ORIGINAL ARTICLE

AGE-RELATED DIFFERENCES IN PROSPECTIVE MEMORY: TURKISH VIRTUAL WEEK (VW-TR)

Turkish Journal of Geriatrics
DOI: 10.29400/tjgeri.2023.351
2023; 26(3):249−257

Gün PAKYÜREK[1]
Banu CANGÖZ TAVAT[2]

Abstract

Introduction: Prospective memory, which involves remembering intended future actions, is a vital function in terms of autonomy, quality of life, and everyday functioning. The primary aim of this study is to examine how aging affects prospective memory performance; its secondary aim is to adapt a laboratory-based prospective memory task, Virtual Week, to the Turkish culture and investigate its efficacy across young and old age groups.

Materials and Method: The study was conducted with 60 young (18–25 years) and 60 old (60–87 years) participants. Participants were included based on their results on cognitive screening tests (Montreal Cognitive Assessment, Activities of Daily Living Scale, Geriatric Depression Scale, and the Beck Depression Inventory). In addition, the Stroop and Trail Making tests were administered to measure executive functions. Finally, the laboratory-based prospective memory task Virtual Week was performed.

Results: Virtual Week has been adapted to Turkish culture and shown to be reliable (Spearman–Brown: 0.82). ANOVA was conducted to analyze the effect of aging on the Virtual Week task, and the results showed that young adults were more successful than older adults in prospective memory tasks ($p < .05$).

Conclusion: The results support the theory of the aging paradox in prospective memory, which suggests that older adults exhibit lower performance in laboratory-based prospective memory tasks. The findings are discussed in the context of the relevant literature.

Keywords: Memory; Cognitive Aging; Executive Function; Young Adult; Aged.

CORRESPONDANCE

[1]Gün PAKYÜREK

Phone : +905358402756
e-mail : gunpakyurek@gmail.com

Received : May 03, 2023
Accepted : Jun 22, 2023

[1] Izmir Democracy University, Psychology, Izmir, Turkey

[2] Hacettepe University, Psychology, Ankara, Turkey

INTRODUCTION

Prospective memory (PM) is defined as memory that allows us to remember that a planned action must be performed based on a specific cue during an ongoing task (1). In laboratory studies, PM is examined in two ways: event-based prospective memory (EB-PM) and time-based prospective memory (TB-PM), depending on the type of cue (1). An event-related environmental cue in tasks carried out in EB-PM requires less self-initiating process and mental effort—for example, remembering to pay the rent when you see the owner. In the case of TB-PM, there is no environmental cue related to the event—the cue is elapsed time. In such a task, spontaneous processes and mental effort are more involved—for example, remembering to take medication at 9 pm (2). Remembering routine tasks (e.g., going to work) requires less retrospective memory (RM), while remembering nonroutine tasks (e.g., a dentist appointment) requires more RM (3).

The multiprocess model of prospective memory (1) and the theory of preparatory attentional and memory processes (PAM) (4) focus on monitoring, the use of limited resources, and related cues.

It has been shown that PM is a critical indicator of functional independent living by investigating PM in laboratory tasks representing daily life (5). The question of how PM performance is affected by the aging process is a matter of interest. In this context, it has been shown that PM performance increases from birth to age 35 and then begins to decline (6). PM requires multistage processes and a relatively high cognitive load. Therefore, PM performance is expected to decline with advancing age (7). Laboratory-based studies have shown that PM performance deteriorates with aging (6, 8, 9), but that older people can be as successful as young people in PM studies compatible with real-life events outside the laboratory (5, 10, 11). On the other hand, there are also studies finding no difference between old and young people in terms of EB-PM (10). The "age prospective memory paradox" (2) posits that older participants perform differently in laboratory and non-laboratory tasks. Haines et al. (12) explain the age PM paradox as a lack of environmental support and cognitive processes, whether automatic or not. Another explanation is that older participants may have developed strategies that use external assistance more in tasks representing real-life events, while young participants may not perform fully due to lack of motivation (5). It was observed that age affected TB-PM and EB-PM differently (8, 9). With aging, TB-PM may deteriorate more than EB-PM (8, 9). Older participants were found to have more success on PM only when they received social feedback (7). The young-old (60–75) and old-old (over 75) age groups showed no significant difference outside the laboratory, although the old-old group failed to perform laboratory-based PM (11).

A new technology that detects spontaneous speech production related to PM found no significant difference between younger and older adults (10). Although various tasks have been developed to evaluate PM objectively, the most structured and highly externally validated one is Virtual Week (VW) (13). Although VW is a laboratory task, it is similar to everyday life, as it involves real-life events. VW was developed in English-speaking regions in North America and Australia. It was then adapted and applied to non-native English cultures (Germany, Poland, and Italy) in Europe (5, 14, 15).

The main purpose of this study is to examine the effect of aging on PM performance. The secondary objective is to examine the functioning of VW (13) in Turkish culture.

METHOD

Participants and Materials

The study was carried out with 120 volunteers, 60 of whom were young (female = 30, male = 30) and 60 of whom were old (female = 30, male = 30). The mean age of the young group was 21.03 (1.48), and

the mean age of the older group was 71.38 (7.43). The two groups were equivalent in terms of years of education, and there was no significant difference between the groups in terms of education level (year) ($t_{(118)} = 1.53$, $p = .13$).

The Beck Depression Inventory (BDI) (16), Geriatric Depression Scale (GDS) (17), Montreal Cognitive Assessment (MoCA) (18), and Functional Activities Questionnaire (FAQ) (19) were used as inclusion criteria. The cut-off points for inclusion criteria were 17 for the BDI (16) for the young group; 15 for the GDS (17); 23 for MoCA (18); and 5 and 7 for the FAQ (19), corresponding to ages 60–70 and 70+, respectively for the older group. The mean and standard deviation scores obtained from the screening tests are summarized in Table 1. Standardization studies, in which all neuropsychological tests and scales were used, were conducted for Turkish culture, and all tests and scales had culture-specific norms.

This research employed the computerized version of the VW developed by Rendell and Henry (13). VW is a task that begins with the subject rolling a virtual dice on a representative game card on the computer screen. Some adaptations have been made because of cultural differences. Turkish culture differs from the Australian and North American contexts in which VW was developed, particularly in terms of eating habits, social habits and/or relations, and celebrities. Thus, culture-specific adjustments were made. For a screenshot of the trial-day dinner, see Figure 1. Minor changes have been made as possible. Foreign dishes in Turkish cuisine such as porridge and cereal replaced with local dishes such as lahmacun and kebab.

Table 1. Screening Tests' mean and standard deviations scores

	BDI M (S)	**GDS** M (S)	**FAQ** M (S)	**MoCA** M (S)
Age Group				
Young (n=60)	8.20 (4.80)	-	-	-
Old (n=60)	-	9.87 (4.81)	14.15 (7.38)	26.43(3.43)

BDI: Beck Depression Inventory, FAQ: Functional Activities Questionaire, GDS: Geriatric Depression Scale, MoCA: Montreal Cognitive Assessment

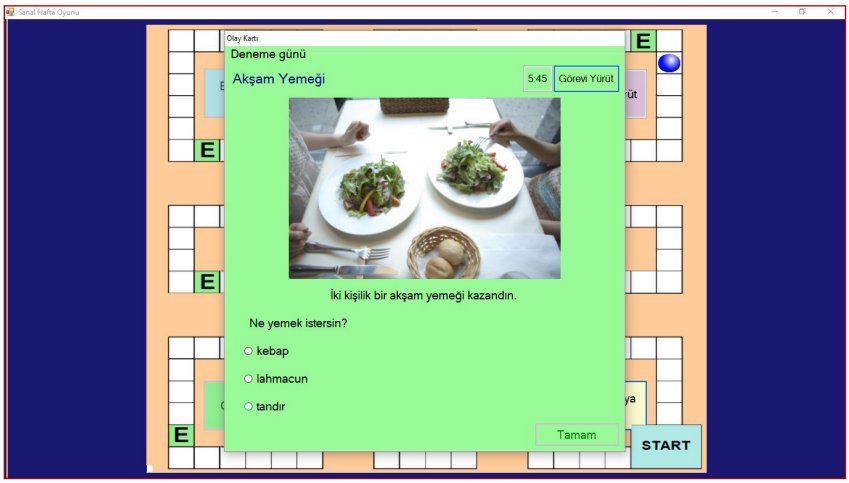

Figure 1. Turkish Virtual Week- a screenshot of the trial-day dinner

Rose et al. (3) found that the reliability coefficient of Spearman-Brown's split halves for VW was .71 in the young group and .93 in the older group. In the Italian version of the VW, the Cronbach's alpha internal consistency coefficient was found to be .64 in the young group and .92 in the older group (15). In the Polish version, Spearman-Brown's split-half reliability coefficient was .75 in the young group and .95 in the old group (14).

Procedure

Informed consent was obtained from all participants before starting the study. Then screening tests/scales were applied. After the screening test/scales, VW was carried out. VW includes regular tasks (e.g., taking medication at breakfast and dinner), regular times (e.g., taking asthma medication at 21:00), and irregular tasks that are specific to a particular day and time (e.g., calling the dentist to make an appointment at 12:00 o'clock). In addition, participants are required to perform lung function tests at two minutes and four minutes in real time. This is called a time-check task.

VW consists of one trial day and three virtual days. Each virtual day included four regular (two event-based, two time-based), four irregular (two event-based, two time-based), and two time-check tasks. A total of 30 PM tasks were used during the experiment. At the end of each virtual day, a recall test was performed to assess whether the tasks were coded by the participants. In the recall test, the planned action (e.g., buying a colored pencil) was selected from the distractors (e.g., receiving a birthday present for the participant's nephew), and the participant was asked to match it with an action cue (e.g., shopping). This study was approved by the Hacettepe University Ethics Committee (reference no. 35853172/431-71).

Statistical Analysis

All analyses were conducted with the Statistical Package for the Social Sciences (IBM®, SPSS 25). A 2 × 2 × 2 ANOVA was conducted age group (young vs. old), PM task (regular vs. irregular), and PM cue (event vs. time) to analyze the performance of VW as a function of age and task demand. The independent variable "age" was manipulated between groups and the other independent variables PM task and PM cue were manipulated within groups. Due to the real-time time-check tasks being quite separate from the virtual times a 2×3 ANOVA was also conducted on the age group (young vs. old), and PM task (regular, irregular vs. time check task) as in the other versions of VW (2, 5, 15). Analyses were conducted regarding the correct answers (responses in the right place and at the right time) to the VW-TR.

RESULTS

The Stroop test (20) and the Trail Making Test (TMT-Part B) (21) were applied to determine the executive functions of the participants. These tests were considered pertinent for prospective memory. The mean and standard deviations of the scores obtained from the Stroop test (20) and the Trail Making Test (TMT-Part B) (21) are presented in Table 2. The young group's duration of completion of the Stroop test was shorter (U = 1074.5, p < .001) and had a lower total score for error and correction (U = 1384.0, p < .05) than the older group (U = 1274.0, p < .05). In addition, the young group's TMT-Part B error and correction total scores (U = 1358.5, p < .05) were lower than those of the older group.

The Cronbach's alpha internal consistency coefficient for the Turkish version of the VW (VW-TR) was .77, and the Spearman–Brown split-half reliability coefficient was .82 in all groups. The Spearman–Brown split-half reliability coefficients were .65 for the young group and .61 for the older group.

Table 3 shows the mean and standard deviation scores of the young and older participants on the VW-TR.

Table 2. Mean and standard deviations of age, education level, and neuropsychological test scores

	Education (years)	Age	TMT-Part B duration of completion	TMT-Part B sum of error and correction	Stroop Test duration of completion	Stroop Test sum of error and correction
	M (S)	M (S)	M (S)	M (S)	M (S)	M (S)
Age Group						
Young	13.68 (.99)	21.03 (1.48)	55.36 (30.79)	.35 (.76)	17.90 (4.99)	0.48 (.93)
Old	13.28 (1.75)	71.38 (7.43)	73.82 (41.93)	.78 (.98)	22.68 (8.18)	0.92 (1.15)

TMT-Part B: Trail Making Test- Part B

Table 3. Mean and standard deviations scores of older and young participants on VW-TR

		Age Group			
		Young		Older	
Task of Type		**Event**	**Time**	**Event**	**Time**
Regular Tasks	M	.82	.79	.72	.37
	SD	.18	.22	.25	.28
Irregular Tasks	M	.85	.57	.65	.13
	SD	.15	.25	.22	.17

The 2x2x2 ANOVA revealed a significant main effect for age ($F_{(1,118)}$=150.7, p<.001, ηp^2=.56); PM task ($F_{(1,118)}$=54.7, p<.001, ηp^2=.32); PM cue ($F_{(1,118)}$=199.4, p<.001, ηp^2=.63); interaction effect of PM cue and age ($F_{(1,118)}$=44.2, p<.001, ηp^2=.27), interaction effect of PM task and PM cue ($F_{(1,118)}$=39.9, p<.001, ηp^2=.25). There was no significant interaction effect of PM task and age ($F_{(1,118)}$=2.5, p>.05), and PM task, PM cue, and age ($F_{(1,118)}$=1.61, p>.05).

Table 4 presents the mean and standard deviation scores on the regular, irregular, and time-check tasks according to the age groups.

The 2x3 ANOVA revealed a significant main effect for age ($F_{(1,118)}$=240.1, p<.001, ηp^2=.67); PM task ($F_{(2,236)}$=40.0, p<.001, ηp^2=.25); and PM task and age interaction ($F_{(2,236)}$=22.2, p<.001, ηp^2=.16). Post-hoc analyses showed that the members of the young group were more accurate in performing regular PM tasks (M=.68, S=.22) than irregular PM tasks (M=.55, S=.22) and time check tasks (M=.48, S=.36), as well as more accurate in performing irregular PM tasks (M =.55, S=.22) than time check tasks (M=.48, S=.36). In addition, post-hoc analyses revealed that the young group (M=.76, S=.14) outperformed the older group (M=.38, S=.13) overall (p<.001, ηp^2=.67): regular task for young (M=.81, S=.16) versus older (M=.54, S=.20); irregular task for young (M=.71, S=.16) versus older (M=.39, S=.14); and time check task for young (M=.75, S=.24) versus older (M=.21, S=.23). In the post-hoc analyses for age and PM task interaction, it was revealed that the older group displayed worse performance on the time check task (M=.21, S=.23) compared to the regular (M=.54, S=.20) and

Table 4. Mean and standard deviation scores on regular, irregular and time check tasks according to the age groups

PM Task	Regular Task M (S)	Irregular Task M (S)	Time Check Task M (S)
Age Group			
Young (n=60)	.81(.16)	.71(.16)	.75(.24)
Older (n=60)	.54(.20)	.39(.14)	.21.(.23)

PM Task: Prospective Memory Task

irregular tasks (M=.39, S=.14) ($p<.05$). The results for the older group were also worse on the irregular task (M=.39, S=.14) as opposed to the regular task (M=.54, S=.20) ($p<.05$). Similarly, the younger group displayed worse performance on the irregular task (M=.71, S=.16) than on the regular task (M=.81, S=.16) ($p<.05$). However, there were no significant differences between regular task (M=.81, S=.16) and time-check task (M=.75, S=.24) performance for the young group ($p>.05$). There were also no significant differences between the irregular task (M=.71, S=.16) and time-check task (M=.75, S=.24) performance for young adults. Since the older adults experienced difficulty using a mouse and requested assistance in this regard, the time to complete the VW-TR was not calculated.

DISCUSSION

As expected, young adults were more successful than older adults on PM. This difference was especially high for the irregular and time-check tasks (*see* Table 3). These results support the age-prospective memory paradox and are in line with previous research showing that older adults perform worse on laboratory tasks than younger adults (2, 5, 8, 9, 11). Due to the necessity of using RM and PM together competently, irregular tasks require a greater cognitive load. Therefore, the performance of the older adults was lower on irregular tasks where the use of RM was especially necessary. This finding is also consistent with the theories of the multiprocess framework (5) and PAM theory (6), which focus on spontaneous retrieval, demands of an ongoing task, and related cues. In a study by Rendell and Thomson (11), when young-old and old-old groups were compared, the old-old group performed worse on laboratory-based PM tasks than the young-old group. Additionally, Bozdemir and Cinan (8) found the performance of the old group was worse than that of the young group on laboratory-based PM tasks.

According to the interdependence model proposed by Imamoğlu (22) for Turkish people, collectivist, and individualist tendencies coexist in this culture. In this context, the fact that the Turkish (Middle Eastern) sample of older adults faced problems in accessing health services and led a relatively sedentary lifestyle associated with living in a collectivist culture (meaning many daily activities are undertaken by the family and relatives of the older individuals) (23) may have led to higher failure rates on laboratory-based VW-TR tasks when compared to the older individuals in the Western sample. Consistent with this evaluation, Turkish older samples' norm values determined through neuropsychological tests, such as MoCA, TMT, and cut-off points, were shown to be lower than those for the same age group in Western culture (18, 21). While neuropsychological tests did not predict naturalistic, self-report, and clinically based prospective memory performance, the obtained

results were found to be consistent with the Age Prospective Memory Paradox, due to the current study being laboratory-based, results are expected opposite to naturalistic prospective memory performance in older individuals (24) Recent research findings also suggest that intercultural differences exist in episodic memory and semantic memory performances (25). It is thought that the lower PM performance of the individuals in the Turkish older sample may be explained by the cultural differences mentioned (22, 23). Additionally, stereotype threats, such as unsuccessful memory performance and a slowdown in response speed, may adversely affect PM performance, especially in older samples (26). Stereotype threat is higher in laboratory-based tasks. Although recreational drug-related deficits in PM have been reported in recent years, we have not asked about it, by reason of they usually have not been honest about drug use in self-reports in Turkish culture (27). Age is also known to influence time-based and event-based performance differently according to the cue type (14, 15). In the other versions of the VW (Western), there was more deterioration in time-based tasks compared to event-based tasks with aging (2, 5, 14, 15). In line with Western culture, there was more deterioration in time-based and time-check VW-TR tasks with aging in Turkish culture.

In the present study, the VW-TR, which is the Turkish version of the VW, was employed. The VW was originally developed in North America and Australia and adapted in Germany, Poland, and Italy. When the reliability scores of Western culture were compared, the results in this study were relatively lower in older adults but compatible with the results for younger adults. One of the reasons for this may be due to the limited use of technology by older Turkish individuals compared to the same demographic in the West. In this study, older participants lost time by having difficulty using the mouse, and by focusing on this issue, they may have struggled to concentrate and follow the ongoing task. The variety of problems encountered, including those related to using the mouse and reading, as well as having to help participants with such difficulties, may have resulted in lower reliability scores on the VW-TR than on other versions of the VW.

The adapted VW-TR was found to be suitable for young participants, while older participants experienced some practical difficulties related to computer use. In this context, using the board game version of VW by hand or using a touchscreen computer can be recommended for future studies. It has been shown that VW, which was developed for Western culture, is also functional in Turkish culture and can be used as a comprehensive laboratory task that measures PM with various (event-based, time-based, time check, regular, irregular) and real-life tasks. It was found that the performance of PM deteriorated with aging in laboratory-based tasks. Niedzwienska et al (7), suggest that the comparison of the daily real-life tasks and laboratory-based tasks of PM will contribute to the literature. As in the study of. Haines et al. (12) the naturalistic PM task performed with older adults in MEMO (smartphone application) may be also studied in older Turkish people, and it is thought that it will be useful to compare the performance of the older adults in the naturalistic PM task with VW, which is the other leg of the '*Age Prospective Memory Paradox*' (3).

Acknowledgements

We would like to express our very great appreciation to Peter Rendell and Simon Haines for helping to use Virtual Week and guiding the research process. Also, we would like to thank Ankara İTÜ Evi, Ayşenur Kanat, Betül Demir, Büşra Dindaş, Şahin Şahinoğulları, Büşra Uğur, and Ayşenur Genç for helping with the data collection of this research work.

REFERENCES

1. McDaniel MA, Einstein GO. Strategic and automatic processes in prospective memory retrieval: A multi-process framework. Appl Cogn Psychol 2000;14:127-44. (DOİ:10.1002/acp.775)
2. Rendell PG, Craik FIM. Virtual week and actual week: Age-related differences in prospective memory. Appl Cogn Psychol 2000;14(7):43-62. (DOİ:10.1002/acp.770)
3. Rose NS, Rendell PG, McDaniel MA, Aberle I, Kliegel M. Age and individual differences in prospective memory during a "virtual week": The roles of working memory, vigilance, task regularity, and cue focality. Psychol Aging 2010;25(3):595-605. (DOI: 10.1037/a0019771)
4. Smith RE. The cost of remembering to remember in event-based prospective memory: investigating the capacity demands of delayed intention performance. J Exp Psychol Learn Mem Cogn 2003;29:347-61. (DOI:10.1037/0278-7393.29.3.347)
5. Aberle I, Rendell PG, Rose NS, McDaniel MA, Kliegel M. The age prospective memory paradox: Young adults may not give their best outside of the lab. Dev Psychol 2010;46(6):1444-53. (DOI: 10.1037/a0020718)
6. Kliegel M, Jager T, Phillips L. Adult age differences in event based prospective memory: A meta-analysis on the role of focal versus nonfocal cues. Psychol Aging 2008;23:203-8. (DOI: 10.1037/0882-7974.23.1.203)
7. Niedźwieńska A, Rendell PG, Barzykowski K, Leszczyńska A. Only social feedback reduces age-related prospective memory deficits in "Virtual Week". Int Psychogeriatr 2014;26(5):759-67. (DOI:10.1017/S1041610214000027)
8. Bozdemir M, Cinan S. Age-related differences in intentional forgetting of prospective memory. Int J Aging Hum Dev 2021;92(3):350-63. (DOI:10.1177/0091415019900165)
9. Yörük A, Cangöz-Tavat B. Aging effect in prospective memory monitoring: An eye tracking study. Turkish Journal of Geriatrics 2022;25(4):640-9. (DOI: 10.31086/tjgeri.2022.323)
10. Haas M, Mehl MR, Ballhausen N, Zuber S, Kliegel M, Hering A. The sounds of memory: extending the age-prospective memory paradox to everyday behavior and conversations. J Gerontol B Psychol Sci Soc Sci 2022;77(4):695-703. (DOI: 10.1093/geronb/gbaa214)
11. Rendell PG, Thomson DM. Aging and prospective memory: Differences between naturalistic and laboratory tasks. J Gerontol B Psychol Sci Soc Sci 1999;54(4):256-69. (DOI: 10.1093/geronb/54B.4.P256)
12. Haines SJ, Randall SE, Terrett G et al. Differences in time-based task characteristics help to explain the age-prospective memory paradox. Cognition 2020;202:1-15. (DOI: 10.1016/j.cognition.2020.104305)
13. Rendell PG, Henry JD. A review of Virtual Week for prospective memory assessment: Clinical implications. Brain Impair 2009;10(1):14-22. (DOİ: 10.1038/s41582-021-00472-1)
14. Niedźwieńska A, Rendell PG, Barzykowski K, Leszczyńska A. Virtual week: Validity and psychometric properties of a Polish adaptation. Eur Rev Appl Psychol 2016;66(2):79-84. (DOİ: 10.1016/j.erap.2016.02.003)
15. Mioni G, Stablum F, Biernacki K, Rendell PG. Virtual week: Translation and adaptation for the Italian population. Neuropsychol Rehabil 2017;27(4):486-506. (DOI:10.1080/09602011.2015.1103758)
16. Hisli N. A study on validity and reliability test of the Beck Depression Scale. Turk J Psychol 1988;6(Suppl 22):118–22. (in Turkish)
17. Ertan T, Eker E. Reliability, validity, and factor structure of the geriatric depression scale in Turkish elderly: Are there different factor structures for different cultures. Int Psychogeriatr 2000;12(Suppl 2):163-72. (DOI: 10.1017/S1041610200006293)
18. Selekler K, Cangoz B, Uluc S. Power of discrimination of Montreal cognitive assessment (MoCA) scale in Turkish patients with mild cognitive impairment and Alzheimer's disease. Turkish Journal of Geriatrics 2010;13(Supple 3):166–71. (in Turkish)
19. Selekler K, Cangöz B, Karakoç E. Adaptation and standardization of the Functional Activities Questionnaire for Turkish culture in people aged 50 and over. Turk Norol Derg 2004;10(2):102-7. (in Turkish)
20. Karakaş S, Erdoğan E, Sak L, et al. Stroop Test TBAG Form: Standardisation for Turkish Culture, Reliability and Validity. Klinik Psikiyatri 1999;2(2):75-88. (in Turkish)
21. Cangoz B, Karakoc E, Selekler K. Trail Making Test: Normative data for Turkish elderly population by age, sex and education. J Neurol Sci 2009;283(1-2):73-8. (DOI: 10.1016/j.jns.2009.02.313)

22. EO Imamoglu. An interdependence model of human development, In: Cigdem Kagitcibasi (Ed). Growth and Progress in cross-cultural psychology. Lisse, Swets and Zeitlinger, The Netherlands 1987, pp 138-45.

23. Aközer M, Nuhrat C, Say Ş. Expectations Regarding Old Age in Türkiye. Sosyal Politika Çalışmaları Dergisi 2011;(27):103-28. (in Turkish)

24. Gryffydd L, Mitra B, Wright BJ, Kinsella GJ. Assessing prospective memory in older age: the relationship between self-report and performance on clinic-based and naturalistic tasks. Neuropsychol Dev Cogn B Aging Neuropsychol Cogn 2022;29(1):104-120. (DOI:10.1080/13825585.2020.1857327)

25. Park DC, Gutchess AH. Aging, cognition, and culture: a neuroscientific perspective. Neurosci Biobehav Rev 2002;26(7):859-67. (DOI: 10.1016/s0149-7634(02)00072-6)

26. Zuber S, Ihle A, Blum A, Desrichard O, Kliegel M. The effect of stereotype threat on age differences in prospective memory performance: differential effects on focal versus nonfocal tasks. J Gerontol B Psychol Sci Soc Sci 2019;74(4):625-32. (DOI:10.1093/geronb/gbx097)

27. Levent A, Davelaar EJ. Recreational drug use and prospective memory. Psychopharmacology (Berl). 2022;239(3):909-922. (DOI:10.1007/s00213-022-06081-0)