Aging and Prospective Memory: Differences Between Naturalistic and Laboratory Tasks

Peter G. Rendell1 and Donald M. Thomson2

1Australian Catholic University, Melbourne, Australia.
2Charles Sturt University, Bathust, NSW, Australia.

The contrasting age-related trends on laboratory and naturalistic prospective memory (PM) studies were investigated with the same participants. In the first two experiments, 380 participants in three age groups (20s, 60s and 80+) were given a naturalistic PM task of logging the time at four set times for one week. There were six between-subjects regimens that varied the complexity of the time schedule, and the opportunity to use conjunction cues and external aids. The 60s and 80+ age groups did not differ and both older adult age groups were consistently superior to the young adults on all regimens. In Experiment 3, the same participants showed a significant age-related decline on retrospective memory tasks, and on event-based and time-based laboratory PM tasks embedded within the retrospective memory tasks. The study confirmed the paradoxical age-related trends on laboratory and naturalistic PM tasks.

Prospective memory is usually defined in contrast to retrospective memory with the former involving remembering to do things in the future, and the latter involving remembering information acquired in the past (McDaniel & Einstein, 1992; Winograd, 1988). Prospective memory (PM) is involved in tasks such as remembering to turn off appliances, to keep appointments, and to take medication. These activities are important to people of all ages, but they are particularly crucial for the maintenance of independence in elderly people (Einstein & McDaniel, 1990; Park, Hertzog, Kidder, Morrell, & Mayhorn, 1997). Despite its obvious importance, the topic has attracted little attention until recently—the pace of research has picked up markedly in the last 6-8 years, however (see Brandimonte, Einstein, & McDaniel, 1996).

The initial studies of PM found no age differences, or even found superior performances by older, compared with younger, adults (Moscovitch, 1982; Poon & Schaffer, 1982, as cited in Poon, 1985)—a marked contrast to the typical age-related decline found on retrospective tasks (Light, 1991). These earlier findings prompted speculation that PM reflected some distinct type of memory process that is spared the effects of aging. Since these early studies, however, several studies employing laboratory paradigms have reported age-related deficits on PM tasks (Cockburn & Smith, 1988; 1991; Einstein, Holland, McDaniel, & Guynn, 1992; Maylor, 1993a; 1996; Park et al., 1997; West, 1988), although the findings are mixed, with no age differences found on some laboratory PM tasks (Einstein et al., 1992; West, 1988), and Zelinski et al., 1990). An interesting task distinction between early and recent studies of PM is that many early studies involved naturalistic methods, while recent studies have typically used laboratory tasks. A closer examination of the two paradigms is warranted.

Naturalistic PM tasks require participants to carry out a task at prescribed times within their daily life: to telephone the researcher (Devolder, Brigham, & Pressley, 1990; Maylor, 1990; Moscovitch, 1982; Poon & Schaffer, 1982, as cited in Poon, 1985; West, 1988), to post cards back to the researcher (West, 1988), to note the date and time on a take-home questionnaire (Dobbs & Rule, 1987), and to simulate the taking of medication (Leirer, Morrow, Pariante, & Sheikh, 1988; Park, Morrell, Frieske, & Kincaid, 1992; Rendell & Thomson, 1993). Six of these studies included comparisons of young and older adults, and four of the six found that older adults performed better than young adults. West (1988) reported no age difference, but hers was the only phone-in study that did not require subjects to respond at a specific time. Dobbs and Rule (1987) found an age-related decline in the ability to fulfill a request that was embedded within the task of completing a questionnaire. A similar pattern of little or no age-related difference has been found in naturalistic studies contrasting young-old and old-old participants. Park and colleagues (1992) reported an age-related decrement comparing adults in their 60s with those aged over 70 years in a study requiring participants to record their own medication, but the difference was relatively small. Other studies (Leirer et al., 1988; Maylor, 1990; Rendell & Thomson, 1993) found no age-related differences within these older age ranges. In general, older adults are as good or better than young adults at real-life PM tasks.

In laboratory PM tasks, the request to perform a future action is embedded within another laboratory task, usually a retrospective memory task. Laboratory PM tasks initially involved one-off incidental requests, such as remembering to ask for a red pen when asked to draw a circle (Dobbs & Rule, 1987), remembering to ask for a hidden belonging, or remembering to deliver a message (Cockburn & Smith, 1988; 1991; West, 1988; Zelinski et al., 1990). Recent studies have used multiresponse tasks involving pressing a particular key on a keyboard when a target word or pattern appeared during a memory task (Einstein et al., 1992; Einstein & McDaniel, 1990; Mäntylä, 1994; Park et al., 1997), and circling the trial number whenever a face had a beard, or wore glasses, during a “famous people naming task” (Maylor, 1993b; 1996). Nearly all these studies have reported some age-related decline, although two exceptions are by Zelinski and colleagues (1990) and Einstein and McDaniel (1990). Unlike naturalistic PM studies, age differences were reported in studies including only adults of at least 50 years of age (Cockburn & Smith, 1988; 1991; Maylor, 1993a; 1996), as
well as in studies comparing young adults with older adults (Dobbs & Rule, 1987; Einstein et al., 1992; Park et al., 1997; West, 1988). In contrast to naturalistic experiments, none of these laboratory studies has found older adults superior to younger adults. Also, age-related declines were found in laboratory studies in comparisons of young-old and old-old groups.

The purpose of the present study is to investigate the paradoxical age-related trends on naturalistic and laboratory PM tasks. Most laboratory PM studies have found an age-related decline, and some have reported no age differences. In contrast, naturalistic PM studies have tended to find older adults superior to young adults, and have tended to find no differences in PM performance between young-old and old-old adults. These different trends may reflect task differences rather than a general difference between laboratory and naturalistic studies, given that such dissociations are not apparent in studies of retrospective memory, where an age-related decline is typically found on both laboratory and naturalistic retrospective memory tasks (Light, 1991). One possible explanation is that naturalistic PM studies have typically required participants to perform a simple action at set times over several days, whereas laboratory PM studies have typically required participants to perform an incidental task that is embedded within a defined session of some other laboratory task.

The present article reports experiments that follow from a previous study by Rendell and Thomson (1993). This study showed that older adults were superior to young adults on a naturalistic PM task, despite an age-related decline on retrospective memory tasks. Participants were given a portable time-logging device and were required to log the time at prescribed times for seven days. The time schedules were simple and regular, and were either once-a-day or four-times-a-day over a week. The first two experiments in the present article examined age-related trends on a naturalistic PM task similar to that used by Rendell and Thomson (1993), but with manipulations of task complexity and opportunities to use external aids. The third experiment investigated the performance on laboratory prospective and retrospective memory tasks of the same participants involved in the first two experiments. Unlike previous research, this enabled a comparison of age-related trends on naturalistic and laboratory PM tasks within the same study, and with the same participants.

One possible factor underlying these different age-related trends is the complexity of the PM task. Einstein and colleagues (1992) found an aging-complexity effect on a laboratory PM task. There were no age differences when there was one PM target, but they found an age-related decrement when the task was made more complex by increasing the number of targets to four. McDaniel's (1990) notion is that time-based tasks lack the environmental cues that are inherent in event-based tasks and, therefore, make more demands on self-initiated retrieval. McDaniel and Einstein (1992) argued that age differences in PM tasks occur when the demands on the retrospective component of the tasks are increased. They proposed that PM tasks have both prospective and retrospective memory components: remembering to implement an action at a future event or time, and remembering the content of the action. Demands on the latter component were suggested to give rise to overall age-related decrements, given that an age-related decline is well established on most retrospective memory tasks. The PM tasks in previous naturalistic studies have not placed heavy demands on remembering the content of the required action. It appears that complexity has not been investigated with naturalistic PM studies. One exception is Rendell and Thomson's (1993) contrast of once-a-day with four-times-a-day time schedules for the time-logging task. The frequency of time-logging did not affect participants' accuracy or the age-related trends, suggesting that the aging-complexity effect reported with laboratory PM studies may not be apparent with naturalistic PM studies.

The opportunity to use external memory aids in naturalistic PM studies has been suggested as an explanation for the contrasting age trends on laboratory and naturalistic PM tasks. Indeed, naturalistic studies have been criticized and even dismissed for the lack of control over participants' use of external aids (Kvavilashvili, 1992; McDaniel & Einstein, 1992). However, the evidence on the contribution of external aids is somewhat mixed. Moscovitch (1982) found that by restricting the use of external aids, the age-related improvement was eliminated on a task requiring telephoning the experimenter at set times. Conclusions from this study need some qualification, however, as it was a small-scale study with just 10 young and 10 older adult participants. In a study of older adults, Park and associates (1992) found that performance on the PM task was unrelated to the level of reported use of external aids. In a sample of participants aged 50 years or older, Maylor (1990) found an interaction between age and type of cue used in an appointment-keeping task with an age-related improvement when participants relied on external or conjunction cues, but an age-related decline when participants relied on internal cues.

Einstein and McDaniel (1990) proposed a distinction between PM tasks with an event or time-based target for the to-be-remembered action, and that participants' performance on the PM task was unrelated to the level of reported use of external aids. In a sample of participants aged 50 years or older, Maylor (1990) found an interaction between age and type of cue used in an appointment-keeping task with an age-related improvement when participants relied on external or conjunction cues, but an age-related decline when participants relied on internal cues.
conjunctive cues could be a factor mediating the good performance of older people on naturalistic PM studies, given that Maylor’s (1990) findings suggest that older adults are more accurate on naturalistic PM tasks when using conjunctive cues.

**EXPERIMENT 1**

In Experiment 1, a group of young adults aged 18 to 28 years, and two groups of older adults, aged 60 to 69 years and 80 to 92 years, performed a naturalistic PM task similar to that employed by Rendell and Thomson (1993). The task required participants to press a button on a time-logging device at set regular times over a week. In this study, a Sharp Organizer was the time-logging device. Participants were required to enter the date and time into the electronic Organizer at four set times, each day, for seven days. There were four conditions manipulating the complexity of the time schedules (regimens)—it was expected that increasing regimen complexity would reduce or reverse the superiority of older adults compared to young adults. This prediction was based on the findings of age-related decrements with increasing complexity of laboratory PM tasks (Einstein et al., 1992). The four conditions varied the regularity of the times.

For example, an irregular regimen was 8:20 a.m., 12:35 p.m., 4:10 p.m., and 8:55 p.m., compared with a regular regimen of 8:00 a.m., 12:30 p.m., 5:00 p.m., and 9:30 p.m. The latter regimen replicated a time schedule used by Rendell and Thomson. The conditions also manipulated whether the regimens were changed each day or remained the same each day. It was proposed that these irregular and/or changing regimens would also minimize the opportunity for conjunctive cues—the opportunity to connect the time-logging task to other routine events in the participants’ lives. Rendell and Thomson suggested that older adults may have more routine and structured lives. Therefore, in addition to an aging-complexity effect, it was expected that minimizing the opportunity for conjunctive cues would also reduce the superiority of older adults on this time-logging task.

As in Rendell and Thomson’s study, the challenge of the task was remembering to carry out a task at set times, rather than the content of the time-logging task. There were two modifications to the task used by Rendell and Thomson. Sharp Organizers were used in Experiment 1 because of their reliability and portability as time-logging devices, although they did increase the demands on remembering the content of the to-be-remembered action. A single button press was required to log the time with the device used by Rendell and Thomson but several keys in a set sequence were required to log the time with the Organizer. In addition, one of the regimens required a number to be entered when logging the time. It was expected that these increased demands would reduce the superiority of older adults on the time-logging task, given Einstein and McDaniel’s (1996) proposal that increasing the demands on retrospective memory increased the likelihood of age-related decrements on PM tasks.

**METHOD**

**Participants**

A total of 280 adult volunteers, aged 18 to 92 years, participated in the experiment. They were living independently in their own homes. Most of the participants were from the Melbourne metropolitan area, and the rest were from rural areas in Victoria. There were three age groups: 20s (18–28 years, M = 20.2 years); 60s (60–69 years, M = 65.0 years); and 80+ (80–92 years, M = 83.2 years). The 20s age group had 120 participants (60 men and 60 women), the 60s age group had 80 participants (32 men and 48 women) and the 80+ age group had 80 participants (40 men and 40 women). The young adult participants were all university students. The occupations of the older participants were mixed. The reported occupations in the 60s age group were either work, home duties, or retired (n = 20, 17, & 43 respectively), and in the 80+ age group, the reported occupations were either home duties or retired (n = 19 & 61 respectively). The older participants were recruited from various community groups such as bowling clubs, Probus, Rotary, a horticultural society, The University of The Third Age, and church groups. There was a significant main effect of age group upon the number of years of education completed by participants, F(2,276) = 40.04, MSE = 6.67, p < .001. Tukey post hoc tests revealed the 20s had significantly more years of education, (M = 14.1, SD = 1.3) than the 60s age group (M = 12.0, SD = 3.2), who, in turn, had significantly more years of education than the 80+ age group (M = 10.9, SD = 3.3). The percentage of participants that were taking medication during the week of time-logging was 8%, 50%, and 66% for the 20s, 60s, and 80+ age groups respectively.

Participants were assigned randomly to each regimen condition, with the restriction that there were equal numbers of men and women in each condition for the 20s and the 80+ age group, and there were 8 men and 12 women in each condition for the 60s age group. Before participating, all participants completed a written consent form that contained a brief explanation of the research project.

**Design**

All participants completed the time-logging task requiring them to remember to press a particular sequence of keys on a Sharp Organizer at four prescribed times a day, for seven days. There were two between-subjects variables: age group (20s, 60s, or 80+); regimen condition (Same Regular, Same Irregular, Different Regular, or Different Irregular). In addition, the between-subjects variable of occupation (work, home duties, or retired) was investigated on a post hoc basis for the 60s age group. Sex of participants was not included as an independent variable, as preliminary analyses revealed that sex was not a significant effect.

The regimen conditions were manipulations of regular versus irregular time schedules, and these schedules were either the same on each day, or different, with a set of four times given each day. In each regimen condition for each day, one prescribed time was in the morning, one around midday, one in the afternoon, and one in the evening, resulting in a total of 28 prescribed times for each participant. Table 1 outlines the time schedules for each of the four regimen conditions. The time schedule for Regimen Same Regular was identical to a schedule used by Rendell and Thomson (1993) on a similar time-logging task. Besides the time-logging device used, there was one difference between Regimen Same Regular and the regimen used by Rendell and Thomson: Regimen Same Regular had an additional number task (see Procedures).

There were two dependent variables: number of *on-time entries* and number of *missed entries*. These measures and the scoring procedures were developed by Rendell and Thomson (1993) for a similar task. Each entry (date and time) was classi-
regimen conditions were given an instruction sheet that had an explanation of the task, prescribed times, and starting and finishing days on one side. On the other side of the instruction sheet was a diagram of the Organizer, highlighting the key press sequence required to enter the date and time. All regimen conditions had the following specific task instructions:

- You are required to enter the current time and date using the "time-stamp" key at, or as soon as possible to, the times below, for seven days. If you forget at any of the prescribed times, then enter the information as soon as you remember.

Regimen Same Regular had the extra task of entering a number at each prescribed time, and involved the following additional instructions:

At each prescribed time you are to enter a number from 1 to 5. The number you are to enter is one more than the number that was entered at the last prescribed time: only go up to 5—after 5, go back to 1 and repeat pattern. For the first entry only, start with 3. Please do not try and check previous entries or write out the numbers, as the task is to remember what number you entered last.

Entering the date and time into the Organizer's memory in all regimen conditions, except Regimen Same Regular, required pressing the following five keys in order: on, memo, shift, time-stamp and enter. In Regimen Same Regular, an extra key press was required—the numerical key, before pressing the shift key in the earlier sequence. When delivering the Organizer, the experimenter demonstrated its operation and went through the instruction sheet with the participant. In this training session, participants were required to practice the key pressing sequence and show that they were competent using the Organizer and logging the time. Participants were shown the display that indicated a successful entry. They were instructed to make sure during the time-logging task that they checked the display to see if an entry was successful. In addition, participants were instructed that if they ever pressed the wrong key or they were not sure an entry was successful, they were to simply repeat the key pressing sequence. They were encouraged to keep the Organizer with them throughout the seven days of the task, including when they were away from home. Participants were also shown the key that would display the current time, so they could use the Organizer to check the time. They were not given instructions about writing notes or about using external aids. In each regimen, participants were given an instruction sheet detailing the prescribed times.

Results

Responses on the time-logging task were nearly all able to be classified as on time (within 5 minutes of the prescribed time),

### Table 1. Time Schedules for Each Regimen Condition in the Time-Logging Task (Experiment 1)

<table>
<thead>
<tr>
<th>Regimen</th>
<th>Prescribed Times</th>
</tr>
</thead>
<tbody>
<tr>
<td>Same Regular*</td>
<td>8:00 a.m.</td>
</tr>
<tr>
<td>Same Irregular</td>
<td>8:20 a.m.</td>
</tr>
<tr>
<td>Different Regular*</td>
<td>base times</td>
</tr>
<tr>
<td>Different Irregular</td>
<td>day 1</td>
</tr>
<tr>
<td></td>
<td>8:25 a.m.</td>
</tr>
<tr>
<td></td>
<td>1:35 p.m.</td>
</tr>
<tr>
<td></td>
<td>5:25 p.m.</td>
</tr>
<tr>
<td></td>
<td>9:15 p.m.</td>
</tr>
<tr>
<td></td>
<td>12:00 p.m.</td>
</tr>
<tr>
<td></td>
<td>12:20 p.m.</td>
</tr>
<tr>
<td></td>
<td>12:10 p.m.</td>
</tr>
<tr>
<td></td>
<td>1:45 p.m.</td>
</tr>
<tr>
<td></td>
<td>1:35 p.m.</td>
</tr>
<tr>
<td></td>
<td>4:45 p.m.</td>
</tr>
<tr>
<td></td>
<td>4:00 p.m.</td>
</tr>
<tr>
<td></td>
<td>4:10 p.m.</td>
</tr>
<tr>
<td></td>
<td>4:40 p.m.</td>
</tr>
<tr>
<td></td>
<td>5:05 p.m.</td>
</tr>
<tr>
<td></td>
<td>5:20 p.m.</td>
</tr>
<tr>
<td></td>
<td>5:25 p.m.</td>
</tr>
<tr>
<td></td>
<td>8:00 p.m.</td>
</tr>
<tr>
<td></td>
<td>8:00 p.m.</td>
</tr>
<tr>
<td></td>
<td>8:20 p.m.</td>
</tr>
<tr>
<td></td>
<td>8:40 p.m.</td>
</tr>
<tr>
<td></td>
<td>8:45 a.m.</td>
</tr>
<tr>
<td></td>
<td>12:25 p.m.</td>
</tr>
<tr>
<td></td>
<td>11:50 a.m.</td>
</tr>
<tr>
<td></td>
<td>5:20 p.m.</td>
</tr>
<tr>
<td></td>
<td>9:25 p.m.</td>
</tr>
</tbody>
</table>

*The time schedule for Same Regular was identical to a schedule used by Rendell and Thomson (1993) on a similar time-logging task.

*Different Regular had different times each day with a regular pattern resembling a bus time table. Each day participants had the same base times and for each day the target times were a set period after or before the same base times. In order of days, the changes to the base times were +15, +45, -15, +30, +15, +45, -15 minutes.
late, or missed entries (failure to log the time at all for a prescribed time). There were very few early entries, and there was an absence of repeated entries (a repeated date and time entry after some delay for the same prescribed time). Therefore, there were just two sorts of errors on the task—late and missed entries. The reported analyses focus on the number of on-time entries and the number of missed entries, but not the number of late entries. Analyses of late responses is not reported in full. Rendell and Thomson (1993) did detail findings in relation to lateness of response. They found the various measures of lateness revealed age-related patterns that were equivalent to those revealed by analysis of on-time responses. Time-deviation score was one of the measures reported by Rendell and Thomson. The findings on this measure for this study are outlined, but the analyses are not reported in full. Analyses of time-deviation scores measured in this study revealed an equivalent age-related pattern of results to the analysis of the number of on-time results.

Separate Age Group (3) × Regimen (4) between groups analyses of variance were applied to the number of on-time entries and the number of missed entries, but not the number of late entries. Analyses of late responses is not reported in full. Analyses of time-deviation scores confirmed the analyses of the measures of on-time and missed entries. The analysis will not be reported in full. The key aspect to report is that the superior performance by the 60s and 80+ compared with the 20s age group and the lack of age differences in performance between the 60s and 80+ age groups with the number of on-time entries was observed with time-deviation scores at each time of day and each day of the week of the time-logging task. This consistent pattern is illustrated in Figures 1 and 2 for day of week and time of day respectively. It is stressed that the older adults were superior at each of the four prescribed times per day, for each of the seven days of the week. There was some effect of time of day and day of the week, but these effects were relatively small and largely limited to the 20s age group.

The reported occupations by participants in the 60s age group were varied: either work, home duties, or retired (n = 20, 17, & 43, respectively). The reported occupations at each of the other levels of age group were not varied: the 80+ age group reported occupation as home duties or retired, and the 20s age group were all university students. One-way analyses of variance of the 60s age group only, with the between-subjects vari-

### Table 2. The Number of On-Time (± 5 minutes) Entries (out of 28) for Each Regimen Condition and Age Group (Experiment 1)

<table>
<thead>
<tr>
<th>Regimen</th>
<th>Age group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>20s 60s 80+</td>
</tr>
<tr>
<td>Same Regular</td>
<td>M 9.03 17.75 19.80</td>
</tr>
<tr>
<td>SD</td>
<td>5.40 7.30 6.82</td>
</tr>
<tr>
<td>Same Irregular</td>
<td>M 11.17 19.00 18.90</td>
</tr>
<tr>
<td>SD</td>
<td>5.97 6.33 6.01</td>
</tr>
<tr>
<td>Different Regular</td>
<td>M 10.13 20.00 20.20</td>
</tr>
<tr>
<td>SD</td>
<td>7.71 5.87 5.55</td>
</tr>
<tr>
<td>Different Irregular</td>
<td>M 11.40 19.75 19.05</td>
</tr>
<tr>
<td>SD</td>
<td>6.45 5.68 6.74</td>
</tr>
</tbody>
</table>

n = 30 for each regimen condition for 20s; n = 20 for each regimen condition for both 60s & 80+ age groups.

### Table 3. The Number of Missed Entries (out of 28) for Each Regimen Condition and Age Group (Experiment 1)

<table>
<thead>
<tr>
<th>Regimen</th>
<th>Age group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>20s 60s 80+</td>
</tr>
<tr>
<td>Same Regular</td>
<td>M 4.60 1.50 1.95</td>
</tr>
<tr>
<td>SD</td>
<td>4.99 1.79 3.19</td>
</tr>
<tr>
<td>Same Irregular</td>
<td>M 5.67 0.95 1.85</td>
</tr>
<tr>
<td>SD</td>
<td>5.44 1.28 2.37</td>
</tr>
<tr>
<td>Different Regular</td>
<td>M 5.40 0.55 1.75</td>
</tr>
<tr>
<td>SD</td>
<td>5.51 0.83 2.55</td>
</tr>
<tr>
<td>Different Irregular</td>
<td>M 3.73 0.55 2.35</td>
</tr>
<tr>
<td>SD</td>
<td>4.09 0.89 3.48</td>
</tr>
</tbody>
</table>

n = 30 for each regimen condition for 20s; n = 20 for each regimen condition for both 60s & 80+ age groups.

![Figure 1. The mean time-deviation score as a function of age group and day of week (Experiment 1).](http://psychsocgerontology.oxfordjournals.org/Downloaded from http://psychsocgerontology.oxfordjournals.org)
able of occupation (work, home duties, or retired), revealed that occupation was not a significant factor determining the number of on-time entries, $F(2,77) = 0.98, MSE = 39.27, p > .05,$ or the number of missed entries, $F(2,77) = 0.03, MSE = 1.71, p > .05.$

After completing the time-logging task, participants were asked some questions. Consistent with age differences in performance, the 20s reported that the time-logging task was difficult more frequently than the 60s or 80+ age groups. Nearly all the participants in the 20s and 60s, and about 80% of the 80+ age group, indicated they found using the Sharp Organizer in the time-logging task either "very easy" or "quite easy."

A larger percentage of the 20s (50%) answered "yes" to "Did you have any difficulties in keeping the Sharp Organizer with you?" than the 60s or 80+ age groups (11% & 20% respectively). An Age Group (3) × Response (2) chi-square analysis revealed that the frequencies of "yes" and "no" responses as a function of age group differed significantly from those expected by chance, $\chi^2 (2, N = 273) = 39.66, p < .001.$ Further Age Group (2) × Response (2) chi-square analyses revealed the 20s did differ significantly from the 60s, $\chi^2 (1, N = 193) = 31.97, p < .001,$ and the 20s differed significantly from the 80+, $\chi^2 (1, N = 193) = 18.46, p < .001,$ but the 60s did not differ significantly from the 80+, $\chi^2 (1, N = 160) = 2.32, p > .05.$

Most participants at all ages reported that they did not develop a plan. The percentage of "yes" responses to whether they developed a plan or strategy to do the time-logging task was 37%, 20%, and 20% for the 20s, 60s, and 80+ age groups respectively. An Age Group (3) × Response (2) chi-square analysis revealed that the frequencies of "yes" and "no" responses as a function of age group differed significantly from those expected by chance, $\chi^2 (2, N = 273) = 9.88, p < .01.$ Comparisons between the 20s and 60s age groups (or the 20s and 80+ age groups) with an Age Group (2) × Response (2) chi-square analysis revealed that the 20s differed significantly from both the 60s and 80+ age groups, $\chi^2 (1, N = 193) = 6.57, p < .05.$

**DISCUSSION**

Older adults were consistently superior to young adults on the time-logging task. The different regimens did not influence the direction or magnitude of the age differences. The age trend found on previous naturalistic PM tasks with simple and regular schedules was confirmed with more complex schedules. The age-complexity effect found with laboratory PM tasks (Einstein et al., 1992) was not found with this naturalistic PM task. This study confirmed the two aspects of the age trend on previous naturalistic studies (Rendell & Thomson, 1993): superior performance by older adults compared with young adults, and no age differences between younger and older elderly participants (60s and 80+ years). The findings of the present study suggest that the age-related trend on naturalistic PM task is robust and is not limited to tasks with simple and regular schedules.

The age trends and the level of performance by participants were similar to that found by Rendell and Thomson (1993) with simpler schedules and a less demanding time-logging task. The mean percentage of on-time entries by participants on a regular four-times-a-day schedule found by Rendell and Thomson was 44%, 67%, and 75% for the 20s, 60s, and 80+ age groups respectively. This is a very similar pattern to the mean percentage of on-time entries found in this study: 37%, 68%, and 70% for the 20s, 60s, and 80+ age groups. In both studies, older adults, on average, were on time about two-thirds of the time, and young adults, on average, were on time just over one-third of the time. It was expected that the time-logging task, in this study, involving a sequence of several key presses would increase the demands on the retrospective memory components of the task and reduce the age superiority. However, it is questionable whether demands on retrospective memory components were increased with the different time-logging device. Most participants, including the older adults, indicated finding using the Organizer and logging the time as either quite easy or very easy.

The additional number task, with the condition Regimen Same Regular, yielded no age differences on the accuracy of entering the appropriate number (85%, 86%, and 82% for 20s, 60s, and 80+ age groups respectively), and had no effect on the time-logging performance. The performance of each age group for this regimen replicated the findings of Rendell and Thomson (1993) on a time-logging task with an identical time schedule that had no additional task. Two conclusions are possible: either increasing the demands on the retrospective component had no effect, or demands on the retrospective memory component were not increased. It is possible that participants were able to work out the pattern of numbers and not have to rely on remembering what number they entered at the last prescribed time. However, there was further, indirect evidence that participants in this task had little difficulty with the retrospective memory component of the PM tasks; consistent with Rendell and Thomson, in this study there was an absence of repeated time-logging entries. This is the equivalent of remembering whether one took medication at the last prescribed time, and avoiding double dosing. This finding is contrary to the predictions based on Einstein and McDaniel’s (1992) componential analysis that age differences on PM tasks occur when there are demands on the retrospective component of the PM task. However, in Einstein and McDaniel’s analysis of PM tasks, they focused on the PM component of remembering to carry out the task and the retrospective memory component of remembering the content of the task. A third component, remembering afterwards whether one has carried out a task, is presumably a retrospective memory component. This third component is a critical aspect of many everyday tasks such as taking medication where repeated doses could be risky as well as wasteful. Interestingly, this component of PM caused no trouble for the young and old participants on a task simulating taking medications.

The use of external cues, in particular conjunction cues, was not directly investigated in Experiment 1. However, the indirect
evidence from Experiment 1 suggests that the use of external cues does not explain the superior performance of older adults on naturalistic PM tasks. It was expected that regimen conditions involving irregular and changing times would reduce the opportunity for conjunction cues. Contrary to expectations, the superior performance of older adults was maintained with schedules that seemingly minimized the opportunity to use conjunction cues. Participants’ use of external aids was assessed indirectly when participants were asked to report whether they used a plan or strategy in completing the task. Most participants of all ages reported they did not use a plan or strategy, and interestingly, fewer older adults than young adults reported using a plan or strategy.

EXPERIMENT 2

The use of external aids in naturalistic PM tasks has been considered a critical factor in the different age-related trends on naturalistic and laboratory tasks (Kvavilashvili, 1992; McDaniel & Einstein, 1992). However, the findings of Experiment 1 were not consistent with the view that the superior performance of older adults on naturalistic tasks simply reflects the greater use of external aids and conjunction cues by older participants. In Experiment 1, regimens of the naturalistic PM task minimizing opportunity for use of conjunction cues did not reduce the accuracy of older adults on the task. However, some caution is needed, as it is difficult to control for the use of external aids by trying to eliminate the use or opportunity to use them. Experiment 2 examined the alternate side: whether maximizing opportunity for use of conjunction cues or access to an external cue for all participants would reduce the superiority of older adults.

Experiment 2 extended the time-logging task in the first experiment by adding two new conditions involving external cues: alarm and choice. These two regimen conditions involved two new groups of participants, but just in the 20s and 60s age groups. The results from the two regimens were contrasted with the results from the four regimens in Experiment 1. The alarm condition involved setting an alarm to go off at the four set times a day, over the 7 days. The Organizer was preset by the experimenter to sound at each of the four prescribed times a day, over the week of time-logging. The alarm facility would not have been apparent to other participants in the non-alarm regimen conditions. The procedure to set the alarm was complex, requiring access to the manual. The alarm sounded for approximately 16 seconds, and the Organizer needed to be fairly near the participant to be heard clearly.

METHOD

Participants

The sample consisted of 100 adults selected in the same way and with the same characteristics as in Experiment 1. There were equal numbers of men and women in both regimen conditions for the 20s age group and there were 8 men and 12 women in both regimen conditions for the 60s age group. Before participating, all participants completed a written consent form that contained a brief explanation of the research project.

Design

The design was similar to Experiment 1 with the same dependent measures of on-time entries (within 5 minutes of prescribed time) and missed entries, and similar between-subjects variables: age group and regimen condition. The variation was that age group only included 20s and 60s age groups, and there were two regimen conditions: Regimen Alarm and Regimen Choice.

An alarm was set at each prescribed time in Regimen Alarm, and this condition had the same times as Regimen Same Irregular: 8:20 a.m., 12:35 p.m., 4:10 p.m., and 8:55 p.m. In Regimen Choice, participants selected one time from each of the following ranges: 8–10 a.m., 12–2 p.m., 4–6 p.m., and 8–10 p.m. Once selected, the times were kept the same throughout the week of time-logging.

Materials

As in Experiment 1, Sharp Organizers were used as the time-logging device. The alarm function of the Organizers was used in Regimen Alarm. The Organizer was preset by the experimenter to sound at each of the four prescribed times a day, over the week of time-logging task. The alarm facility would not have been apparent to other participants in the non-alarm regimen conditions. The procedure to set the alarm was complex, requiring access to the manual. The alarm sounded for approximately 16 seconds, and the Organizer needed to be fairly near the participant to be heard clearly.

Procedure

The procedures followed those used in Experiment 1. Participants were given instruction sheets with the same basic instructions as in Experiment 1. In Regimen Alarm condition the instruction sheet included details of the alarm and the following additional instructions:

The Organizer has been set so an alarm will sound at the prescribed times. The device will need to be on your person or close at hand for you to hear the alarm.

In Regimen Choice condition, participants chose their set times before commencing the task, and the instruction sheet had the following additional instructions concerning the choice of times:

You must select four times within the above time ranges before commencing. These times will be kept the same over the 7 days. Select times that you think will give you the best chance of remembering. Try to select a time that you can connect to some regular event.

The alarm was demonstrated for participants in the Regimen Alarm and the experimenter emphasized the need to have the Organizer close at hand if the alarm was to be heard. Participants in the choice regimen chose the four times during the task briefing, and these times were noted by the experimenter before commencing the task.
RESULTS
Separate 2 × 6 analyses of variance were applied to the number of on-time entries and missed entries. The two factors were between-subjects variables of age group (20s or 60s) and regimen condition (Same Regular, Same Irregular, Different Regular, Different Irregular, Alarm, or Choice). The results for the 20s and 60s participants on the four regimen conditions in Experiment 1 (Same Regular, Same Irregular, Different Regular, or Different Irregular) were included in the analysis. The mean numbers of on-time and missed entries as a function of age group and regimen condition for Experiment 2 are shown in Table 4. The highest possible number of on-time entries was 28, as each participant had four prescribed times a day, for seven days. There was a significant main effect of age group upon number of on-time entries, $F(1,288) = 141.90$, $MSE = 38.63, p < .001$. The 20s had significantly fewer on-time entries ($M = 11.70$) than the 60s age group ($M = 20.43$). There was also a significant main effect of regimen condition, $F(5,288) = 7.36$, $MSE = 38.63, p < .001$. Age and regimen did not significantly interact. Tukey post hoc tests revealed that there were significantly more on-time entries for Regimen Alarm ($M = 19.78$) than for all the other regimen conditions ($M =$ 12.52, 14.30, 14.08, 14.74, & 15.72 for Regimens Same Regular, Same Irregular, Different Regular, Different Irregular, and Choice respectively), but no other comparisons of the mean number of on-time entries for each regimen condition differed significantly.

There was a significant main effect of age group upon number of missed entries, $F(1,288) = 76.45$, $MSE = 15.34, p < .001$. The 20s had significantly more missed entries ($M = 4.88$) than the 60s age group ($M = 0.84$). Regimen condition was not a significant main effect, $F(5,288) = 1.03$, $MSE = 15.34, p > .05$, and did not significantly interact with age group.

DISCUSSION
The findings of Experiment 2, together with the findings of Experiment 1, suggest that the superior performance by older adults on the time-logging task was not simply due to the use of external cues and the opportunity to use conjunction cues. The opportunity to use conjunction cues was minimized by the regimens with irregular and changing time schedules in Experiment 1. In contrast, use of conjunction cues was maximized by Regimen Choice in Experiment 2. Participants' time-logging performance did not differ significantly between Regimen Choice and the irregular and changing regimens. Older adults were consistently superior on the time-logging task with regimens minimizing and maximizing conjunction cue use.

The findings on the alarm condition suggest that the use of external aids does not explain the superior performance by older adults when compared with young adults on naturalistic PM tasks. Regimen Alarm improved the performance of both young and older adults, but did not change the magnitude or direction of age differences on the time-logging task. This suggests that the performance of older age groups on the other conditions was not as ceiling, and that the lack of age differences between the 60s and 80 + age groups was not due to ceiling effects. The alarm condition provided all participants with access to a relatively reliable and effective cue for carrying out the PM task at the prescribed times. However, the effectiveness of the alarm was limited, given that the alarm could only be heard if the Organizer (time-logger) was in close proximity at the prescribed times. Young adults did report having more difficulty keeping the Organizer with them. It should be pointed out that this is a feature of many strategies or external aids in PM. For example, for individuals to use a diary effectively, they need to remember to keep the diary with them and to remember to check entries. This qualifies, but does not undermine, the conclusion that the use of external aids is not critical in age-related trends in naturalistic PM tasks. It should be noted that young adults were consistently inferior at each time of day and each day of the week. It would be expected that the ease or difficulty of keeping the Organizer at hand would vary for different prescribed times.

EXPERIMENT 3
Experiment 3 investigated the age-related trends on laboratory prospective memory (PM) tasks and retrospective memory tasks, with the same participants that completed the naturalistic PM task in Experiments 1 and 2. Within two months of completing Experiment 1 or 2, all the participants were given the same laboratory tasks session. The session involved retrospective memory tasks, recall and recognition tasks, and completion of a questionnaire. Participants were shown a practice list of words followed by a free recall test, then shown a study list of 24 unrelated words and asked to complete a free recall test followed by a recognition test. The session finished with a questionnaire. Embedded in the session were two laboratory PM tasks: the stop-clock task and the note-finish task. At the start of the session, participants were requested to turn off a stop-clock after seven minutes (stop-clock task) and to note the time they finished the questionnaire (note-finish task).

The laboratory PM tasks investigated the event and time-based distinction proposed by Einstein and McDaniel (1990). They argued that age-related decrements are more likely on time-based tasks than event-based tasks, as time-based tasks lack the environmental support of an event cue to perform the prospective action. The stop-clock task had the time of seven minutes from start of session as the target for carrying out the task of stopping the stop-clock. The note-finish task had the event of finishing the questionnaire as the target for carrying out the task of noting the time. Based on Einstein and McDaniel’s (1990) proposal, it was expected that age differences would be larger on the time-based task. The note-finish task closely followed a task by Dobbs and Rule (1987), and

<p>| Table 4. The Number of On-Time (± 5 minutes) and Missed Entries (out of 28) for Each Regimen Condition and Age Group (Experiment 2) |</p>
<table>
<thead>
<tr>
<th>Age group</th>
<th>On-time entries</th>
<th>Missed entries</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>20s</td>
<td>60s</td>
</tr>
<tr>
<td>Regimen</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alarmb</td>
<td>M</td>
<td>16.83</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>5.66</td>
</tr>
<tr>
<td>Choicec</td>
<td>M</td>
<td>11.63</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>7.74</td>
</tr>
</tbody>
</table>

*a = 30 for each regimen condition for 20s; n = 20 for each regimen condition for 60s.

b In Regimen Alarm, time schedule was same as the Regimen Same Irregular in Experiment 1. Alarm set on Organizer for each of the four daily times.

c In Regimen Choice, the participants chose the four times for each day. The times were the same throughout the week.
used their strict and lenient definition of performance that attempted to separate the retrospective and prospective memory components of the task. The strict definition focused on the content of the task that required the finish time to be noted in a specific position on the questionnaire. The lenient definition focused on the PM component by scoring whether the finish time was noted, and ignoring the accuracy of where the finish time was located. It was expected that age differences would be larger on the strict definition where the demands on the retrospective component were greater. This expectation was based on the proposal by Einstein and associates (1992), that age differences in event-based tasks occur when the demands on the retrospective component of the PM tasks are increased.

METHOD

Participants

All except five of the 20s age group (n = 175), all the 60s age group (n = 120), and all the 80+ age group (n = 80) from Experiment 1 and 2 participated. The results of 50 participants on the recognition test were lost due to breakdown of the computer equipment during testing leaving 126, 120, and 79 participants for the 20s, 60s, and 80+ age groups respectively.

Design

The retrospective memory tasks had the independent variable of age group (20s, 60s, or 80+). The dependent variables were the number of words recalled from the study list, and the number of words recognized from it.

The laboratory PM tasks had the independent variable of age group (20s, 60s, or 80+). In the stop-clock task (time-based task), the dependent measures were the frequency of occasions on which the stop-clock was stopped early, on time, late, or not stopped at all. In the note-finish task (event-based task) the dependent measure was the frequency with which participants noted time correctly, incorrectly, or not at all. The laboratory PM tasks were the same for all participants. Sex of participants was not included as an independent variable, as preliminary analyses revealed that sex was not a significant effect for any of the tasks in the laboratory session.

Materials

The retrospective memory tasks involved a practice list of 12 proper nouns and one study list of 24 words of mixed word frequency. The study list words had either one or two syllables, and were either nouns or adjectives. In the study list, 12 words were classified high frequency (100-200 per million) and 12 were classified low frequency (1 per million). The words selected were rated in these frequency ranges by both Thorndike and Lorge (1944) and Kucera and Francis (1967) word frequency counts. The words were also selected to avoid associations between them. The recognition test list of 48 words included all 24 words from the study list and 24 distracters. The distracters were selected according to the same criteria as the study list. Word frequency was not included as an independent variable. Preliminary analysis revealed that there was no interaction between age group and word frequency on the retrospective memory measures.

The practice list, study list, and recognition test list were presented on a Sharp laptop computer (PC6741/CE671B) with a monochrome screen that was 204 mm × 154 mm. The words were displayed in lowercase with letters about 1.5 cm high. The laboratory PM tasks included a desktop stop-clock with a real-time clock attached. Both units had digital displays with numbers about 2 cm high. The questionnaire in the laboratory tasks session contained questions eliciting basic biographical details from the participant, as well as a couple of questions on how they found remembering names and remembering to do things like keep appointments and take medication in their daily life.

Procedures

The laboratory tasks session was about 20 minutes long and included the retrospective memory tasks, a questionnaire, and two laboratory PM tasks. The six were, in order: presentation of practice list of 12 words, free recall of practice list, presentation of study list of 24 words, free recall of study list, recognition test of study list, and the questionnaire that was unrelated to the word list tasks. Instructions were given at the start of the session for the two PM tasks. Consistent with the naturalistic PM task in Experiments 1 and 2, there was no cover story given for the laboratory PM tasks. No further reference was made to the PM tasks during the session.

Participants were tested individually in a quiet place in their homes or in a laboratory room at the university. Participants sat at a table with the laptop computer immediately in front of them. The stop-clock and real-time clock were just to the left of the computer, within easy reach, and visible throughout the session. The session began with the following preliminary instructions:

There will be three word memory tasks. You will be shown a list of words and you will be asked to see how many words you can remember. At the end of the three word tasks you will get this general questionnaire sheet to fill in which does not relate to the word memory tasks. Could you please put the time you finish the questionnaire on the top left hand corner of the sheet (the note-finish task). The time you write down should be the time shown on this clock (the real-time clock). This other clock is a stop-clock. I am going to start this clock and could you stop it after seven minutes. That is, could you simply push this button to stop this clock in seven minutes time when a seven will show here on the display (the stop-clock task). The first list will be a practice list. Words will be shown one at a time on the screen. Try to remember as many as you can.

Participants were then shown the practice list at a rate of 3.3 seconds for each word. Immediately after the last word on the practice list was shown, they were asked to, “Write down all the words you can remember, in any order.” The instructions and procedure were then repeated for the study list. The recall tests of the practice and study list were untimed. After the participants finished recalling words from the study list, the participants were given the recognition test. Participants were told they would be shown words one at a time, and that some of the words appeared on the study list and some did not. They were asked to press the key labeled “yes” if they remembered seeing the word, or press the key labeled “no” if they did not recognize the word as occurring before. The recognition test list was shown at the same rate as the study lists, 3.3 seconds for each word. After the recognition test was finished, participants were given the questionnaire sheet to complete and were told this was the last activity.
The target time for the stop-clock task, seven minutes, occurred after participants spent some time recalling words from the study list. The recognition task was not started until the seven minutes had clearly elapsed. Participants finished completing the questionnaire around 20 minutes from the start of the session.

RESULTS

Retrospective Memory Tasks

Table 5 shows the mean number of words recalled and recognized (hit rates) out of the 24 presented. Separate one-way analyses of variance were conducted on the number of words recalled and number of words recognized. The independent variable was the between-subjects variable of age group. Age group was a significant main effect for the number of words recalled, F(2,372) = 102.31, MSE = 7.01, p < .001, and the number of words recognized, F(2,322) = 40.76, MSE = 13.43, p < .001. Tukey post hoc tests revealed that the 20s recalled and recognized significantly more words than the 60s or 80+ age groups, and the 60s recalled and recognized significantly more than the 80+ age group. Age group was not a significant effect for the number of false positives on the recognition task, F(2,322) = 2.10, p > .05. Further Age Group (2) X Response (4) chi-square analysis revealed that the frequencies as a function of age group differed significantly from those expected by chance, \( \chi^2 (6, N = 375) = 63.49, p < .001 \). Table 6 shows the percentage of each response on the stop-clock task. Across the three age groups, as age increases, the percentage of on-time responses decreases, and the percentage of missed responses (failure to stop the clock) increases.

Laboratory PM Task: Stop-clock

Participants' responses to the stop-clock task were classified according to whether the stop-clock was stopped as requested at the start of the session, and how close to the target time participants were if the clock was stopped. Table 6 shows the percentage of participants early (11–60 seconds), on time (±10 seconds), late (11–60 seconds), very late (>60 seconds), and missed (failing to stop the stop-clock) as a function of age group. The clock was stopped more than 10 seconds before the target time by only three participants, each from the 60s age group. These early participants were all within 60 seconds of the target time. Interestingly, the on-time category only included four participants who stopped the clock before the target time: one each from the 20s and 80+ groups, and two from the 60s group.

Chi-square analysis was applied to the frequencies of responses as a function of age group and response category. The response categories of early and late by 11 to 60 seconds were combined, resulting in four response categories. An Age Group (3) X Response (4) chi-square analysis revealed that the frequencies as a function of age group differed significantly from those expected by chance, \( \chi^2 (6, N = 375) = 63.49, p < .001 \). Further Age Group (2) X Response (4) analyses revealed that the 20s did differ significantly from the 60s, \( \chi^2 (3, N = 295) = 20.37, p < .001 \), the 20s differed significantly from the 80+ age group, \( \chi^2 (3, N = 255) = 58.63, p < .001 \), and the 60s differed significantly from the 80+ age group, \( \chi^2 (3, N = 200) = 16.33, p < .01 \). Table 6 shows the age-related decline on the stop-clock task. Across the three age groups, as age increases, the percentage of on-time responses decreases, and the percentage of missed responses (failure to stop the clock) increases.

Laboratory PM Task: Note-finish

Table 7 shows the percentage of responses on the note-finish task. Responses were classified according to a strict and a more lenient definition closely following the distinction made by Dobbs and Rule (1987) in a version of this task. In the strict classification, a correct response was noting the correct time in the correct position. This required participants to note the time they finished the questionnaire on the top left hand corner of the questionnaire sheet. In the lenient definition, responses were classified without regard for position as: correct time, wrong time, or missed. Correct time was noting the time the questionnaire was finished, and wrong time was noting the time of some other event. In both classifications, a missed response indicated a failure to note any time. Table 7 shows the age-related decline on the note-finish task. Across the three age groups, as age increased, the percentage of
correct responses decreased and the percentage of missed responses (failure to note any time) increased. The age-related decline was confirmed by separate Age Group (3) × Response (3) chi-square analyses that revealed that the frequencies of responses differed significantly from those expected by chance for both the strict and lenient classifications, χ²(4, N = 375) > 74, both ps < .001. Further Age Group (2) × Response (3) analyses revealed the following age differences. The 20s differed significantly from the 60s age groups, for both the strict and lenient classifications, χ²(2, N = 295) > 22, both ps < .001. The 20s differed significantly from the 80+ age group, for both the strict and lenient classifications, χ²(2, N = 255) > 72, both ps < .001. The 60s differed significantly from the 80+ age group, for both the strict and lenient classifications, χ²(2, N = 200) > 18, both ps < .001.

**Discussion**

The typical age-related decline was found on the retrospective memory tasks. There was a significant age-related decline across the three age groups; 20s, 60s, and 80+ years. The age-related trends on both laboratory PM tasks paralleled the trends on the retrospective memory tasks. The findings differed from those trends found with the same participants on the naturalistic PM task in Experiments 1 and 2 in two key ways. Older adults were superior to young adults on the naturalistic PM task and inferior on the laboratory PM tasks; the 80+ were inferior to the 60s age group on the laboratory PM tasks, but the older age groups did not differ on the naturalistic PM task.

The findings provided no evidence to support Einstein and McDaniel’s (1990) proposal of a distinction in the age-related trends on event versus time-based PM tasks. On both the time-based and event-based laboratory PM tasks, there was a significant and substantial age-related decline. However, there is some doubt about whether the two tasks precisely tested the distinction. There was an environmental cue present in both tasks. The stop-clock and real-time clock were in close proximity and in full view throughout the task. In addition, both tasks had time monitoring requirements. The note-finish task had an event target, finishing the questionnaire, but the content of the task involved noting the time. There seemed to be evidence that an age-related decline was found on both retrospective and prospective memory components of the note-finish task. This conclusion is limited to the extent that the strict and lenient definitions separately tested the demands on the retrospective and prospective memory components. It appears that the distinctions between time and event-based tasks, and the attempts to identify retrospective and prospective memory components of PM tasks, have not been helpful in identifying differences in age-related trends on the PM tasks.

**General Discussion**

The major finding of the three experiments is the paradoxical age-related differences between naturalistic and laboratory PM tasks. On the laboratory PM tasks and the retrospective memory tasks, there was an age-related decline, with young adults performing significantly better than the 60s age group who, in turn, performed significantly better than the 80+ age group. In contrast, on the naturalistic PM task, there were no differences between the two older age groups, and both older age groups were significantly and consistently more accurate than the young adults. This confirms the paradoxical age-related trends observed in the review of naturalistic and laboratory PM studies in the introduction. Unlike previous studies, however, this contrast was observed within the same study, and with the same participants. In addition, despite showing the typical age-related decline on retrospective memory tasks, the older adults were consistently superior in their performance on the naturalistic PM task across every variation of the task that included various attempts at making the task more demanding. Interestingly, making the task demanding seems to be the feature that had produced robust age-related declines in performance on laboratory PM tasks (Einstein et al., 1992; Mäntylä, 1994; Park et al., 1997). However, the study has not revealed any satisfying explanation of the paradoxical age-related trends. Instead, the study renders several possible explanations as somewhat unlikely: motivation, level of activity, use of external aids, use of conjunction cues, and the distinction between time-based and event-based tasks. It does not seem to be simply an issue of real versus artificial tasks, as reliable age differences have been found with a variety of naturalistic and laboratory retrospective memory tasks (Light, 1991).

Motivation seems an unlikely explanation for the superior performance of older adults on the naturalistic PM task. The contrasting age-related effects were observed with the same participants. There is no reason to suggest that older adults had different commitments to the different tasks. Even if older adults were more motivated than younger adults on the naturalistic task, it does not diminish the accurate level of performance demonstrated by both young-old and old-old adults on the naturalistic task. Level of activity also seems an unlikely explanation for the paradox. Performance on the naturalistic PM task did not differ among the three occupation categories for the 60s age group: retired, working, or performing home duties. In addition, the general impression of participants was that older adults, who were all living independently, were as active or more active than their younger counterparts. This was consistent with the impressions of participants in previous naturalistic PM studies (Devolder et al., 1990; Moscovitch, 1982).

Interestingly, both motivation and level of activity are often suggested as possible explanations when people are told about the superior performance of older adults on naturalistic PM tasks. A common response is an inclination to explain away the age-related improvement. In addition to suggestions related to motivation, there is the suggestion that young adults' poor performance was due to the difficulty in keeping the time-logging device with them. The fact that young adults might have failed to organize their day to ensure the timer was with them does not dispose of the matter. That aspect of organization is a critical component of memory and, in particular, remembering to do things in daily life. Taking medication often requires individuals to remember to keep a medication container with them. In addition, the paradoxical age-related trends on naturalistic and laboratory PM tasks have two aspects. In contrast to laboratory PM tasks, there is the inferior performance on the naturalistic PM task by young adults compared with older adults, and there is also the lack of age differences on naturalistic PM task between the young-old and the old-old. Explanations that involve the suggestion that the young adults have not given their best performance do not deal with the lack of age differences between the young-old and old-old. An interesting issue that should not be overlooked for future research is an explanation for the consistently inferior performance of young adults on the
naturalistic PM task in this study. Consistent with similar studies, the young adults in this study were all university students. A possible question for future research is whether the relatively inaccurate performance on the naturalistic task is a general feature of the age group or is limited to undergraduate university students. Future research might examine performance on naturalistic PM tasks of young adults in various occupations. It must be noted that some of the older adults in the study were once university students, so if the inferior performance by the young is a feature of university students then an interesting question deserving further research is what changes to produce the accurate performance by older adults on the naturalistic PM task.

In the introduction it was noted that the opportunity to use external aids, or the lack of control over use of external aids, in naturalistic PM studies has been suggested as an explanation for the paradoxical findings on naturalistic and laboratory tasks (Kvavilashvili, 1992; McDaniel & Einstein, 1992). In three areas, the findings of the present study were not consistent with this view: the variations of regimens in the naturalistic task in Experiment 1 and 2, the reported use of external aids on the naturalistic task, and the presence of an external cue in the laboratory PM task. The age-related trends on the naturalistic PM tasks were constant over the variations of the task designed to minimize and maximize the opportunity for the use of conjunction cues and the use of external aids. The naturalistic PM task in this study involved logging of time, and the regimens were varied to minimize and maximize opportunity to use both conjunction and external cues. Conjunction cues involved connecting the prescribed PM task to other regular events in the daily life of participants. There were no differences in performance on the time-logging task with the various regimens of constant or changing times, and regular or irregular times. The only regimen variation that had a significant effect was the setting of the alarm on the Organizer (time-logging device). This improved the performance of both the young and older adults, but did not change the direction or the magnitude of the age-related difference on the task.

If the use of external aids was a critical factor, then an increased use of external aids by older adults would be expected on the naturalistic PM task. However, fewer older adults than younger adults reported that they used a plan or strategy in completing the task. This is consistent with the findings by Rendell and Thomson (1993), that older adults reported less use of external aids than their younger counterparts during a naturalistic PM task.

The presence of an external cue in the laboratory PM tasks provides further support for the view that the opportunity for the use of external aids or external cues does not seem to be the critical factor in difference between laboratory and naturalistic PM tasks. There does not seem to be an absence versus a presence of external cues distinction between the laboratory versus the naturalistic PM tasks. An external cue was present in the two laboratory PM tasks in this study—the note-finish and the stop-clock task. In both tasks, participants had a stop-clock and real-time clock in full view and close proximity throughout the task. Presumably, these would have provided strong external cues throughout the task.

The distinction between time-based and event-based tasks proposed by Einstein and McDaniel (1990) and based on the environmental support model of aging and memory (Craik, 1986) was not helpful in explaining the paradox. As described in the discussion of Experiment 3 results, an age-related decline was found on both a time-based and an event-based laboratory PM task. Some doubt was expressed as to whether the tasks were distinctly time-based and event-based. However, the naturalistic PM task in the present study required logging time at set times, and older adults were consistently superior to young adults on this task. This result is contrary to the prediction by Einstein and McDaniel (1990) of a greater likelihood of age differences on time-based compared with event-based PM tasks.

The findings on the naturalistic PM task also seem contrary to the environmental support model, given that older adults were consistently superior on the naturalistic PM task with regimens minimizing and maximizing the opportunity to use conjunction cues. However, it is possible that the model is useful in explaining the paradoxical findings on laboratory and naturalistic PM tasks. A key distinction between the two is in the different background activities. The naturalistic task is embedded in the everyday life of participants, where they have the opportunity to use the organizational framework provided by the structured pattern of daily activities inherent in living independently in a familiar environment. The content of the daily routine, with specific events providing conjunction cues, does not seem to be a factor in an explanation of the paradox. However, it is speculated that older adults may have a more predictable and consistent pattern of daily activities that provides an overall framework to remember to execute actions. Just as increasing organization of to-be-remembered material helps retrospective memory, an organized or structured pattern of daily activities may facilitate PM. If the support provided by the structure of daily life does explain the paradoxical findings, then this support must be robust enough to support irregular as well as routine schedules. Further research is needed for what is currently an open question regarding the possible contribution of differences in the nature of background activities in laboratory and naturalistic PM, in particular the possible contribution of structured patterns of daily life. It also remains a possibility that paradoxical age-related trends are explained, as Craik and Kerr (1996) suggest, by older adults developing a range of compensatory strategies in their everyday life that offset the age-related losses in memory functioning. Being more organized in daily life may be one of those strategies. The relatively accurate performance on a naturalistic PM task was found on a task that simulated taking medication. Older adults may be more experienced with such a task, given that older adults are more likely to be taking medication and are more likely to have complex regimens (Gilbert, Luszcz, & Owen, 1993; Simons et al., 1992). In this study, more older than younger participants reported that they were taking medication during the naturalistic PM task in Experiments 1 and 2.

Prospective memory tasks in laboratory and naturalistic studies differ in the nature of the task, as well as in the nature of background activities. It is speculated that tasks used by laboratory and naturalistic studies may both represent PM tasks in everyday life, but different types of everyday PM tasks. The time-logging task in this study was typical of PM tasks in previous naturalistic studies that required a simple response at prescheduled times over several days. This represents one kind of PM task in everyday life, such as taking medication and keeping appointments, where there is clear and advanced notice of the appointments. The PM task was the focus of the briefing for the naturalistic PM task, giving clear and advanced notice, whereas...
the laboratory PM task was somewhat incidental in the instructions for the laboratory tasks. There is another sort of PM task in everyday life—tasks that crop up unexpectedly during the day. A structured pattern of daily life and familiarity with routine tasks such as taking medication may not be helpful with the crop-up type of PM task. The laboratory PM tasks may be closer to the crop-up type of task than the taking medication or keeping appointment type of tasks. Future research needs to consider the possible contribution of crop-up versus routine tasks to the paradoxical findings on PM tasks. In addition, the extent to which the task is explicit in the instruction warrants further investigation. Another distinction between naturalistic and laboratory PM tasks is the extent of demands on the content of the to-be-remembered task. The PM tasks in naturalistic studies typically require simple actions that place little demand on remembering the content, whereas laboratory PM tasks tend to make some demand on remembering the content of action (Einstein et al., 1992). In everyday life, some PM tasks place little demand on remembering content, but many other PM tasks in daily life involve some content, even if it is just remembering to buy bread, but not milk, on the way home.

It seems likely that an explanation for the paradoxical age-related trends lies in a consideration of task differences. The findings of this study support the view of recent reviewers (Craik & Kerr, 1996; Crowder, 1996; Dobbs & Reeves, 1996; Ellis, 1996) that aging and PM studies are tapping a collection of tasks, rather than one distinct memory process. It would seem that closer analogues between laboratory and naturalistic PM tasks are needed to resolve the paradox. Nevertheless, in the context of an aging society, the findings of the present study have optimistic implications for older adults. The findings suggest that older adults, in their everyday life, are capable of accurate and reliable performance on important tasks such as remembering to keep appointments and to take medication, even with complex regimens. This is despite the well-established finding of an age-related decline on a variety of memory tasks. Furthermore, the findings suggest that this positive expectation holds for independent living adults in their 60s through to those in their 80s.

ACKNOWLEDGMENTS

These studies formed part of a doctoral dissertation undertaken by Peter Rendell at Monash University, Melbourne, Australia. We wish to thank the participants for their assistance. We thank Erik Craik for helpful comments on earlier versions of this article.

Address correspondence to Dr. Peter Rendell, Department of Psychology, Australian Catholic University, P.O. Box 213, Oakleigh, Victoria, Australia, 3166. E-mail: prendell@christ.acu.edu.au

REFERENCES


Received June 24, 1998
Accepted December 18, 1998