They were asked to produce a typical sequence by pressing keys on the computer; the person’s key press caused the corresponding circle to fill in. No time limit was imposed, but participants were advised to respond quickly without agonizing over their choices. No feedback was given. Finally, people completed a questionnaire concerning the production task. The questions were as follows:

1. Do you have anything to report regarding your experience with this task?
2. Were you able to remember a typical sequence or do you feel you were forced to guess?
3. Were you able to remember any portions of a typical sequence? (If so, please describe.)
4. Any other comments?

The computation span task administered at the end of Session 4 was modeled after Salthouse and Babcock (1991). People were required to solve a series of simple arithmetic problems (e.g., $2 + 4 = ?$) while retaining the second digit of each problem (e.g., 4) and then to report back all of these retained digits when requested. As in Salthouse and Babcock, people were given three trials (each consisting of the appropriate number of arithmetic problems plus a recall attempt) at each of the seven span lengths ranging from 1 to 7. The constraints placed on the arithmetic problems and the details of administration were identical to those of Salthouse and Babcock except that we tape-recorded the task, allowing 3 to 4 s between arithmetic problems and approximately 2 s for each item to be recalled at the end of each trial. We also added a short practice sequence to ensure that people understood the task. As in Salthouse and Babcock, a participant’s computation span was taken to be the highest span at which the participant was perfect (at both the arithmetic and recall component) for two out of the three trials.

Results and Discussion

Are the three age groups equated on overall accuracy? We made several procedural changes in the present experiment in an attempt to equate the age groups on overall accuracy, but as the accuracy scores in Table 2 reveal, we met with only partial success. Although as we had hoped, accuracy for the older groups was lower for Experiment 2a than for Experiment 1, accuracy also declined for the young group, with the result that there was still a significant main effect of age on overall accuracy in the present experiment, $F(2, 15) = 9.22$, $p < .01$, $MSE = 0.0118$.

The age difference in overall accuracy was due, however, to the difference between the young versus the old-old group, $F(1, 10) = 31.62$, $p < .001$, $MSE = 0.0068$, and the young-old group versus the old-old group, $F(1, 10) = 5.50$, $p < .05$, $MSE = 0.0128$. The young and the young-old groups did not differ significantly from each other in overall accuracy, $F(1, 10) = 2.53$, $p > .10$, $MSE = 0.1587$. Therefore, in the analyses of Experiment 2a that follow, any age differences in pattern learning between the young and young-old groups were unlikely due to overall age differences in accuracy.

Are people of all three age groups sensitive to the regularity? In agreement with Experiment 1, Figures 5 and 6 indicate that they were. Figure 5 shows that response times declined over days for all three age groups, with pattern trials being faster than random ones. In the accuracy data shown in Figure 6, for all three age groups, accuracy declined only slightly over sessions for the pattern trials but declined more dramatically for the random ones. These observations were confirmed by ANOVAs including the data from all three age groups; the Session $\times$ Trial Type (pattern vs. random) interaction was significant for both accuracy, $F(5, 75) = 33.01$, $p < .0001$, $MSE = 0.0002$, and response times, $F(5, 75) = 10.71$, $p < .0001$, $MSE = 0.0594$. In addition, separate ANOVAs performed on each of the three age groups revealed that there was a significant difference between pattern and random trials ($p < .05$ or less), that is, a significant trial type effect, for every age group for both the accuracy and the response time data. The magnitude of the trial type effects for both dependent measures is shown in Table 2. As had been the case in Experiment 1, every participant revealed significant sensitivity to the pattern. By using ANOVAs on individual data (as described in Experiment 1), the main effect of trial type, $F(1, 240)$ was significant for at least one of the dependent measures for every person. Pattern trials were significantly better than random for both accuracy and response time measures for all 6 young people, for 5 of 6 young-old people, and for 4 of 6 old-old people. For the remaining 3 people, the effect on response times was either marginal (1 young-old person and one old-old person) or did not approach significance (1 old-old person).

Are there age differences in sensitivity to the pattern? De-
Despite the fact that all individuals of all ages revealed pattern learning, this experiment again yielded clear age-related deficits in this learning, with age differences in the magnitude of the trial type effect getting larger over sessions on both dependent measures. These trends are apparent in Figures 5 and 6 and in Table 2. When all three age groups were included, the Age × Trial Type × Session interaction was significant for both response times, $F(10, 75) = 2.46, p < .02, \text{MSE} = 20.0594$, and accuracy, $F(10, 75) = 8.16, p < .001, \text{MSE} = 0.0002$.

Furthermore, young people differed from young-old people in the trial type effect and, in turn, young-old differed from old-old people. When only the young and the young-old groups
were compared, the Age × Trial Type × Session interaction was significant for both response times, $F(5, 50) = 4.17, p < .05$, $MSE = 14.6147$, and accuracy, $F(5, 50) = 6.20, p < .001$, $MSE = 0.0002$. When only the young-old people and old-old people were compared, this triple interaction was significant for accuracy, $F(5, 50) = 2.75, p < .05$, $MSE = 0.0001$, but not for response times, $F(5, 50) = 0.45$.

These significant triple interactions make it clear that there are age differences in the time course of pattern learning. As in Experiment 1, to gain a better picture of this time course, we examined group and individual data to find the earliest session in which a significant trial type effect occurred. At the group level, by the response time measure, as Table 2 shows, both older groups showed a significant effect ($p < .05$) in Session 1, but the younger group did not do so until Session 2. This unexpected finding was due to the fact that although all 12 older people were in the direction of having longer Session 1 response times for random than for pattern trials, 1 young person was in the opposite direction. By the accuracy measure, the trial type effect was first significant for the young and young-old groups on Day 2, but not until Day 3 for the old-old group. For the response time measure, the young group did not show a significantly greater trial type effect than the young-old group until Session 6, and the young-old and old-old groups did not differ from each other on this trial type effect for any session. For the accuracy measure, the trial type effect was greater for the young than for the young-old group for the first time on Session 2 and greater for the young-old group than the old-old group for the first time on Session 3.

At the individual level, as had been the case in Experiment 1 there is some evidence of pattern sensitivity appearing earlier for younger people. For example, as Table 3 shows, when the first session at which a person showed significant sensitivity on the accuracy measure was examined, 2 young people, 1 young-old person, and 0 old-old people did so on Session 1.

In general, then, and in keeping with Experiment 1, there were age-related deficits in pattern sensitivity, in the form of an Age × Trial Type interaction, and these age differences increased over sessions.

Were people able to describe the regularity? The responses to the end-of-session questionnaires and postexperimental interview were similar to those of the first experiment, with most people of all ages reporting that they had looked for a pattern but had been unable to find one that they could describe. As before, many said they felt as though they might be responding on the basis of a regularity they could not describe. One young-old person described it as having had a "premonition" of where the target was going. At least one person of each age reported feeling as though their fingers were responding to some regularity that their mind could not detect. The guesses they produced when encouraged to describe any regularities (postexperimental interview Question 5) were again incorrect and usually vague. As in the first experiment, the most common description was that events had a tendency to repeat (e.g., there were often two consecutive occurrences of a given position). Although this does not accurately describe the pattern structure—all pairwise transitions were equally likely—it may nonetheless reflect a kind of explicit (but superstitious) influence on responding.

To determine if responding was influenced by this repetition superstition, we conditionalized the performance measures for each trial on whether the previous event was the same as or different from the response. The results for responses on correct random trials are shown at the top of Table 4 and those for pattern trials at the bottom. In keeping with the participants’ reports, these data revealed a repetition bias for both pattern and random trials and across each of the three age groups. We conclude this because participants responded more quickly (by 65 ms on the average) and accurately (a difference of .12 on the average) when the response was identical to the event on the immediately previous trial. Similar improvements occurred for the pattern trials in that overall people were faster (by 53 ms) and more accurate (a difference of .05) on repetition than on nonrepetition trials. Error trial data were also consistent with the repetition bias, but they are not presented because of low cell frequencies.

These findings indicate that participants in all three age groups developed explicit, albeit inaccurate, knowledge of the pattern sequence, which appears to have influenced their responding. It is important to note, however, that this repetition superstition cannot account for the pattern learning effects we observed because repetitions were, in fact, equally likely on pattern and random trials.

Responses to the queries after the production task were similar to those following the main task. In response to Question 2 regarding whether they "remembered a typical sequence" or were "forced to guess" in producing their sequences, the number choosing "guess" was 4, 3, and 3 for the young, young-old, and old-old groups, respectively, and the number choosing "remember" was 0, 1, and 1, respectively. The remaining 2 people of each age reported "both," "neither," or "not sure."

What is being learned, and are there age differences therein? As in Experiment 1, we looked for evidence of higher order pattern learning by comparing pattern triplets with high-frequency random ones, reasoning that if these trial types diverge, then people must be learning something more than simply which triplets occur frequently. The mean response times for these triplet types are shown with open symbols in Figures 7a, 7b, and 7c for young, young-old and old-old groups, respectively, and the corresponding accuracy data are shown in Figures 8a, 8b, and 8c, respectively.

Consistent with Experiment 1, these figures suggest that the younger people were sensitive to higher order regularities, but neither of the older groups were, with the trends again being particularly apparent for the accuracy data. In keeping with these observations, an ANOVA on the accuracy data for high-frequency triplets only for all three age groups revealed a significant Age × Trial Type interaction, $F(2, 15) = 10.79, p < .01, MSE = 0.0007$, and a significant Age × Trial Type × Session interaction, $F(10, 75) = 2.00, p < .05, MSE = 0.0002$. Subsequent separate ANOVAs on each age group revealed that for the younger group, but for neither of the older groups, there was a significant effect of trial type, $F(1, 5) = 11.50, p < .02, MSE = 0.0019$, and a Trial Type × Session interaction, $F(5, 25) = 3.55, p < .02, MSE = 0.0003$. Furthermore, the young group differed from the young-old group in higher order learning: a separate ANOVA of the high-frequency triplets for these
two age groups indicated that though there was no age difference in overall accuracy ($p > .10$), there was a significant Age × Trial Type interaction, $F(1, 10) = 14.27, p < .01$, $MSE = 0.0010$, and a significant Age × Trial Type × Session interaction, $F(5, 50) = 2.36, p < .053$, $MSE = 0.0002$. For this experiment, the ANOVAs of response times for high-frequency triplets yielded no significant effects of trial type for any age groups.

The analyses of triplets from Experiment 1 and the present experiment suggest that people of all ages must be learning which triplets are relatively frequent. This follows from the finding that people of all three age groups performed better on pattern than on random trials, and this second-order (triplet) structure is the lowest level of structure that differentiates pattern trials from random ones. We reasoned that if people are, in fact, using this knowledge to anticipate events, then the errors on random trials should be more likely to be pattern consistent than pattern inconsistent. Therefore, for each participant, we counted the number of times errors on random trials were consistent with the triplet structure of the pattern for each individual and determined the ratio of consistent to inconsistent errors. Because 16 of the 64 possible triplets are pattern consistent, chance responding would lead to a consistent:inconsistent ratio of .25 on random trials. Higher ratios indicate a tendency to anticipate pattern-consistent events.

The results of this analysis supported our triplet-learning hypothesis, with all age groups producing ratios that were reliably greater than chance (young $M = 0.630$): $t(5) = 10.19, p < .0001$; (young-old $M = 0.573$), $t(5) = 5.57, p < .01$; and (old-old $M = 0.449$), $t(5) = 3.23, p < .02$. Furthermore, this ratio was above chance for every individual in each group. In keeping with the analyses of accuracy and response times as indicated above that showed age differences in the magnitude of pattern sensitivity, there was a trend toward older people producing lower consistent:inconsistent error ratios than younger people, $F(2, 15) = 2.99, p < .09$, $MSE = 0.0172$. The only pairwise comparison to approach significance was that between the young and old-old groups, $t(5) = 2.02, p < .10$.

**Does the production test reveal evidence of the pattern learning, and are there age differences therein?** The production data were analyzed by determining the frequency of pattern-consistent and pattern-inconsistent triplets produced by each individual. Pattern-consistent triplets were those that occurred with high frequency during learning because of the structure introduced by the pattern for that individual. In contrast, pattern-inconsistent triplets reflected the random events and occurred only infrequently during learning. Therefore, the production of more pattern consistent than inconsistent triplets during production is taken as evidence of pattern learning.

The mean frequency of pattern-consistent and pattern-inconsistent triplets and the difference between them are shown in Table 5 for each group. These values were determined by counting the number of triplets actually produced. There are 312 (78 trials × 4 blocks) opportunities for each person to produce 64 ($4^3$) distinct triplets. Chance responding would yield a uniform distribution or 4.875 instances of each triplet type (312 occasions/64 triplets). Of the 64 triplets, 16 were pattern consistent, and the remaining 48 were pattern inconsistent. (The mean frequencies reported in Table 5 are above the chance frequency because people never produced all of the 64 possible triplets.) An ANOVA on these data revealed that pattern-consistent triplets were produced more frequently than inconsistent ones, in that there was a main effect of triplet type (pattern consistent vs. pattern inconsistent), $F(1, 15) = 12.20, p < .01$, $MSE = 1.0191$. However, despite the suggestion of age differences in the table, the Triplet Type × Age interaction was not significant, $F(2, 15) = 1.39$.

### Table 4

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Mean reaction time</th>
<th>Mean proportion correct</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Different</td>
<td>Same</td>
</tr>
<tr>
<td><strong>Experiment 2a</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Random trials</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Young</td>
<td>354 18</td>
<td>293 23</td>
</tr>
<tr>
<td>Young old</td>
<td>458 22</td>
<td>414 31</td>
</tr>
<tr>
<td>Old old</td>
<td>554 100</td>
<td>463 44</td>
</tr>
<tr>
<td>Mean</td>
<td>455 390</td>
<td>65</td>
</tr>
<tr>
<td>Pattern trials</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Young</td>
<td>324 16</td>
<td>283 23</td>
</tr>
<tr>
<td>Young old</td>
<td>433 22</td>
<td>409 40</td>
</tr>
<tr>
<td>Old old</td>
<td>540 101</td>
<td>445 42</td>
</tr>
<tr>
<td>Mean</td>
<td>432 379</td>
<td>53</td>
</tr>
<tr>
<td><strong>Experiment 2b</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Random trials</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accurate (young)</td>
<td>364 22</td>
<td>322 14</td>
</tr>
<tr>
<td>Pattern trials</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accurate (young)</td>
<td>346 25</td>
<td>314 17</td>
</tr>
</tbody>
</table>
Figure 7. Mean of median (Med) reaction times (RTs) for Experiment 2a as a function of testing session and trial type for the young, young-old, and old-old people. Data for random trials are sorted into high-frequency triplets (H; open squares) and low-frequency triplets (L; filled squares). Error bars of one standard error (within-subjects computed on block-level data) are plotted for each point but are only visible when greater in magnitude than the size of the symbol. P = pattern; R = random.
Figure 8. Mean proportion (Prop.) correct for Experiment 2a as a function of testing session and trial type for the young, young-old, and old-old people. Data for random trials are sorted into high-frequency triplets (H; open squares) and low-frequency triplets (L; filled squares). Error bars of one standard error (within-subjects computed on block-level data) are plotted for each point but are only visible when greater in magnitude than the size of the symbol. P = pattern; R = random.
Nonetheless, there is a strong suggestion of age differences on the production test on two other counts. First, *t* tests indicated that the young group was influenced by the pattern they encountered earlier, in that they produced significantly more pattern-consistent triplets than pattern-inconsistent ones, *t*(5) = 3.76, *p* < .02. In contrast, neither older group showed significant evidence of such sensitivity; in neither case was the difference score significantly different from zero (*p* > .10). Second, the data from individuals showed a similar pattern of age-related deficits; whereas all 6 of the young people were in the direction of producing more consistent than inconsistent triplets, this was true for 5 of 6 young-old participants, but only for 2 of 6 old-old participants.

Are age differences in pattern sensitivity due simply to age differences in overall accuracy? Two kinds of comparisons argue against this interpretation. First, in Experiment 2a, despite the fact that the young and young-old groups did not differ significantly from each other in overall accuracy, the critical triple interaction of Age × Trial Type × Session was significant for both accuracy and response time when pattern and random trials were compared for these two groups. In addition, when we examined only the high-frequency triplets for evidence of higher order learning, comparison of these two age groups again yielded significant Trial Type × Age interactions, indicating that they differed in whether they showed higher order learning. Second, as Table 2 shows, although we did not manage to match age groups precisely on overall accuracy within Experiment 2a, we did do so across experiments. That is, the young people in Experiment 1 and the young-old people in Experiment 2a both had overall accuracy of .92, and yet as the table shows, the overall trial type effect appeared to be greater for the young people in Experiment 1 than for the young-old people in Experiment 2a for both accuracy and response time. ANOVAs comparing these two groups confirmed this observation, revealing a significant Group × Trial Type × Session interaction for response times, *F*(5, 50) = 3.73, *p* < .01, *MSE* = 16.3172, though not for accuracy, *F*(5, 50) = 1.30, *p* > .10, *MSE* = 0.0002.

Of course, this triple interaction for the response time data must be interpreted with caution because the different age groups were drawn from experiments that differed from each other in a number of ways, so that age and experiment were completely confounded. However, there is good evidence that it is age, and not experiment, that was causing the interaction because when the young group from Experiment 1 was compared with the young group from Experiment 2a, an ANOVA on the response times for pattern versus random trials indicated that no main effects or interactions of group approached significance (*p* > .10 in all cases). This suggests that the triple interaction, obtained when young people from Experiment 1 were compared with young-old people from Experiment 2a, reflects an age difference.

There is also clear evidence from such cross-experiment comparisons that younger people showed more higher order learning than young-old people, even when the age groups made exactly the same overall error rate. Comparisons between the young people of Experiment 1 and the young-old people of Experiment 2a on the high-frequency random versus pattern triplets revealed a significant Age × Trial Type interaction for both accuracy, *F*(1, 10) = 17.88, *p* < .001, *MSE* = 0.0006, and for response times, *F*(1, 10) = 5.92, *p* < .05, *MSE* = 99.9778. This interaction must be due to age differences, not experiment, because when the young people of the two experiments are compared with each other, no interaction between group and trial type approached significance for either dependent measure.

Experiment 2b

It appears, then, that even when the age groups were equivalent in overall accuracy, there were significant age-related deficits in pattern sensitivity; older people differed from younger ones in the magnitude of the pattern versus random difference and in higher order learning. To put this conclusion to an additional test, we conducted another experiment in which we tested only younger people. The procedure was identical to that of Experiment 2a, except that we attempted to push these people to be as accurate overall as the old-old group of Experiment 2a.

### Method

**Participants.** Six younger volunteers were recruited from the same sources as the previous experiments. Their characteristics are shown in Table 1, and *t* tests established that this younger high-accuracy group did not differ from the younger group in Experiment 2a on any of these characteristics.

**Stimuli and apparatus.** These were identical to Experiment 2a.

**Procedure.** This was identical to Experiment 2a except that the instructions stressed that the person should try to maintain a near-perfect accuracy level. On any block when the accuracy dropped below 95%, the person was told to pay more attention to accuracy, slowing down if necessary.

### Results and Discussion

Were the high-accuracy young people as accurate as the old-old people? The young people tested in the present experiment averaged 96% correct across the six sessions, perfectly matching the level observed for the old-old group of Experiment 2a. Con-

### Table 5

Mean Number of Pattern Consistent and Pattern Inconsistent Triplets Produced per Person for Each Age Group During the Production Blocks of Experiments 2a and 2b

<table>
<thead>
<tr>
<th>Triplet type</th>
<th>Experiment 2a</th>
<th>Experiment 2b</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Young</td>
<td>Old-old</td>
</tr>
<tr>
<td>Consistent</td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Inconsistent</td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Difference</td>
<td>M</td>
<td>SD</td>
</tr>
</tbody>
</table>

* *p* < .05.

* p < .05.

The data from individuals showed a similar pattern of age-related deficits, whereas all 6 of the young people were in the direction of producing more consistent than inconsistent triplets, this was true for 5 of 6 young-old participants, but only for 2 of 6 old-old participants.

Nonetheless, there is a strong suggestion of age differences on the production test on two other counts. First, *t* tests indicated that the young group was influenced by the pattern they encountered earlier, in that they produced significantly more pattern-consistent triplets than pattern-inconsistent ones, *t*(5) = 3.76, *p* < .02. In contrast, neither older group showed significant evidence of such sensitivity; in neither case was the difference score significantly different from zero (*p* > .10). Second, the data from individuals showed a similar pattern of age-related deficits; whereas all 6 of the young people were in the direction of producing more consistent than inconsistent triplets, this was true for 5 of 6 young-old participants, but only for 2 of 6 old-old participants.

Are age differences in pattern sensitivity due simply to age differences in overall accuracy? Two kinds of comparisons argue against this interpretation. First, in Experiment 2a, despite the fact that the young and young-old groups did not differ significantly from each other in overall accuracy, the critical triple interaction of Age × Trial Type × Session was significant for both accuracy and response time when pattern and random trials were compared for these two groups. In addition, when we examined only the high-frequency triplets for evidence of higher order learning, comparison of these two age groups again yielded significant Trial Type × Age interactions, indicating that they differed in whether they showed higher order learning. Second, as Table 2 shows, although we did not manage to match age groups precisely on overall accuracy within Experiment 2a, we did do so across experiments. That is, the young people in Experiment 1 and the young-old people in Experiment 2a both had overall accuracy of .92, and yet as the table shows, the overall trial type effect appeared to be greater for the young people in Experiment 1 than for the young-old people in Experiment 2a for both accuracy and response time. ANOVAs comparing these two groups confirmed this observation, revealing a significant Group × Trial Type × Session interaction for response times, *F*(5, 50) = 3.73, *p* < .01, *MSE* = 16.3172, though not for accuracy, *F*(5, 50) = 1.30, *p* > .10, *MSE* = 0.0002.

Of course, this triple interaction for the response time data
sequently, our objective was achieved, and in the following analyses we compared this high-accuracy young group to the old-old group of Experiment 2a.

Are the high-accuracy young people sensitive to the regularity? Our analyses revealed that they were. As was the case with people of all ages in both earlier experiments, every high-accuracy young person showed statistically significant pattern-random differences on both the response time and accuracy measures.

However, the high-accuracy young people revealed a smaller pattern-random difference than their age mates in Experiment 2a. When these two groups were compared, the Trial Type (pattern/random) \times Group (high- vs. low-accuracy young people) \times Session interaction was significant for the accuracy data, \( F(5, 50) = 12.20, p < .0001, MSE = 0.0001 \), as was the Trial Type \times Group interaction, \( F(1, 50) = 40.01, p < .0001, MSE = 0.0008 \). For the response time data, the latter interaction approached significance, \( F(1, 50) = 4.31, p < .07, MSE = 73.7389 \). These findings indicate that pushing young people to be highly accurate reduced the size of the pattern-random difference, and hence the evidence for pattern sensitivity, and so it suggests that at least some of the age differences in pattern learning observed between age groups differing in overall accuracy in earlier experiments could be due to the accuracy difference.

Are there age differences in pattern sensitivity when the age groups are matched on overall accuracy? Age differences in pattern sensitivity between the high-accuracy young people from Experiment 2b and the old-old people in Experiment 2a did not appear for the response times in this experiment (Figure 9), but they remained for accuracy (Figure 10). This occurred despite the extremely high (and equivalent) levels of accuracy for both age groups. For accuracy, the Trial Type \times Age Group interaction was significant, \( F(1, 50) = 31.76, p < .001, MSE = 0.0001 \), indicating that the difference in accuracy between pattern and random trials, although small, was reliably greater for the high-accuracy young than for the old-old group; \( t \) tests indicated this was the case for all six sessions. When overall accuracy across the six sessions was compared for individuals, no overlap occurred in the two age groups (the smallest pattern-random difference for a high-accuracy young person was .023, and the largest difference for an old-old person was .020).

Additional evidence for age differences in pattern sensitivity can be seen in Table 3; whereas all 6 of the high-accuracy young people showed significant pattern learning by Day 2 on the accuracy measure, only 2 of the old-old people were doing so.

Were the high-accuracy young people able to describe the regularity? These participants' reports were similar to those in the previous experiments. Most said they had tried to identify a pattern and suspected one might be there, but they could not figure out what it was. As in the earlier experiments, several people said that repetitions occurred often.

In keeping with these reports, Table 4 shows that this group’s responses exhibited the superstitious repetition bias observed for all groups in the previous experiment. These high-accuracy young people were faster and more accurate on repetitions than on nonrepetitions for both random events (by 42 ms in response time and .08 in accuracy) and pattern events (by 32 ms and .02). Hence, encouraging accuracy did not eliminate the tendency to develop a repetition superstition.

What is being learned, and are there age differences therein? The mean response times for pattern versus high-frequency ran-
Figure 10. Mean proportion (Prop.) correct as a function of testing session, age, and trial type, for the young high-accuracy group of Experiment 2b (Yacc; circles) and the old-old group of Experiment 2a (O-O; triangles). Error bars of one standard error (within-subjects computed on block-level data) are plotted for each point but are only visible when greater in magnitude than the size of the symbol. P = pattern; R = random.

dom trials are shown for the high-accuracy young people from Experiment 2b by open symbols in Figure 11, and the corresponding accuracy data are shown in Figure 12. As had been the case for all age groups in Experiment 2a, response times comparing pattern versus high-frequency random triplets failed to reveal significant higher order learning.

When the high-accuracy young people from Experiment 2b were compared with the old-old people of Experiment 2a, response times did not reveal significant higher order learning, but the trial type effect did not interact with age group, thus revealing no significant age difference in the high-frequency pattern versus high-frequency random separation. It is likely that the lack of a significant interaction with age here was due to the small size of the trial type effect resulting from the extremely high overall accuracy in both groups. This interpretation is consistent with the finding indicated above that when two age groups were equated at a lower overall accuracy (.92 for the young of Experiment 1 and for the young-old of Experiment 2a), then significant Age × Trial Type (pattern vs. high-frequency random) interactions emerged.

The secondary analysis comparing the proportion of pattern-consistent to pattern-inconsistent errors agreed with the results of Experiment 2a in revealing age differences in the nature of the errors. Despite the extremely low error rate here, high-accuracy young people produced more pattern-consistent than pattern-inconsistent errors ($M$ ratio $= 0.572$), $t(5) = 12.03$, $p < .0001$, with all 6 people producing proportions greater than the chance level of .25. This ratio was marginally greater than that observed for the old-old group in Experiment 2a ($M$ ratio $= 0.449$), $t(5) = 2.31$, $p < .07$. Thus even when matched perfectly on overall accuracy, younger people still showed more evidence than older people of using what they have learned about the triplet probabilities to anticipate the next item.

**Does the production task reveal that pattern learning had occurred, and are there age differences therein?** The answer in both cases is yes, as shown in Table 5. The high-accuracy young people were like their age mates in Experiment 2a in producing significantly more consistent than inconsistent triplets during the production task, $t(5) = 5.54$, $p < .01$. In fact, the high-accuracy young group was in the direction of revealing more pattern learning than the lower accuracy young people from Experiment 2a on this measure, though the difference was not significant ($p > .10$). As had been the case for the low-accuracy young people, every high-accuracy young person produced more pattern-consistent than pattern-inconsistent triplets.

The evidence for age differences on the production task is also stronger in this experiment than it was when the age groups were not matched on accuracy. Here, unlike the previous experiment, the Age × Triplet Type (consistent vs. inconsistent) interaction approached significance, $F(1, 10) = 3.60$, $p < .09$, $MSE = 1.1945$. Hence, it is clear that the trend toward age differences on the production measure in Experiment 2a could not be attributed to differences in overall accuracy.

**General Discussion**

The present experiments, using a serial pattern learning task in which pattern events alternate with random ones, lead to several conclusions. First, we found that every participant per-
formed better on pattern than on random trials, despite the fact that no one was able to describe accurately anything about the nature of the regularity. We conclude from this that people of all ages spontaneously and unknowingly learn about subtle sequential regularities of the sort studied here. Presumably, such implicit learning enables people to adapt throughout life to new routines and environments, such as those that accompany moving to a new residence or taking up new activities.

Second, despite the fact that everyone showed sensitivity to the pattern, there were reliable age differences in the magnitude of implicit pattern learning; performance on pattern and random trials diverged to a greater degree for younger than for older people in all three experiments. Furthermore, in the second experiment there were differences in pattern learning between the young-old and old-old groups, suggesting that this variety of implicit learning declines with age even when only people over 65 years of age are compared.

Third, not only did we find this quantitative age difference in the amount of implicit learning, but we also found qualitative age differences; older people differed from younger ones both in what is learned and in the conditions under which such learning is revealed. Regarding what is learned, we found that participants of all ages acquire at least second-order statistical information about the event sequences (i.e., information about runs of three events or triplets). This follows from the fact that the four events occurred equally often overall for pattern and for random trials, as did the 16 possible pairs of events. Hence, the stimulus sequences contained neither zero- nor first-order information in a statistical sense. So to show better performance on pattern than on random trials, as people of all ages did, they must have learned at least second-order information (i.e., which triplets are relatively likely to occur). This interpretation receives further support from the evidence that when people made errors on random trials, the errors were more likely to be pattern consistent than pattern inconsistent for all age groups. The additional finding that this pattern-consistent error effect was marginally greater for younger than for older people, even when they were equated on overall accuracy (Experiments 2a and 2b), is also consistent with our conclusion that there are age differences in the magnitude of implicit learning.

In contrast, our analysis of three-trial sequences or triplets indicated that the young people also acquired higher order information about the sequences, but the older people did not. This follows from the finding that for younger people, but not for older ones, pattern trials diverged from high-frequency random ones. Hence, the older participants appeared to have learned only which three-trial sequences were most likely, but the younger participants acquired some higher order information as well. Whether older people would learn higher order contingencies if given more training is a question for further research.

The other qualitative difference in learning for which we found more tentative evidence concerns the conditions under which learning is revealed. Our younger adults demonstrated that they had learned about the pattern in the alternating SRT task, not only by their performance on the data-driven SRT task itself, but also in the subsequent conceptually driven production test in which they attempted to recreate (without feedback) sequences like those they had encountered earlier; younger people produced reliably more pattern-consistent than pattern-inconsistent triplets during the production task, whereas older adults did not. Thus, older people, despite having learned about the pattern during the data-driven SRT task, were not able to
Figure 12. Mean proportion (Prop.) correct as a function of testing session and trial type for the high-accuracy young people in Experiment 2b. Data for random trials are sorted into high-frequency triplets (H; open squares) and low-frequency triplets (L; filled squares). Error bars of one standard error (within-subjects computed on block-level data) are plotted for each point but are only visible when greater in magnitude than the size of the symbol. P = pattern; R = random.

demonstrate this learning subsequently when they had to produce sequences themselves rather than respond to stimuli that were present. This age difference in the production task was unlikely due to forgetting by the older people because the production task occurred within a few minutes of the end of the SRT task. Our findings here are consistent with the fact that age-related deficits are always observed on the similar generation task that has been used with the original SRT task (e.g., D. V. Howard & J. H. Howard, 1989, 1992). Given that people received explicit memory instructions for the present production task (i.e., they were told to try to make sequences like those encountered earlier), this age difference seems to be one more example of a circumstance in which older adults are poorer than younger at engaging in voluntary retrieval.

Nonetheless, as stated earlier, we do not conclude that the performance of the younger people on this production task relies on explicit learning; by the definition of implicit-explicit learning we are using here, it does not. No one of any age was able to give an accurate description of the nature of the regularity in the pattern, even after the production task. The one regularity people reported most frequently (i.e., that repetitions are more frequent than nonrepetitions) was inaccurate, and though our analyses indicated that this repetition superstition did influence responding, it would not contribute to the predominance of pattern-consistent over pattern-inconsistent triplets the younger people produced during the production task.

Are Age Differences in Learning Due Only to Overall Accuracy Differences?

It is clear that they are not. We compared age groups matched perfectly at two different levels of overall accuracy: namely, young people from Experiment 1 matched with young-old people of Experiment 2a at a level of .92 overall accuracy, and young people from Experiment 2b with old-old people of Experiment 2a at a level of .96 accuracy. In both cases, significant age differences in pattern sensitivity were obtained, including age differences in higher order learning, despite the fact that the overall error rates were identical for the two age groups.

Nonetheless, cross-experiment comparisons did suggest that the magnitude of the implicit learning measure did vary with overall accuracy. It is not clear whether this means that the actual amount of implicit learning is greater when people make more errors, or, as seems more likely, that there is just more room to show learning (the task is more sensitive) when people are making more errors. That is, when people maintain high accuracy, there is little room to show learning on the accuracy measure, and it is this measure that is the more sensitive indicator of learning in this task.

Therefore, in future studies, attempts should be made to equate overall accuracy in all age groups to be compared. The 92% level seems appropriate because it allows enough room to reveal age differences in higher order learning, and it is possible to guide both younger and older people to perform at this level by providing postblock feedback and reminders of the sort we used in Experiments 2a and 2b.

Age Deficits in Implicit Serial Learning

Our observation of an age deficit in implicit learning deserves further discussion because at first glance it seems at odds with a variety of previous findings for serial pattern learning from our own and other laboratories (e.g., D. V. Howard & J. H.
Howard, 1992; Segor, 1994). However, the alternating SRT task differs from the original in at least three important ways, and so comparing the two tasks offers insights into the nature of the age-deficits observed here. The alternating SRT task differs in what must be learned, in the inclusion of potentially distracting stimuli, and in the demands it places on working memory. Any or all of these features might be leading to the age deficits observed here, and so we consider each below. Although the present data do not allow us to decide among them, this analysis generates hypotheses to be investigated in future studies.

First, age differences in implicit learning may appear with the alternating SRT task, but not with the original, because the tasks differ in what must be learned. Earlier research on aging has followed the convention established in the original Nissen and Bullemer (1987) experiment of presenting a repeating pattern of stimulus events. Sequences generated in this way contain structure at a variety of levels (e.g., Reed & Johnson, 1994). For example, individual events and event pairs typically occur with different frequencies reflecting low-order information absent in the present sequences. In contrast, the alternating SRT task requires that the person acquire, at a minimum, second-order sequence information. Previous work with both younger (e.g., Stadler, 1992) and older (e.g., Jackson & Jackson, 1992) participants has demonstrated that the statistical structure of a sequence influences its difficulty of learning. According to this alternative, then, age differences occur only in the learning of second- or higher order sequence information; so earlier studies using the original SRT task did not detect such age deficits because learning could be based solely on low-order information that people of all ages acquire equally. This explanation is consistent with the present finding of qualitative differences in the sequential dependencies learned by younger versus older people.

Second, in the alternating SRT procedure, unlike the original, random events alternate with pattern ones. These random events may disrupt organizational processes during acquisition, and this disruption may be greater for old than for younger individuals. Previous studies have demonstrated that superimposing secondary tone counting on the original SRT task can interfere with pattern learning (e.g., Cohen, Ivry, & Keele, 1990; Nissen & Bullemer, 1987). Stadler (1995) has argued that this reflects the secondary task’s interference with pattern organization, and he has demonstrated similar disruption when temporal variation is introduced into the pattern. The random events in the alternating task may similarly interfere with organizational processes in pattern acquisition. In addition, the older participants have greater within-subjects variability in response times than the younger people; for Experiment 2a, the mean within-subjects variability in response times was 42 ms ($SD = 10$) for the young, 68 ms ($SD = 14$) for the young-old, and 83 ms ($SD = 42$) for the old-old. Because the person’s response time on a given trial contributes to the interstimulus interval, this greater within-subjects variability in the older people results in their receiving a stimulus sequence with more temporal variation than the younger people, with resulting greater opportunity for disrupting organizational processes during learning. Furthermore, greater disruption may occur for old than young because older adults are sometimes more susceptible to interference than the young (e.g., Park, Smith, Dudley, & LaFronza, 1989; but see Light & Prull, 1995).

The third possibility, concerning differing demands on working memory, is suggested by the work of Frensch and his colleagues, who proposed that implicit serial learning depends on both short-term memory capacity and the activation level of pattern elements (e.g., Frensch & Miner, 1994). They have shown that the magnitude of implicit learning is related to memory span (people with larger spans learn more) as well as to the rate of element presentation (slower presentation rates lead to lower activation and, hence, less learning; Frensch & Miner, 1994). These factors should be especially important in the alternating SRT task because learning the second- and higher-order structures that characterize such patterns would be particularly demanding of working memory. Because working memory capacity declines in old age (Salthouse & Babcock, 1991), the alternating SRT task would be particularly difficult for older people. Indeed, the computation span measure of memory span assessed in Experiment 2a showed reliable age differences, and computation span correlated highly with both the speed ($r = .54, p < .05$) and the accuracy ($r = .74, p < .001$) measures of implicit learning. Furthermore, our older participants effectively received a slower presentation rate than the younger participants because of their longer response times. According to Frensch’s theory, the resulting lower activation would also serve to inhibit pattern learning for these individuals. This explanation, like the other two, is consistent with the present results; deciding among them most await further research.

In summary, the present experiments revealed that people of all ages are sensitive to subtle regularities in a repeating serial sequence but that there are age-related deficits in both quantitative and qualitative aspects of this implicit learning. The older groups in these studies showed less implicit pattern learning than the younger group. Also, unlike the younger people, the older ones failed to learn about third-order contingencies and failed to reveal pattern learning later when tested in a more conceptually driven production test. The mechanisms underlying these age-related deficits in implicit serial learning are yet to be determined.

References


Received October 21, 1996
Revision received February 4, 1997
Accepted February 4, 1997

---

**Dannemiller Appointed Editor of Developmental Psychology, 1999–2004**

The Publications and Communications Board of the American Psychological Association announces the appointment of James L. Dannemiller, PhD, University of Wisconsin, as editor of *Developmental Psychology* for a 6-year term beginning in 1999.

Effective January 1, 1998, manuscripts should be directed to

James L. Dannemiller, PhD  
Developmental Psychology Journal Office  
Room 555 Waisman Center  
University of Wisconsin—Madison  
1500 Highland Avenue  
Madison, WI 53705-2280  
email: jldannem@facstaff.wisc.edu