Baseline Performance and Learning Rate of Procedural and Declarative Memory Tasks: Younger Versus Older Adults

Eli Vakil and Dafna Agmon-Ashkenazi

Psychology Department, Bar-Ilan University, Ramat-Gan, Israel.

In recent years there has been an increasing interest in the residual learning capabilities of amnesic patients and of elderly people. The dissociation between impaired and preserved memory task performance has been proposed to reflect different systems — declarative versus nondeclarative (Squire, 1994) or explicit versus implicit (Schacter, 1987). While this dissociation is well established in the literature on amnesia, it is still controversial with regard to the effect of age. Declarative memory is typically assessed with the use of tasks such as free recall and recognition. Performance on these types of tasks is impaired in amnesic patients (Squire, 1987) and has been shown to be adversely influenced by age (for review, see Burke & Light, 1981; Light, 1991).

Nondeclarative memory consists of several memory subtypes (Squire, 1994). One of the subtypes of nondeclarative memory is skill/procedural learning (Cohen & Squire, 1980), which is measured by tasks such as the Tower of Hanoi puzzle (THP). The THP is a well-studied procedural task, and the ability to solve it has been found preserved in amnesics (Cohen, Eichenbaum, Deacedo, & Corkin, 1985).

In this task, participants are required to move a number of disks placed on one of three pegs to another peg in a minimum number of moves. Another procedural task is the Porteus mazes (PM) (Porteus, 1950), in which participants are repeatedly asked to solve a number of mazes. Performance on this task was also found preserved in amnesics (Brooks & Baddeley, 1976).

The effect of age on procedural learning is controversial. Moscovitch, Winocur, and McLachlan (1986) presented young, elderly, and memory-impaired participants with either normal sentences or sentences in which all letters were rotated. Participants were then tested for both recognition and reading speed. Although the younger group read faster than the older group, both groups improved at the same rate over sessions (i.e., nonsignificant Group × Session interaction). Another procedural task, serial reaction time, was originally introduced by Nissen and Bullemer (1987). Here, the subject's task is to push keys in the corresponding repeated sequence of asterisks appearing on a computer screen. The learning of the sequence, which is evident by the reduction of reaction time over training, is demonstrated even in participants who report no awareness of the repeated sequence (Willingham, Nissen, & Bullemer, 1989). Howard and Howard (1989) have compared younger and older adults on this task. In their study, although the older group's overall reaction time was slower than that of the younger group, both groups displayed a parallel decrease in reaction time over training sessions (i.e., nonsignificant Group × Session interaction). Additional procedural tasks that have been found preserved are rotor pursuit (Eslinger & Damasio, 1986; Heindel, Butters, & Salomon, 1988), and mirror tracing (Gabrieli, Corkin, Mickel, & Growdon, 1993; Mickel, Gabrieli, Rosen, & Corkin, 1986). Other studies, however, have reached the opposite conclusion, i.e., that procedural learning is age-sensitive. Wright and Payne (1985) reported that older adults' performance on rotor pursuit and mirror tracing tasks was found inferior to that of younger adults.

Of particular interest is the study by Davis and Bernstein (1992), who compared the performance of older and younger adults on the THP. In their study, all participants made four solution attempts in four sessions (i.e., a total of 16 trials). Participants ranging in age from their twenties to their eighties were compared on the average number of moves required for the four attempts in each session. Results showed that the older the participants, the more moves they required to solve the THP. Unfortunately, it is not clear from Davis and Bernstein's report whether the interaction of Age Group × Sessions is significant.

In our opinion, a distinction should be made between the
initial baseline performance level and the rate of learning. We see the latter as the crucial measure of procedural learning, while the former reflects the basic ability of problem solving (which itself is probably dependent on several other capacities, such as working memory). This view is in accord with conclusions made by Moscovitch et al. (1986) and Howard and Howard (1989), as reported above. Specifically, despite slower reading time or reaction time, the lack of significant interaction between age group and learning session was interpreted as indicating similar procedural learning for the different age groups.

The goal of the present study is twofold: first, to compare older and younger adults on their baseline ability as well as on their rate of learning, and second, to make this comparison in reference to both procedural and declarative tasks. This paradigm allows for comparison between the learning rates of the groups for both types of task. Accordingly, only in the case of a dissociation between the learning rate of the declarative and the procedural tasks (i.e., a steeper learning rate by the younger group on the declarative but not on the procedural task), could we conclude that the first but not the second is age-related.

**Method**

**Participants**

A group of older adults and a group of younger adults participated in the present study. The older adult group consisted of 25 participants (12 males and 13 females) whose age ranged from 60 to 84 years ($M = 69.76$) and whose educational level ranged from 8 to 16 years of schooling ($M = 12.44$). All older individuals participated in a special series of courses for elderly people offered at Bar-Ilan University (Israel). Although all of them claimed to be active and alert, the Mini-Mental State Exam (Folstein, Folstein, & McHugh, 1975) was administered as an objective screening test. The range of scores on the test was from 27 to 30 points ($M = 28.48, SD = 0.87$). Thirty points is the maximum possible score for this test, and the normal range is 26 to 29 points ($M = 27.6, SD = 1.7$) (Folstein et al., 1975). One participant was rejected because of performing below the normal range in this test, and two more were rejected for taking psychiatric medication. All the elderly people were alert and oriented to time and place when tested. They claimed to be in good health and had no uncorrected vision or hearing problems.

The group of young adults consisted of 25 volunteers (12 males and 13 females) whose age ranged from 21 to 35 years ($M = 27.32$) and whose educational level ranged from 12 to 18 years of schooling ($M = 14.36$). For comparison purposes, the younger group was also administered the Mini-Mental State Exam. The range of scores on the test was from 27 to 30 points ($M = 29.36, SD = 0.87$).

All participants volunteered for the experiment. Participants in both groups were proficient in Hebrew and had no history of mental illness, CNS disease, alcoholism, or drug use. A large proportion of the Israeli population arrived in Israel as new immigrants. Since many of the elderly participants immigrated as youngsters, they had to leave school and go to work, and therefore did not have the opportunity to fulfill their intellectual potential. For this reason, we felt that their employment level would better reflect their intellectual capacity than would their formal educational level. We used a 3-point employment scale: a score of 1 was given for nonprofessional work, a score of 2 for professional work, and a score of 3 for managerial-academic employment. Six of the elderly people were classified as 1, eleven as 2, and eight as 3. Although the groups differ in number of years of schooling, $t(48) = 3.36, p < .003$, participants in the younger group were exactly matched to the older group for employment level.

**Tests and Procedure**

Participants were tested individually, in two sessions held one week apart. Two tasks were employed to test procedural memory: THP (Cohen et al., 1985) and PM (Porteus, 1950). Two declarative tests were employed as well: Visual Paired Associates (VPA), a subtest of Wechsler Memory Scale—Revised (WMS–R; Wechsler, 1987), and the Rey Auditory-Verbal Learning Test (AVLT) (Lezak, 1983). Three considerations determined selection of the declarative tasks: First, they are standard memory tests; second, we chose tests that provide memory measures that are parallel to those measured by the procedural tasks (i.e., baseline and learning rate). Finally, to have a broad sample of declarative memory ability, we selected a verbal and a visual memory test.

The THP and the PM tasks were administered three times: twice in the first session, half an hour apart, and once in the second session, one week later. Thus, participants performed each procedural test (i.e., THP and PM) a total of three times. With regard to the declarative tests, half of the participants from each group were administered one of the tests (e.g., AVLT) in the first session, and the second test (e.g., VPA) was administered in the second session. For the other half of the participants, tests were given in reverse order.

**Tower of Hanoi puzzle (THP).** — The task consists of five plastic disks and three wooden pegs numbered from 1 to 3 from left to right, respectively. At the outset, five disks are arranged on the left-most peg with the largest disk at the bottom and the smallest disk at the top. Participants are told that the goal is to move the five disks from the left-most peg (i.e., no. 1) to the right-most peg (i.e., no. 3) in a minimum number of steps. They are also told that they can move only one disk at a time, they may not place a large disk on a small disk, and they can use the middle peg as well. The optimal solution for five disks requires 31 moves. The experimenter recorded the number of moves and time required to solve the puzzle.

**Porteus mazes (PM).** — Nine mazes are presented in order of difficulty from the simplest (i.e., age 6 level) to the most difficult (i.e., adult level). Participants are asked to mark the way from the starting point to the exit point with a pencil. Participants are also asked not to lift the pencil while solving the maze, not to enter a dead-end alley, and to avoid transection of lines of the maze; doing any one of these is recorded as an error. The experimenter recorded
both the number of errors and the time required to solve each maze.

**Visual Paired Associates (VPA).** — This subtest of WMS–R (Wechsler, 1987) consists of a set of six different colors paired with six nonsense shapes. Each card contains one pair. The same set is repeated at least three consecutive times in a different order each time. At this point, if the participant has successfully learned all six pairs, the repetition stops. If not, the list of pairs is repeated until all six pairs are learned, or until all six trials are administered. Following each set of six cards, six testing cards consisting only of shapes are presented. The participants are presented a folder with eight different colors and are asked to point to the color pair associated with the presented test card shape. One more matching test trial is repeated half an hour after completion of the first set of trials. Notice that unlike the first trials, participants in the delayed trial are presented with only the testing cards to test their retention of the visual pairs. The VPA was administered in the first session for half of the participants and in the second session for the other half of the participants.

**Rey Auditory-Verbal Learning Test (AVLT).** — The Hebrew version of the test (Vakil & Blachstein, 1993) was administered in standard fashion (Lezak, 1983). The test consists of 15 common nouns that are read to the participants for five consecutive trials (Trials 1 to 5); participants are asked to recall as many words as possible. Each trial is then followed by free recall. In Trial 6 an interference list of 15 new common nouns is presented, followed by a free recall test of these new nouns. In Trial 7 participants are asked to recall the first list again. Twenty minutes later, participants are once again asked to recall the first list. The half of the participants who were administered the VPA test in the first session were administered the AVLT in the second session and vice versa.

The testing order of the first session was as follows: (1) procedural tasks (i.e., the THP and the PM in counterbalanced order); (2) the learning phase of one of the declarative tests (i.e., either the three trials of the VPA or the first five learning trials of the AVLT); (3) both procedural tasks in the same order as in Stage 1; (4) the delayed phase of the declarative test in Stage 1 (i.e., either the fourth trial of the VPA or the delayed recall of the AVLT). The order of the second session was as follows: (1) learning phase of the declarative task (i.e., the one not presented in first session); (2) the procedural tasks in the same order as in Stage 1 of the first session; (3) delayed phase of the declarative test in Stage 1.

**RESULTS**

**Tower of Hanoi Puzzle**

The task was administered three times (immediate, half an hour, and one week later). Two separate dependent measures were employed to analyze the data: number of moves for solution and solution time.

**Number of moves.** — The results were submitted to a mixed analysis of variance (ANOVA) to analyze the effect of Age Group (younger and older adults) by Time Delay (immediate, half an hour, and 1 week). The former is a between-subjects factor, and the latter is a within-subjects factor. The two main effects reached significance, but the interaction between them did not. As can be seen in Table 1, the older group needed more moves than did the younger group to solve the puzzle, F(1,48) = 5.09, p < .03 (MSE = 369.14). Overall, there was a significant improvement over time, F(2,96) = 9.34, p < .001 (MSE = 222.13). The Age Group × Time Delay interaction did not reach significance, F(2,96) = .18, p = .83 (MSE = 222.13). The observed power at the .05 level is .08. This finding should be interpreted cautiously because of the low power of the analysis. Thus, these results failed to provide evidence against the claim that both groups improved at the same rate.

**Solving time.** — Analysis of the same variables as above again indicates that both main effects reached significance, but the interaction between them did not. As can be seen in Table 1, it took longer for the older group to solve the puzzle than the younger group, F(1,48) = 18.13, p < .001 (MSE = 12.45). Overall, there was a significant decrease in solving time over time, F(2,96) = 41.82, p < .001 (MSE = 5.31). The Age Group × Time Delay interaction did not reach significance, F(2,96) = .62, p = .54 (MSE = 5.31). The observed power at the .05 level is .15. As noted above, this finding should be interpreted cautiously. Thus, these results failed to find evidence against the claim that both groups improved at the same rate.

**Porteus Mazes**

The task was administered three times (twice in the first session and once in the second session one week later). Two separate dependent measures were employed to analyze the data: number of errors (e.g., entrances to blind alley) and the total time it took to solve all nine PM.

<table>
<thead>
<tr>
<th>Delay Times</th>
<th>Young</th>
<th>Half Hour</th>
<th>One Week</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Moves — THP</strong></td>
<td>(17.47)</td>
<td>(18.50)</td>
<td>(13.13)</td>
</tr>
<tr>
<td>Young</td>
<td>56.24</td>
<td>48.36</td>
<td>41.64</td>
</tr>
<tr>
<td>Old</td>
<td>61.28</td>
<td>56.08</td>
<td>50.12</td>
</tr>
<tr>
<td><strong>Time — THP</strong></td>
<td>(2.24)</td>
<td>(2.06)</td>
<td>(1.80)</td>
</tr>
<tr>
<td>Young</td>
<td>6.60</td>
<td>3.60</td>
<td>3.20</td>
</tr>
<tr>
<td>Old</td>
<td>9.64</td>
<td>5.68</td>
<td>5.44</td>
</tr>
<tr>
<td><strong>Errors — PM</strong></td>
<td>(3.54)</td>
<td>(2.42)</td>
<td>(2.65)</td>
</tr>
<tr>
<td>Young</td>
<td>4.28</td>
<td>2.88</td>
<td>2.20</td>
</tr>
<tr>
<td>Old</td>
<td>13.88</td>
<td>8.72</td>
<td>7.52</td>
</tr>
<tr>
<td><strong>Time — PM</strong></td>
<td>(100.49)</td>
<td>(90.13)</td>
<td>(78.95)</td>
</tr>
<tr>
<td>Young</td>
<td>277.60</td>
<td>220.36</td>
<td>218.68</td>
</tr>
<tr>
<td>Old</td>
<td>565.00</td>
<td>465.40</td>
<td>459.04</td>
</tr>
</tbody>
</table>

*Young group, n = 25; Old group, n = 25.*
Number of errors.—As above, results were submitted to a mixed ANOVA to analyze the effects of Age Group and Time Delay. Table 1 shows that overall the elderly group made more errors than did the younger group, $F(1,48) = 29.84, p < .001$ (MSE = 60.17). The number of errors significantly decreased over time, $F(2,96) = 23.11, p < .001$ (MSE = 10.62). A steeper decrease in number of errors over time was demonstrated by the older group than by the younger group, $F(2,96) = 6.42, p < .002$ (MSE = 10.62). To detect the source of this interaction, the two age groups were compared on the amount of improvement from (a) the immediate to the half-hour delay measure and (b) the half-hour to the one-week delay measure. The first comparison indicates that the amount of improvement of the older adults was greater than that of the younger adults ($5.16$ and $1.40$, respectively), $t(48) = 3.01, p < .005$. However, in the second comparison the amount of improvement of the two age groups was not significantly different ($1.20$ and $0.68$ for the older and younger group, respectively), $t(48) = .52, p = .61$. Additional analyses were conducted to test the possibility that the interaction is due to ceiling effect in the younger adults. A t-test comparing immediate versus half-hour performance within the young group and again within the old group was significant for both age groups, $t(24) = 2.10, p < .05$ and $t(24) = 4.81, p < .001$, respectively. Thus, the possibility that the interaction is due to ceiling effect in the younger adults was not conclusively demonstrated. However, the possibility of ceiling effect still exists, because it is possible that the younger group’s improvement, although significant, could have been even greater.

Solving time.—Analysis of the same variables as above indicates that both main effects reached significance. As can be seen in Table 1, it took longer for the older group than the younger group to solve the PM, $F(1,48) = 52.55, p < .001$ (MSE = 47354.99), and overall there was a significant decrease in solving time over time, $F(2,96) = 33.30, p < .001$ (MSE = 3244.05). The Age Group X Time Delay interaction did not reach significance, $F(2,96) = 2.59, p > .05$ (MSE = 3244.05). The observed power at the .05 level is .51. Thus, these results failed to offer evidence against the claim that both groups improved at the same rate.

Visual Paired Associates (VPA)

As previously described, participants were tested on the correct matching of a color to six nonsense shapes in three consecutive trials and once again after half an hour. Two separate analyses were conducted: the first three trials as a measure of learning and the fourth trial as a measure of retention over time.

Learning.—The results were submitted to a mixed ANOVA to analyze the effect of Age Group (younger and older adults) by Learning Trials (1 to 3). Both main effects and the interaction between them reached significance. The younger group learned more pairs than did the older group, $F(1,48) = 68.54, p < .001$ (MSE = 4.67). As can be seen in Figure 1, there was a significant overall increase in number of pairs learned from trial to trial, $F(2,96) = 24.05, p < .001$ (MSE = .97). The age groups’ learning rates differed, as indicated by the significant Age Group X Learning Trials interaction, $F(2,96) = 6.30, p < .003$ (MSE = .97). To detect the source of this interaction, the two age groups were compared on the extent of improvement from the first to the second trial, and from the second to the third trial. The first comparison indicates that the extent of improvement of the younger adults was greater than that of the older adults ($1.32$ and $0.20$, respectively), $t(48) = 2.97, p < .005$. However, in the second comparison the extent of improvement of the two age groups was not significantly different ($0.68$ and $0.52$ for the younger and older group, respectively), $t(48) = .41, p = .68$.

Retention.—The younger group outperformed the older group when compared on delayed trial performance, $t(48) = 8.04, p < .001$.

Rey Auditory-Verbal Learning Test (AVLT)

The younger and older groups were compared on the learning and retention measures of the AVLT. Learning was assessed by the first five learning trials of the task, and retention was measured by the number of words correctly recalled after a 20-minute delay.

Learning.—The results of both groups in the first five trials of the AVLT were submitted to a mixed ANOVA. The two main effects and the interaction between them reached significance. Overall, the younger group recalled more words than did the older group (see Figure 2), $F(1,48) = 23.32, p < .001$ (MSE = 14.72). There was also a significant increase in the number of words recalled from trial to trial, $F(4,192) = 161.04, p < .001$ (MSE = 1.60). The age groups’ learning rates differed, as indicated by the significant Age Group X Learning Trials interaction, $F(4,192) = 2.86, p < .03$ (MSE = 1.60). To detect the source of this interaction, the two age groups were compared on the extent of improvement from one trial to the next (i.e., trials 1 to 5). In none of these comparisons were the groups significantly different. The transition from trial 4 to 5 came the closest to reaching significance, $t(48) = 1.88, p = .066$.
PROCEDURAL AND DECLARATIVE MEMORY

Figure 2. The mean number of words recalled on the five learning trials of the AVLT, and the number of words correctly recalled after the 20-minute delay, by the younger and older groups.

However, the significant interaction of Age Group × Learning Trials indicates that the overall learning rate of the younger group was steeper than that of the older group.

Retention — Retention was measured by comparing the number of words recalled on the fifth trial with the number of words recalled 20 minutes later. As can be seen in Figure 2, the younger group recalled more words overall than the older group, $F(1,48) = 29.90, p < .001$ (MSE = 9.33), and fewer words were recalled on the delayed trial as compared with the fifth trial, $F(1,48) = 43.36, p < .001$ (MSE = 1.16). The forgetting rate of the older group was steeper than that of the younger group, as suggested by the significant interaction between these two factors, $F(1,48) = 4.55, p < .04$ (MSE = 1.16).

DISCUSSION

Younger and older adults were compared on two declarative and two procedural learning tasks. As mentioned in the Methods section, the particular declarative and procedural tasks were chosen to enable a more specific comparison of the different components of the tasks, that is, baseline and learning rate. The younger group baseline performance was better than that of the older group for all tasks employed in this study. A dissociation between the learning rate of declarative and procedural tasks was demonstrated in elderly participants. The younger group demonstrated a steeper learning rate than the older group in the declarative tasks. By contrast, procedural task performance differed. For the THP task, whether measuring number of moves or time, and for the time measure of the PM task, the learning rates of the two groups did not differ consistently. The number of errors measured in the PM task was the only exception in which the elderly group showed a significantly steeper learning rate than that of the younger group. It is unlikely that these results on the PM task reflect a reliable advantage for the older group. As mentioned above, this finding, although not proven conclusively, could possibly reflect a ceiling effect in the younger group's performance on this task. In any event, the present results demonstrate a clear differential age effect on the learning rate of the declarative and the procedural tasks. While the younger group showed a steeper learning rate than the older group on the declarative tasks, on the procedural tasks the two groups did not differ consistently in their learning rate, and in one measure the older group even showed a steeper learning rate.

A clear advantage of these results is that this dissociation is demonstrated in the same group of participants. This indicates that these findings do not simply reflect a failure to reject the null hypothesis, but rather they reflect a predicted pattern of dissociations. Another strength of the present study is the use of standard testing instruments for the declarative task, rather than the arbitrarily selected word lists employed in many other studies. However, the differences between the declarative and the procedural tasks themselves raise some concerns as to whether differential rates of improvement are due to fundamental differences in the procedural/declarative nature of the tasks or whether they are due to differences in some other elements of the tasks.

The results obtained reinforce the importance of distinguishing, in procedural learning, between baseline performance and the rate of learning of a task, since the former but not the latter was found to be adversely affected by age. A similar pattern of results was previously found with other procedural tasks by Moscovitch et al. (1986), using reading speed, and Howard and Howard (1989), using serial reaction time.

The different baseline inherent in comparing younger and older adults raises questions with regard to the appropriate comparison of learning rate. For example, should learning be calculated as a proportional/percentage improvement or as absolute improvement? We preferred the latter option primarily because previous studies of procedural memory and aging have applied this same approach (i.e., Howard & Howard, 1989; Moscovitch et al., 1986). An alternative possibility to addressing this issue in future studies is to equalize performance of younger and older adults at the baseline, either by enabling more practice or by administering a less difficult version of the task to older adults. Finally, the distinction between baseline and learning rate in procedural tasks has not only conceptual but also diagnostic value in further characterizing the specific memory deficits associated with normal aging.

ACKNOWLEDGMENTS

This study was supported by a grant from the Brookdale Institute of Gerontology and Adult Human Development in Israel, and ESHEL, Association for the Planning and Development of Services for the Aged in Israel.

Address correspondence to Dr. Eli Vakil, Psychology Department, Bar-Ilan University, Ramat-Gan, 52900, Israel. E-mail: vakile@ashur.cc.biu.ac.il

REFERENCES


Received March 6, 1996
Accepted April 2, 1997