# Research Report

# OPTIMAL TIME OF DAY AND THE MAGNITUDE OF AGE DIFFERENCES IN MEMORY

Cynthia P. May, Lynn Hasher, and Ellen R. Stoltzfus
*Duke University*

**Abstract**—*Across two studies comparing younger and older adults, age differences in optimal performance periods were identified (Study 1), and then shown to be an important determinant of memory differences (Study 2). A norming study showed that while most younger adults were Evening or Neutral types, as determined by a standard questionnaire, the vast majority of older adults were Morning types. A second study compared the recognition performance of younger and older adults tested in the morning or in the late afternoon. Substantial age differences were found in the late afternoon, when younger but not older adults were at their optimal times. However, no age differences in memory performance were found in the morning, when older but not younger adults were at their peak period. Thus, synchrony between optimal performance periods and the time at which testing is conducted may well be a critical variable in determining group differences in intellectual performance, particularly between older and younger adults.*

An informal survey taken at a recent academic meeting of cognitive gerontologists suggests that over 60% of both younger and older adults who participate in experiments do so in the afternoon.[1] In the past, this consistency in time of testing across age groups would have been considered an appropriate control for well-known time-of-day influences on performance (Baddeley, Hatter, Scott, & Snashall, 1970; Broadbent, Broadbent, & Jones, 1989; Folkard, Monk, Bradbury, & Rosenthall, 1977; Kleitman, 1963; Millar, Styles, & Wastell, 1980). Recent research on circadian rhythms and cognition, however, suggests that the general effect of time of day is moderated by individual differences in optimal performance periods (Anderson, Petros, Beckwith, Mitchell, & Fritz, 1991; Bodenhausen, 1990; Horne, Brass, & Pettitt, 1980; Petros, Beckwith, & Anderson, 1990), with best performance occurring at an individual's peak time during the day. These findings raise concerns about studies which compare subjects who differ on one dimension, such as age, and who, in fact, may also differ in their optimal performance periods.

With respect to aging, there is indeed some suggestion of differences in optimal time of day. Tune (1969) and Webb (1982) found that sleep-wake patterns shift with age toward early rising and retiring. Additionally, other researchers (Hoch et al., 1992; Mecacci, Zani, Rocchetti, & Lucioli, 1986), using the Horne-Ostberg (1976) Morningness-Eveningness Questionnaire, have found significantly higher morningness ratings for older adults relative to younger adults. Thus, if the suggestion that younger and older adults have different optimal periods during the day is correct, and if our survey of cognitive gerontologists is representative of procedures in many laboratories in which older and younger adults are tested, then there is indeed reason to be concerned about the generality of the widely reported age differences in cognition. Indeed, these may prove to be somewhat exaggerated.

The present investigation concerns the contribution that synchrony between optimal periods and the time of testing makes to performance differences seen between older and younger adults. We report two studies, one normative and one experimental, suggesting that there is substantial reason to be concerned about the time of day at which younger and older adults are tested in cognitive studies. In the first, we report norms for samples of younger and older adults confirming age differences in optimal times of day. In the second, we consider the impact of synchrony on age differences in a memory task.

## STUDY 1

In this normative study, we used Horne and Ostberg's (1976) Morningness-Eveningness Questionnaire to assess Morningness-Eveningness tendencies for samples of younger and older adults. The test is a pencil-and-paper questionnaire which provides a score assigning people to three main categories: Morning types, Neutral types, and Evening types. Psychometric assessments have shown the questionnaire to have good reliability (Buela-Casal, Caballo, & Cueto, 1990; Smith, Reilly, & Midkiff, 1989), and scores on this test have been shown to correlate with circadian variations in oral temperature, sleep-wake behavior, and periods of perceived alertness and performance (Buela-Casal et al., 1990; Horne & Ostberg, 1977; Mecacci & Zani, 1983; Smith et al., 1989). Furthermore, individual differences on this test have been shown to correlate with the cognitive performance of younger adults on several tasks (Anderson et al., 1991; Bodenhausen, 1990; Petros et al., 1990).

### Method

#### *Subjects*

Subjects in the norming study included 210 younger adults (ages 18–22) and 91 older adults (ages 66–78). The younger adults were introductory psychology students at Duke University; they participated as one way of satisfying a course requirement. The older

Address correspondence to Cynthia P. May or Lynn Hasher, Department of Psychology: Social and Health Sciences, Duke University, Durham, NC 27706.

1. A questionnaire was distributed at the Fourth Biennial Cognitive Aging Conference (April 1992), Atlanta, Georgia, to assess the time of day at which researchers typically test both younger and older subjects. Results from this questionnaire indicated that across researchers, 62% of older subjects and 71% of younger subjects are tested in the afternoon hours (noon to 6:00 p.m.).

 Copyright © 1993 American Psychological Society 

adults were community-dwelling volunteers obtained through the Duke University Center for the Study of Aging and Human Development. Demographic information from 60 of these older adults suggested that our sample was highly educated, having completed, on average, 16.8 years of formal education, and having achieved a mean verbal ability score of 37.9 on the Extended Range Vocabulary Test (ERVT; Educational Testing Service, 1976).

*Materials and procedure*

All subjects completed the Morningness-Eveningness Questionnaire. Younger subjects were tested in four large groups and were given as much time as needed to complete the questionnaire. The older adults received a copy of the Morningness-Eveningness Questionnaire in the mail and were asked to complete and return it in a self-addressed, stamped envelope; of the 150 older adults who received the questionnaire, 91 completed and returned it.

The questionnaires were scored according to the guidelines delineated by Horne and Ostberg (1976), with each response given a scaled score. These scores were added to determine a subject's Morningness-Eveningness rating. Total possible scores range from 16 to 86, with Morningness-Eveningness tendencies determined by the following scores: 16–30 (Definitely Evening), 31–41 (Moderately Evening), 42–58 (Neutral), 59–69 (Moderately Morning), and 70–86 (Definitely Morning).

## Results

The distribution of older and younger participants into the Morningness-Eveningness categories is shown in Table 1. The majority of older adults had high Morningness ratings, with a mean score of 64.5 ($SD$ = 7.6). Over 73% of older adults were either Moderately Morning or Definitely Morning types; no older adults could be classified as an Evening type. By contrast, 94% of the young subjects were either Evening or Neutral types, with a mean rating of 43.3 ($SD$ = 9.1). Less than 10% of younger subjects were Moderately Morning types, and no younger adult was a Definitely Morning type.

Table 1. *Distribution of Morningness-Eveningness scores for younger and older subjects in Study 1*

| | Morningness-Eveningness type | | | | |
|---|---|---|---|---|---|
| Group | Definitely Evening (16–30) | Moderately Evening (31–41) | Neutral (42–58) | Moderately Morning (59–69) | Definitely Morning (70–86) |
| Young ($N$ = 210) | | | | | |
| $n$ | 15 | 78 | 105 | 12 | 0 |
| % | 7 | 37 | 50 | 6 | 0 |
| Old ($N$ = 91) | | | | | |
| $n$ | 0 | 0 | 24 | 45 | 22 |
| % | 0 | 0 | 26 | 50 | 24 |

An analysis of variance (ANOVA) indicated a significant difference between younger and older adults on the Morningness-Eveningness Questionnaire, $F(1, 299) = 374.74$, $MS_e = 76.48$, revealing a reliable shift toward morningness with age.

## Discussion

Older and younger adults clearly differed in their Morningness-Eveningness tendencies: In the present samples, older adults tended to be Morning types while younger adults showed strong eveningness tendencies. These results are generally consistent with the literature on similar populations. The skew toward morningness our older subjects showed is consistent with that found by Mecacci et al. (1986; see also Hoch et al., 1992) and with data showing a shift in sleep-wake patterns with age (Tune, 1969; Webb, 1982). The low proportion of Morning types among our younger adults is consistent with evidence reported by other investigators (e.g., Anderson et al., 1991; Mecacci & Zani, 1983), and is also consistent with the suggestion that evening activities associated with a college setting may cause an increase in the tendency toward eveningness (Webb & Bonnet, 1978).

The clear age differences in Morningness-Eveningness patterns found here suggest that one can expect optimal performance from most older adults in the morning and optimal performance from many younger adults in the afternoon and evening. Thus, the outcomes of cognitive studies comparing older and younger adults can, to some degree, be expected to vary depending on when subjects of different ages are tested.

# STUDY 2

Older adults show a wide range of cognitive deficits, particularly on tasks assessing memory (for a review, see Kausler, 1991). In Study 2, we employed a particularly difficult memory task, verbatim recognition of sentences from a series of paragraphs. The experiment compared the performance of younger and older adults who were tested in either the early morning or the late afternoon. Subjects were selected from the populations normed in Study 1. Each participant selected had a personal optimal time that was consistent with the age norms reported in Study 1: Younger participants all had evening peak periods, and older ones all had morning peak periods. Prior to completing the normative study, we had intended a fully crossed design of Age × Morningness-Eveningness × Time of Testing for the present experiment. This design was not possible, however, since few of our younger subjects were Morning types and none of our older subjects were Evening types.

The task we used required subjects to read and remember a series of paragraph-length stories, then to recognize sentences from the passages they had read. To make the test particularly difficult, the foil items were plausible inferences based on the content of each story and were closely matched in style and vocabulary to the original texts. Thus, false alarms might be expected to be high, particularly at subjects' nonoptimal times.

## Method

### Subjects

Subjects were selected from the norming samples reported in Study 1 on the basis of their scores on the Morningness-Eveningness Questionnaire, with all older adults having strong Morningness scores ($M$ = 70.2, $SD$ = 3.7) and all younger adults having strong Eveningness scores ($M$ = 29.3, $SD$ = 3.5). Twenty-two older adults ($M$ age = 70.5 years; range: 66–78) and 20 younger adults ($M$ age = 18.8 years; range: 18–20) participated in the reading study. All subjects were given the ERVT (Educational Testing Service, 1976) to assess verbal ability. As is often the case, the older participants had reliably higher vocabulary scores ($M$ = 39.7) than the younger participants ($M$ = 25.7), $F(1, 40) = 61.48$, $MS_e = 30.25$.

### Materials and procedure

Half of the subjects in each age group were tested in the morning (at either 8:00 or 9:00), and half were tested in the afternoon (at either 16:00 or 17:00), such that half of each age group was tested at that group's optimal time of day (i.e., older adults who participated in the morning and younger adults who participated in the afternoon), while the other half was tested at the group's nonoptimal time (i.e., older adults who participated in the afternoon and younger adults who participated in the morning). All subjects were tested individually, with the experiment lasting about 1 hr.

Materials consisted of 10 short stories (each approximately 250 words in length).[2] Each story was presented on a computer screen, one line at a time. Subjects read silently at their own pace, advancing through the lines of a story by pressing the space bar on the computer keyboard, with the computer programmed to record reading time. Subjects were instructed to read the stories carefully and were informed that they would be tested for their verbatim memory of story sentences when they had read all 10. There were two orders of story presentation, used equally often for each age group at each time of testing.

Immediately after reading all 10 stories, subjects took the recognition test. For this test, 30 old test sentences, 3 from each of the 10 stories, and 30 new sentences (called foils), again 3 for each of the stories, were presented. Thus, half of all test sentences were old and half were new.

Sentences appeared one at a time. Subjects were required to judge whether each sentence had been presented, in verbatim form, in one of the original stories. All sentences from the same story appeared sequentially, and stories were tested in the same order as in their initial presentation.

### Design

The design of this experiment combined age and time of testing factorially, with dependent measures of story reading time and recognition memory accuracy. Age was fully confounded with Morningness-Eveningness, since no older participants were Evening types and no younger participants were Morning types. The central question was whether the time of day at which testing occurs affects the magnitude of age differences in recognition memory.

## Results and Discussion

The significance level was set at .05 for all analyses. Four older subjects, two tested in the morning and two in the afternoon, failed to follow instructions in the recognition portion of the task, and thus their data were deleted from further analyses.

### Reading time

Mean reading times for the two groups are displayed in Table 2. A 2 (Age: young vs. old) × 2 (Time of Testing: morning vs. afternoon) ANOVA was performed to assess reading time (per word) of the 10 experimental stories. Younger subjects read at a faster rate ($M$ = 434 ms per word) than older subjects ($M$ = 595 ms per word), $F(1, 36) = 21.66$, $MS_e = 12{,}303.19$. There was no main effect of time of testing, $F(1, 36) = 1.58$; however, there was a significant interaction between age and time of testing, $F(1, 36) = 6.00$, $MS_e = 12{,}303.19$. While the reading rate of younger subjects did not differ significantly from morning to afternoon ($F < 1$), that of older subjects was significantly slower in the morning than in the afternoon, $F(1, 17) = 8.86$, $MS_e = 9{,}007.26$. Taken together with the findings reported below, these results suggest the possibility that older subjects are better able or are more willing to read closely at their optimal time than at their nonoptimal time.

Table 2. *Mean reading times per word (in ms) for younger and older adults by testing time in Study 2*

| | Time of testing | |
|---|---|---|
| Group | a.m. | p.m. |
| Young | 412.8 | 455.9 |
| | (90.6) | (149.2) |
| Old | 669.6 | 535.6 |
| | (68.2) | (111.3) |

*Note.* Standard deviations are in parentheses.

### Recognition accuracy

Table 3 presents hits for old items and false alarms to foils for younger and older adults. A cursory inspection of these data reveals the need for a corrected recognition score (here, hits minus false alarms; see Snodgrass & Corwin, 1989). In particular, note that for young adults, hits and false alarms vary across the day in such a way as to suggest heightened general accuracy in the afternoon: Their hit rate is higher in the afternoon than in the morning, and the reverse is true for their false alarm rate. For older adults, the situation is more complex. Their hit rate remains fixed across the day, but their false alarm rate increases substantially from the morning to the afternoon. Thus, analyses are reported for corrected recognition scores.

As is often the case, younger subjects

2. Lynne Reder generously provided these stories, which have been used in a number of previous studies (e.g., Reder, 1982; Reder, Wible, & Martin, 1986).

Table 3. *Mean percentage of hits, false alarms, and corrected recognition for younger and older adults by testing time in Study 2*

| | Hits | | False alarms | | Corrected recognition[a] | |
|---|---|---|---|---|---|---|
| Group | a.m. | p.m. | a.m. | p.m. | a.m. | p.m. |
| Young | 75.7 | 87.5 | 29.5 | 22.5 | 46.2 | 65.0 |
| Old | 84.4 | 85.7 | 37.2 | 55.3 | 47.2 | 30.4 |

[a] Corrected recognition = hits − false alarms.

Acknowledgments—This research was supported by Grant AGO 4306, awarded by the National Institute on Aging to Lynn Hasher and Rose T. Zacks, and by a National Science Foundation Fellowship awarded to Cynthia May. We greatly appreciate the help of Rose Zacks and the Duke University Center for the Study of Aging and Human Development at various stages of the project. We also extend thanks to the researchers at the Fourth Biennial Cognitive Aging Conference (April 1992) for their cooperation in our survey. Ellen Stoltzfus is now at Kenyon College.

were more accurate than older subjects, $F(1, 36) = 9.98$, $MS_e = 265.62$. However, the age effect was tempered by an interaction between age and time of testing, $F(1, 36) = 11.12$, $MS_e = 265.62$. In the morning, there was no significant difference between younger and older adults ($F < 1$). In the afternoon, however, younger adults held a substantial advantage over older adults, $F(1, 19) = 20.51$, $MS_e = 290.13$. Furthermore, the performance of younger adults improved from morning to afternoon, $F(1, 19) = 7.5$[illegible], $MS_e = 232.11$, while the performance of older adults declined from morning to afternoon, $F(1, 17) = 4.08$, $MS_e = 303.31$. These results, reflecting the fact that each age group's optimal performance is seen at that group's peak period, create an overall pattern in which older and younger adults are equivalent in the morning and in which older adults are substantially disadvantaged in the afternoon.

## GENERAL DISCUSSION

Differences in optimal performance periods clearly affect recognition memory for both younger and older adults: Younger adults showed better recognition memory when tested at their peak performance period (the afternoon) than when tested at their off-peak period (the morning). Older adults showed better recognition memory when tested at their optimal time (the morning) than when tested at their off-peak period (the afternoon). Thus, the effects of time of day on performance are not as consistent across individuals as was once thought (e.g., Broadbent et al., 1989; Folkard, 1979; Oakhill, 1986). Instead, the effects of time of day vary as a function of the synchrony between individual optimal performance periods and the time at which testing occurs. These findings are consistent with others showing an impact of optimal time on the cognitive performance of younger adults (e.g., Bodenhausen, 1990; Petros et al., 1990), and extend these findings to older adults as well.

As is common in cognitive gerontology, overall age differences in recognition memory were observed, with younger adults demonstrating better memory performance than older adults. However, the magnitude of these age differences varied greatly across the day: Exaggerated age differences were observed in the afternoon, when younger but not older adults were at their peak time, yet no age differences were observed in the morning, when older but not younger adults were at their optimal time.

The present findings should serve as general caution to researchers investigating group differences of any sort. For example, if groups differ in optimal performance periods, and if all participants are tested at approximately the same time, the outcome, whether showing differences or similarities, may be misleading. Given the age differences in Morningness-Eveningness tendencies, then, age differences in cognitive performance will be estimated best when the synchrony between optimal performance and time of testing is controlled. Ideally, both younger and older adults would be tested at their optimal time periods. However, our informal survey of cognitive gerontologists suggests that in many laboratories, the majority of participants are tested in the afternoon. Thus, it is conceivable that at least for some tasks (and perhaps many), the average age difference reported in the literature is somewhat exaggerated.

## REFERENCES

Anderson, M., Petros, T.V., Beckwith, B.E., Mitchell, W.W., & Fritz, S. (1991). Individual differences in the effect of time of day on long-term memory access. *American Journal of Psychology, 104*, 241–255.

Baddeley, A.D., Hatter, J.E., Scott, D., & Snashall, A. (1970). Memory and time of day. *Quarterly Journal of Experimental Psychology, 22*, 605–609.

Bodenhausen, G.V. (1990). Stereotypes as judgmental heuristics: Evidence of circadian variations in discrimination. *Psychological Science, 1*, 319–322.

Broadbent, D.E., Broadbent, M.H.P., & Jones, J.L. (1989). Time of day as an instrument for the analysis of attention. *European Journal of Cognitive Psychology, 1*, 69–94.

Buela-Casal, G., Caballo, V.E., & Cueto, E.G. (1990). Differences between morning and evening types in performance. *Personality and Individual Differences, 11*, 447–450.

Educational Testing Service. (1976). *Kit of factor-referenced cognitive tests*. Princeton, NJ: Author.

Folkard, S. (1979). Time of day and level of processing. *Memory & Cognition, 7*, 247–252.

Folkard, S., Monk, T.H., Bradbury, R., & Rosenthall, J. (1977). Time of day effects in school children's immediate and delayed recall of meaningful material. *British Journal of Psychology, 68*, 45–50.

Hoch, C.C., Reynolds, C.F., Jennings, J.R., Monk, T.H., Buysse, D.J., Machen, M.A., & Kupler, D.J. (1992). Daytime sleepiness and performance among healthy 80 and 20 year olds. *Neurobiology of Aging, 13*, 353–356.

Horne, J., Brass, C., & Pettitt, S. (1980). Circadian performance differences between morning and evening types. *Ergonomics, 23*, 29–36.

Horne, J., & Ostberg, O. (1976). A self-assessment questionnaire to determine morningness-eveningness in human circadian rhythms. *International Journal of Chronobiology, 4*, 97–110.

Horne, J., & Ostberg, O. (1977). Individual differences in human circadian rhythms. *Biological Psychology, 5*, 179–190.

Kausler, D.H. (1991). *Experimental psychology, cognition, and human aging* (2nd ed.). New York: Springer-Verlag.

Kleitman, N. (1963). *Sleep and wakefulness*. Chicago: University of Chicago Press.

Mecacci, L., & Zani, A. (1983). Morningness-eveningness preferences and sleep-waking diary data of morning and evening types in student and worker samples. *Ergonomics, 26*, 1147–1153.

Mecacci, L., Zani, A., Rocchetti, G., & Lucioli, R. (1986). The relationships between morningness-eveningness, ageing and personality. *Personality and Individual Differences, 7*, 911–913.

Millar, K., Styles, B.C., & Wastell, D.G. (1980). Time of day and retrieval from long-term memory. *British Journal of Psychology, 71*, 407–414.

Oakhill, J. (1986). Effects of time of day on the integration of information in text. *British Journal of Psychology, 77*, 481–488.

Petros, T.V., Beckwith, B.E., & Anderson, M. (1990). Individual differences in the effects of time of day and passage difficulty on prose memory in adults. *British Journal of Psychology, 81*, 63–72.

Reder, L.M. (1982). Plausibility judgments versus fact retrieval: Alternative strategies for sentence verification. *Psychological Review, 89*, 250–280.

Reder, L.M., Wible, C., & Martin, J. (1986). Differential memory changes with age: Exact retrieval versus plausible inference. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12*, 72–81.

Smith, C.S., Reilly, C., & Midkiff, K. (1989). Evaluation of three circadian rhythm questionnaires with suggestions for an improved measure of morningness. *Journal of Applied Psychology, 74*, 728–738.

Snodgrass, J.G., & Corwin, J. (1989). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology: General, 117*, 34–50.

Tune, G.S. (1969). The influence of age and temperament on the adult human sleep-wakefulness pattern. *British Journal of Psychology, 60*, 431–441.

Webb, W.B. (1982). Sleep in older persons: Sleep structures of 50- to 60-year-old men and women. *Journal of Gerontology, 37*, 581–586.

Webb, W.B., & Bonnet, M.H. (1978). The sleep of 'morning' and 'evening' types. *Biological Psychology, 7*, 29–35.

(RECEIVED 8/10/92; REVISION ACCEPTED 1/19/93)

**Teachers College, Columbia University**

**Applied Educational Psychology: School Psychology Assistant/Associate Professor**

**Two-year Appointment Tenure-track**

The Program in School Psychology, located in the Department of Developmental and Educational Psychology, follows a scientist-practitioner model with a focus on learning and behavior problems in school contexts. The program is accredited by the American Psychological Association. The candidate we are seeking needs to have demonstrated expertise in consultation and experience in work with culturally diverse populations.

Requirements: Ph.D. or Ed.D. in school psychology; evidence of demonstrated potential for a program of empirical research; teaching competence; eligibility for NYS licensure and certification as a school psychologist. Interest in one or more of the following areas is highly desirable: family issues as related to assessment and intervention in school settings; alternative assessment procedures; preschool populations; pre-referral interventions. Candidates need to have a clear commitment to education and advocate the prevention model of school psychology. Individuals with other areas of expertise are invited to apply.

Responsibilities: Teaching, dissertation advisement, research, and supervision of clinical experiences.

Applications will be reviewed commencing October 31, 1993. Candidates whose applications are received after that date cannot be promised full consideration. Send CV, letter of application including research interests and applied experiences, publications, and 3 letters of reference to: Dr. Ann E. Boehm, Box 1, Teachers College, Columbia University, 525 West 120 Street, New York, NY 10027.

*Teachers College as an institution has long been committed to a policy of equal opportunity in employment. In offering higher education in the discipline areas of education, psychology, and health services, the College is committed to providing expanding employment opportunities to minorities, in its own activities and in society. Candidates whose qualifications and experience are directly relevant to complementary College priorities (e.g., urban and minority concerns) may be considered for a higher rank than advertised.*