Age, Time of Testing, and Proactive Interference

LYNN HASHER, CHRISTIE CHUNG, University of Toronto
CYNTHIA P. MAY, College of Charleston
NATALIE FOONG, University of Toronto

Abstract A four-list version of a release from proactive interference paradigm was used to assess the degree to which older and younger adults tested at optimal and nonoptimal times of day are vulnerable to interference effects in memory, effects that may increase at nonoptimal times. Morning type older adults and Evening type younger adults were tested either early in the morning or late in the afternoon. Standard buildup and release effects were shown for all age groups except for older adults tested in the afternoon; they failed to show release. Recall and intrusion data suggested that older adults are more vulnerable to proactive interference than younger adults and that for older adults at least, interference effects are heightened at nonoptimal times of day. The data are discussed in terms of an inhibitory model of control over the contents of working memory (Hasher, Zacks, & May, 1999).

Circadian rhythms reflect 24-hour cycles of increases and decreases in a range of biological and physiological functions, including body temperature, heart rate, and hormone secretion (e.g., Hrushesky, 1994; Moore-Ede, Sulzman, & Fuller, 1982). Circadian rhythms are also seen in cognitive functioning (e.g., Folkard, 1982). Recent work has shown, however, that the general circadian patterns in cognition are substantially moderated by reliable individual (and related age-group) differences in the degree to which people are alert and aroused early in the day (Morning types) versus the degree to which alertness occurs considerably later in the day (Evening types; see e.g., Anderson, Petros, Beckwith, Mitchell, & Fritz, 1991; Petros, Beckwith, & Anderson, 1990; Yoon, 1997; Yoon, May, & Hasher, 1998).1 In particular, a “synchrony effect” has now been widely reported, with Morning types performing better in the morning than in the afternoon and Evening types showing the reverse pattern on a range of cognitive tasks, including negative priming, false memory, recognition and recall of prose and span materials, categorization, impression formation, judgment and control over distraction (Bodenhausen, 1990; Intons-Peterson, Rocchi, West, McLellan, & Hackney, 1998; Intons-Peterson, West, McLellan, Hackney, & Rocchi, 1999; May, 1999; May & Hasher, 1998; May, Hasher, & Stoltzfus, 1993; Petros et al., 1990; Yoon, 1997).

Many of the tasks showing circadian effects are ones that require inhibitory processes for performance to be at its peak. According to Hasher, Zacks, and May (1999), inhibitory attentional processes regulate the flow of information from both thought and perception by (a) limiting access to consciousness to goal relevant information; (b) deleting irrelevant and no longer relevant information from consciousness; and (c) restraining strong responses so they can be evaluated for appropriateness. Each of these functions has been shown to be more efficient for younger adults than for older adults, and of particular relevance here, more efficient for both age groups when they are tested at optimal as compared to nonoptimal times of day (see e.g., May 1999; May & Hasher, 1998; Yoon et al., 1998).

Here we consider the possibility that the inhibitory control process of deletion plays an important role in a classic forgetting paradigm, namely the buildup and release from proactive interference (PI; see Wickens, 1972). In commonly used versions of this task, participants are exposed to three or more short lists of related

1 Circadian arousal patterns are reliably measured by the Horne-Ostberg Morningness-Eveningness Questionnaire (1976; see Kerkhof, 1984), a paper and pencil inventory that correlates well with physiological measures such as body core temperature, hormone secretion, and pulse rate (Buela-Casal, Caballo, & Garcia-Cueto, 1990; Horne & Ostberg, 1977; Smith, Reilly, & Midkiff, 1989). Considerable work in North America and elsewhere demonstrates that although a high proportion of adults over age 60 (70-75%) are Morning types, relatively few younger adults are (7% on North American college campuses). By contrast, a substantial proportion of college students (35%) are Evening types, with few older adults falling into this category (less than 5%; all proportions based on norms reported in Yoon et al., 1998).
items, each presented for recall after a brief (say 10-20 s) filled interval. A final trial then consists either of additional related items or of items from a distinctively different set of materials (e.g., early lists might consist of furniture words, while the final list consists of food words). The “buildup” of PI is typically measured across the first series of recall trials as a decline in recall. As well, intrusion errors can serve as a measure of PI buildup (see e.g., Chiappe, Hasher, & Siegel, 2000), as individuals who are unsuccessful in deleting prior items may be more likely to mistakenly produce those inappropriate items on subsequent trials, and this may be especially true when later trials contain related items. “Release” from PI is seen either by an improvement in recall from the pre to postswitch lists or by the difference in recall between groups switching to items from a distinctive set versus those remaining with items from the same set.

The deletion function of inhibition plays a major role in determining the amount of PI because it suppresses no longer relevant items from consideration, enabling retrieval to be narrowly focused on only currently relevant items. In the buildup and release task, as in many other tasks (e.g., working memory span; see Chiappe et al., 2000; Lustig, May, & Hasher, 2001; May, Hasher, & Kane, 1998), the items on each successive list become irrelevant immediately after they have been recalled. If deletion is efficient, the items from the just-recalled list will be suppressed, enabling the items on the new list to be the sole focus of attention, both at encoding and at retrieval, thus limiting competition from prior list items and so benefiting recall (Underwood, 1957). If deletion is inefficient, the consequences may include both impaired recall and heightened intrusion rates. As it happens, we know from a series of studies on both explicit and implicit memory tasks that deletion is more efficient for younger than it is for older adults (e.g., Hasher, Zacks, & May, 1999; Zacks & Hasher, 1994) and is more efficient for both groups at their (different) optimal than at their nonoptimal times of day (e.g., May & Hasher, 1998; May et al., 1999). Taken together, this would suggest that the buildup of PI should be greater for older than for younger adults and should be greater at optimal than at nonoptimal times of day, with the buildup measured both as a decline in recall and as an increase in intrusion errors. As well, we would expect to see greater release effects for younger than older adults and greater release for everyone tested at their optimal as compared to their nonoptimal time.

The present study is a test of these predictions comparing Morning type older adults and Evening type younger adults tested early in the morning (8 or 9 a.m.) or late in the afternoon (4 or 5 p.m.) on a task in which they studied and recalled from each of four successive lists comprised of categorized items. This design choice enables us to explore the synchrony effect by which individuals are expected to show greater inhibitory control at their optimal as compared to nonoptimal times of day. Each list consisted of 10 words, with two from each of five categories. The same five categories were used, always with different items, on the first three lists that were studied and recalled. On the fourth list, all participants again received 10 items, this time with two instances selected from each of five new categories. As will be seen, there are indeed differences tied to time of testing and also to the age of the participants that are generally, if not entirely, consistent with our specific predictions.

**Method**

**PARTICIPANTS**

Forty-eight younger adults (ages 18-32 years) and 48 older adults (ages 58-78) participated in this study. Younger adults were undergraduate students from the University of Toronto and the College of Charleston. Older adults were community-dwelling volunteers from participant registries at both institutions. All participants were selected on the basis of their performance on the Horne-Ostberg (1976) MEQ, such that younger adults were Evening types and older adults were Morning types. Younger adults received course credit or $10 upon completion of the present study. Older adults received $10 in remuneration for their participation.

**MATERIALS**

Sixty nouns (6 words in each of 10 categories) were selected and equated for frequency of generation according to category frequency norms (Howard, 1979). Within each category, the six nouns selected represent exemplars that have relatively low frequencies of generation (i.e., the six most common exemplars generated from each category were not selected). Using these items, we created six lists as follows: First, the 10 categories were randomly sorted into two groups of five categories each, Group A (fruit, metal, trees, sports, words), and Group B (animals, cloth, family, occupation, weather). Then, within each group of categories, three 10-item lists were generated by randomly assigning two words from each of the five categories to each list. Order within a list was randomly determined with the constraint that the two category exemplars within a list could not occupy successive list positions. As well, the lists were equated for frequency of occurrence using the Krucera and Francis norms (1960).

Each participant received four lists, with half of each age group and time of testing condition receiving three
A lists followed by a B list, and the other half receiving three B lists followed by an A list. Two randomly
determined orders of presentation were used for the
first three lists, and each order was paired equally often
with one of the three lists of the switched category set,
creating 12 unique orders of lists, each used two times
within each of the four Age x Time of Testing condi-
tion.

PROCEDURE
Half of the participants were tested in the morning
(from 8:00 a.m.-9:15 a.m.; an optimal time for older
adults and a nonoptimal time for college students), and
the other half were tested in the afternoon (from 4:30
p.m.-5:15 p.m.; an optimal time for younger adults and
a nonoptimal time for older adults). Morning type
older adults had MEQ scores ranging from 59 to 81,
with a mean of 67.3 ($SD = 5.71$). Evening type college
students had MEQ scores ranging from 24 to 38, with a
mean of 19.7 ($SD = 6.84$).

Participants completed the memory task first, followed
by a health questionnaire, the Mill Hill vocabulary
scale (Raven, 1965), and the Extended Range Vocabulary Test (ERVT), version three (Educational

The word lists were presented on a computer screen
with participants instructed to focus on the screen
throughout the memory task. The words were shown
one at a time in the middle of the computer screen, at
a rate of 3 s/word. After the presentation of each list, a
brief filler task was administered. A three-digit number
was presented and participants were instructed to
count backwards by threes for 10 seconds. Participants
were then asked to recall all the words that they could
remember, in any order, from the list that they had just
seen. The same procedure was repeated for all four
lists, with no warning of a change in materials prior to
the fourth list. The number of words recalled as well
as the number of intrusions was recorded on a check-
list by the experimenter who was present throughout
the entire experiment.

Results

Participants
Younger adults ($M_{age} = 20.29$ years, $SD = 3.01$, range =
18-32) had a mean score of 19.7 ($SD = 6.84$) on the
ERVT, and of $13.56$ ($SD = 1.97$) on the Mill Hill test. Older adults ($M_{age} = 67.96$, $SD = 4.70$, range = 58-78)
had a significantly higher score on the ERVT ($M = 31.34$,
$SD = 9.43$), $F(1, 94) = 47.89$, $MSE = 67.84$, and also on
the Mill Hill ($M = 15.81$, $SD = 2.54$), $F(1, 94) = 26.00$,
$MSE = 4.67$. There were no main effects or interactions
with testing time and, as a result, within each age
group, those randomly assigned to afternoon versus

morning test times did not differ on age or vocabulary
levels.

Synchrony Effects in Recall: Lists 1-3
To explore overall differences in recall, mean recall
across the first three lists was determined for each par-
ticipant and an Analysis of Variance was performed
with age and time of day (morning versus afternoon
testing) as between S factors (see Figure 1). Overall,
older adults recalled fewer words ($M = 4.94$, $SD = 1.9$)
then younger adults ($M = 6.20$, $SD = 1.84$; $F(1, 92) =
22.81$, $MSE = 4.99$, $p < 0.05$). The basic synchrony
effect can be seen in Figure 1, with Morning type older
adults recalling more in the morning than in the after-
noon and Evening type younger adults recalling more
in the afternoon than in the morning (Age x TOD effect:
$t(1, 92) = 5.77$, $MSE = 4.99$, $p < 0.05$). As can be seen
in Figure 1, the performance of younger and older
adults was very similar in the morning, but the pattern
changed across the day, with a reliable improvement
for younger adults ($t = 2.34$, $p < 0.05$) and an unreliable
decline for older adults ($t = 1.14$). As a result of this
TOD effect, there were reliable differences between
younger and older adults tested in the late afternoon,
$t = 5.15$, $p < .05$. No other effects were significant.

PI Buildup: Recall Across Lists and Intrusion Errors
Recall on Lists 1 to 3 (PI). Performance across the first
three trials is the standard index of the buildup of PI
and is depicted here in Figure 2. A mixed $3 \times 2 \times 2$
ANOVA was conducted on number recalled, with list (1, 2, 3) as a within-subject factor, and age (older, younger) and time of testing (TOD; morning vs. afternoon) as between-subject factors. As is commonly found in this literature (e.g., Wickens, 1972), recall declined across the lists, $F(2, 184) = 37.471$, $MSE = 1.95$, $p < 0.05$, indicating the buildup of PI. No other effects involving List (List x Age, List x TOD, and List x Age x TOD, largest $F < 1$) were significant, suggesting that the declines across lists did not differ across age and between people tested at their optimal and nonoptimal times of day.

**Intrusion Errors.** Intrusion errors can be used as an index of the buildup of PI, particularly when there is, as is the case here, similarity across lists (e.g., Chiappe et al., 2000). For present purposes, we included all items that were produced on a recall test that had not been presented on a particular test trial. Because all participants received items from a distinctive set of categories on List 4, and intrusions were extremely low on this trial (a floor effect), we considered intrusions only from Lists 1-3. Although intrusion rates were low, we did find a reliable increase in intrusions from List 1 to List 3 (overall means of .18, .52, .78, $F(2, 184) = 20.3$, $MSE = 0.43$). The total number of intrusions made by each person was assessed in a $2 \times 2$ between subject ANOVA with age and TOD as factors (see Figure 3). Older adults made more intrusions ($M = 1.98$, $SD = 1.60$) than younger adults, ($M = 0.65$, $SD = 0.91$; $F(1, 92) = 26.12$, $MSE = 1.65$, $p < 0.05$). There was also a reliable Age x TOD interaction, $F(1, 92) = 4.31$, $MSE = 1.63$, $p < 0.05$, with older adults tested in the afternoon making more intrusions ($M = 2.42$, $SD = 1.80$) than those tested in the morning ($M = 1.54$, $SD = 1.29$; $t = 1.94$, $p < 0.058$). By contrast, younger adults produced slightly more intrusions in the morning ($M = .75$) than in the evening ($M = .54$), though this difference was not reliable ($t < 1$), potentially because of floor effects. As will be seen in the discussion, it is possible that the greater tendency of older adults to produce intrusions may have obscured some of the differences in recall that we were expecting.

**Release From PI.** To assess the release effect, we compared recall on Lists 3 and 4 using a $2 \times 2 \times 2$ mixed ANOVA on recall with List (3, 4) as the within-subject factor, and Age (older, younger) and TOD (morning, evening) as the between-subject factors (see Figure 2). As expected, there was an overall release effect, with better recall for items on List 4 than for items on List 3 ($M_{list\ 3} = 4.82$, $SD = 1.86$; $M_{list\ 4} = 5.21$, $SD = 2.02$; $F(1, 92) = 3.36$, $MSE = 2.12$, $p < 0.07$). Younger adults recalled more items across both lists than did older adults, $F(1, 92) = 36.5$, $MSE = 3.79$, $p < .05$. Two second, order interactions are notable: The first was between Age and TOD, $F(1, 92) = 6.93$, $MSE = 3.79$, $p < .05$. As before, performance of young adults tended to improve from morning to afternoon testing times ($t = 1.93$, $p < 0.06$) while that for older adults tended to decline over that time period ($t = 1.79$; $p < 0.08$). The second interaction was a borderline effect for the List x Age interaction, $F(1, 92) = 3.36$, $MSE = 2.12$, $p < 0.07$. Post hoc tests showed that release was reliable for young adults,
This latter finding is clear evidence of the greater disruptive impact the past can have on older than on younger adults. However, an inspection of Figure 2 suggests the presence of the three-way interaction between age, list, and time of testing, such that three of four groups show release (both young groups along with older adults tested in the morning) while one group (older adults tested in the afternoon) does not. This interaction was not significant ($F < 1$): Older adults did not show reliable release whether tested in the morning or in the afternoon, $p$s > .20.

**Discussion**

The present investigation examined younger and older adults’ performance in a classic PI paradigm at peak and off-peak times of day, with the assumption that PI buildup and release are determined, at least in part, by the efficiency of inhibitory mechanisms that operate to delete obsolete and irrelevant information from working memory. Previous research has demonstrated that the deletion function of inhibition is more efficient for younger than older adults, and for individuals tested at peak relative to off-peak times of day (e.g., Hasher et al., 1999; May & Hasher, 1998; May et al., 1999), and so here we expected PI buildup to be greater for older adults and for those tested at nonoptimal times, and for release from PI to be diminished for these groups. We also examined the possibility that time of day effects might be greater for older than younger adults, as age-related impairments in inhibitory processes may render them particularly vulnerable to other assaults on inhibitory functioning. The pattern of data from our three measures of PI (recall rates, intrusion errors, and PI release) indicate that our predictions were generally, although not entirely, confirmed.

In assessing the buildup of PI, we considered both the recall data from Lists 1-3 and the intrusion errors from those lists together. Evaluating the recall measure alone may well be misleading, particularly when assessing group differences in PI when groups may vary in their ability to delete old items (as is the case in the present study; see Chiappe et al., 2000).

Assume that because of inefficient inhibitory control over deletion, obsolete and even irrelevant items clutter working memory (Hasher & Zacks, 1988), making it difficult for individuals to distinguish between targets and distractors. If individuals are impaired in their ability to suppress obsolete and other irrelevant items (and possibly in the ability to distinguish between them; e.g., Johnson, Hashtroudi, & Lindsay, 1993; Underwood, 1957), they may adopt one of two very different strategies for completing the task, both of which provide evidence for heightened PI, but which lead to very different patterns of performance. Poor inhibitors might elect to set a very high criterion for response, reporting only those items that they are certain are targets, with the consequence of reduced recall (relative to efficient inhibitors). On the other hand, poor inhibitors might elect to set a very low response criterion, reporting any items in working memory, and perhaps even generating possible targets from semantic knowledge. Such a strategy could inflate correct responses given the categorized items used here, and this strategy would also be associated with high intrusion rates. Our data provide evidence for this latter strategy by older adults, particularly those tested at off-peak times of day.

Consider first the recall data from Lists 1-3; there is PI buildup for all groups, regardless of age and testing time. Although older adults generally recalled fewer items on average than younger adults, the decrease in recall rates across lists was equivalent for younger and older adults. Furthermore, the decrease was equivalent across testing times for both age groups. Thus, with respect to target recall, there is no evidence for enhanced PI with age or nonoptimal times.

A close look at the intrusion errors, however, suggests that older adults, especially those tested at nonoptimal times, did indeed suffer greater PI. Intrusion rates increased across lists, were higher for older than younger adults, and were higher for older adults tested at off-peak relative to older adults tested at peak times. Together with the recall data, then, these findings are consistent with the interpretation that the efficiency of inhibition is reduced for older adults (particularly so for those tested at nonoptimal times), resulting in a reduced ability to distinguish between target items and obsolete (or irrelevant) ones. Participants with reduced inhibition appear to have more items in working memory and to report all candidates in working memory, regardless of their status as current or previous targets.

As further evidence for ineffective deletion in older adults, particularly those tested in the late afternoon, consider the release from PI data. Older adults failed to show any improvement in performance when new categories were introduced on List 4, suggesting that they were so impaired in their ability to escape the past and redirect to the present that even an external shift in materials could not facilitate this process. A similar finding was reported by May and Hasher (1998, Experiment 1) where evidence of the ability to escape the past came from an implicit test of both relevant and no longer relevant words from a garden path sentence processing task. There too, Morning type older adults tested in the late afternoon were unable to escape the past (in that study a self-generated, highly likely word
that completed a sentence) to acquire a new, experimenter-provided, less likely final word. Note that although in the present study, the time of testing effect for older adults was not reliable, the means were at least consistent with the pattern that PI susceptibility was greater at off-peak times.

The finding that older adults are more susceptible to PI than younger adults confirms a large literature showing a similar pattern (Kane & Hasher, 1995; Lustig & Hasher, 2001; Winocur & Moscovitch, 1983). The present data indicate that this impairment is heightened at off-peak times of day, and that age differences will be magnified if older adults are tested in the afternoon and evening. Indeed, for overall recall rates, intrusion error rates and release from PI rates, age differences were reliably greater in the afternoon (an optimal time for young but not old) than in the morning (an optimal time for old but not young). This pattern of exaggerated age differences in the evening replicates a number of similar findings (e.g., Hasher et al., 1999; May, 1999; May et al., 1993; Yoon et al., 1998), and taken together with the large proportion of older adults who are Morning-type people, reiterates the importance of time of testing as an essential consideration when making comparisons across age groups.

This pattern of findings is also consistent with emerging reports (e.g., Intons-Peterson et al., 1998; Petros et al., 1990; Yoon et al., 1998) that individual and age-related differences in circadian arousal patterns are nontrivial contributors to cognitive performance across the day.

Many people helped with this project and we acknowledge our appreciation to Lance Bloom, Sunghan Kim, Rachelle Ta-Min, Mark Leung, Jacqueline Hartley and Kelly Jacobs. This work was supported by grants RO1 12753 and R37 4306 from the US National Institute on Aging. This work was supported by grants RO1 12753 and R37 4306 from the US National Institute on Aging. Christie Chung is now at Claremont Graduate University. Send reprint requests to Lynn Hasher at hasher@psych.utoronto.ca.

References


La capacité de réguler de l’information qui a cessé d’être pertinente est un facteur essentiel de la performance mnésique globale de la mémoire explicite (Hasher, Zacks et May, 1999). Cette capacité, qui est en partie attribuable à la composante de détérioration du contrôle inhibitoire, peut varier chez un individu et d’un individu à l’autre. Dans la présente étude, nous avons examiné cette variabilité en effectuant des comparaisons entre la performance des jeunes gens et des personnes âgées, à qui l’on a administré des tests aux heures creuses ou aux heures de pointe de la journée. La tâche choisie pour examiner la fonction de détérioration du contrôle inhibitoire était une version inspirée du relâchement classique du phénomène d’interférence proactive, selon lequel le rappel diminue pendant la présentation de listes successives présentant des items verbaux similaires, pour ensuite augmenter (être « relâché ») lorsque la liste renferme des modifications.

Nous soutenons que la performance supérieure « relâché ») lorsque la liste renferme des modifications.

Si on tient compte tant des données de rappel obtenues lors de la présentation des trois premières listes que des données d’intrusion, les conclusions que nous avons dégagées laissent croire que, chez les per-
sonnes âgées, en particulier chez celles à qui l’on a administré le test pendant l’après-midi ou durant les périodes creuses de la journée, les effets de l’interférence étaient plus marqués que chez les adultes. Nous expliquons ces résultats par le fait que les participants qui font montre d’une inhibition réduite stockent un grand nombre d’items dans la mémoire de travail, ce qui serait en partie attribuable au fait qu’ils sont incapables d’éliminer efficacement l’information devenue non pertinente. Lors du test, ils ont réussi à reconnaître tous les items proposés, que ceux-ci constituent des informations pertinentes ou des informations devenues non pertinentes. Par ailleurs, les performances des personnes âgées ayant subi le test pendant la période creuse de la journée, n’ont pas permis de démontrer clairement une amélioration du rappel lorsque nous passions des catégories de la troisième liste à celles de la quatrième. Ces résultats vont dans le sens d’une perte de capacité accrue, chez les personnes âgées, qui les empêche de s’éloigner efficacement des informations passées et de déplacer leur attention vers des informations présentées récemment. D’un point de vue général, il convient de remarquer que les conclusions sur le rappel exposées dans le présent article sont conformes à d’autres conclusions selon lesquelles les différences d’âge au chapitre de la mémoire explicite s’accentuent, si une personne âgée effectue le test en après-midi.