





Siri Vikram¹ 
 Sharon Ashley¹ 
 Gagan Bajaj¹ 
 Dasmine Fraclita D'Souza¹ 

Keywords

Prospective Memory
 Social Context
 Psychometric Properties
 Healthy Aging
 Cognitive Assessment

Evaluation of the Context-Based Prospective Memory (CBPM) task in neurotypical middle-aged and older adults: psychometric properties and sensitivity to age-related cognitive changes

ABSTRACT

Purpose: Prospective Memory (PM) is the ability to remember intentions for future execution which plays an important role in communication and is vulnerable to age-related decline. Context-Based Prospective Memory (CBPM) task, a novel test in Indian scenario evaluated psychometric properties among young adults. This study aimed to evaluate the psychometric properties of CBPM task among the neurotypical middle-aged and older adults and to investigate its sensitivity to age-related changes. **Methods:** A total of 80 participants (40 middle-aged, 40 older adults) completed the CBPM task and the Memory for Intentions Screening Test (MIST). Psychometric properties like internal consistency, test-retest reliability and convergent validity were evaluated. Age-related differences were examined using previously established CBPM task data from young adults. **Results:** CBPM task demonstrated strong internal consistency (Cronbach's $\alpha=0.875$ for middle-aged and 0.715 for older adults), Excellent test-retest reliability for total performance scores (ICC=0.93 and 0.81) for middle-aged and older adults respectively and error classification (ICC= 0.97 and 0.99). Similarly inter-rater reliability also indicated strong reliability for the total PM performance (ICC=0.91 and 0.85), and error classification (ICC=0.97 and 0.99). Significant convergent validity was also observed with MIST, suggesting that both the tasks perform similarly. On assessing the tool's sensitivity to age-related changes in PM abilities, significant declines were observed in middle-aged and older adults compared to young adults affirming the age-related cognitive deterioration. **Conclusion:** The CBPM task is an effective and reliable tool to measure PM across age groups, demonstrating strong psychometric properties and sensitivity to identify the age-related decline.

Correspondence address:

Gagan Bajaj
 Department of Audiology and Speech
 Language Pathology, Kasturba Medical
 College Mangalore, Manipal Academy
 of Higher Education
 Manipal, Karnataka 576104, India.
 E-mail: gagan.bajaj@manipal.edu

Received: May 20, 2025

Accepted: August 26, 2025

Editor: Ana Carolina Constantini.

Study conducted at Kasturba Medical College Mangalore - KMC Mangalore – Manipal (Karnataka), India.

¹ Department of Audiology and Speech Language Pathology, Kasturba Medical College Mangalore, Manipal Academy of Higher Education, Manipal, India

Financial support: nothing to declare.

Conflict of interests: nothing to declare.

Data Availability: Research data is available in the body of the article.



This is an Open Access article distributed under the terms of the Creative Commons Attribution license (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

INTRODUCTION

Prospective memory (PM), or the capacity to recall and achieve the aimed task at the right time⁽¹⁾, is a cornerstone of daily life. Whether it's delivering a message, making a scheduled phone call, or remembering to take medication, PM underpins our ability to navigate personal, social, and professional responsibilities. Among aging adults⁽²⁾, PM has a particularly pivotal role in preserving independence and ensuring wellness across various domains⁽³⁾. Unlike retrospective memory, which involves recalling past events or information, PM is characterized by self-initiated retrieval processes, where one must independently act on a prior intention when prompted by a relevant cue, rather than external demands to remember⁽⁴⁾. From attending appointments to paying bills, PM tasks are deeply woven into the fabric of everyday functioning, underscoring their vital importance in preserving an individual's welfare and independence⁽⁵⁾.

Beyond its role in everyday tasks, PM is also fundamental to effective communication⁽⁶⁾. It enables crucial activities like recalling factual details in conversations, adhering to response timelines through email, initiating or answering telephone calls promptly or re-dialing without delay, being available for crucial meetings or planned commitments, and relaying critical messages accurately⁽⁶⁾. PM skills are also needed to fit into social norms, i.e., greeting others, making eye contact, and engaging appropriately in dialogues thus making communication more effective and stronger. PM's contribution to communication is indispensable in meeting social expectations and fostering meaningful relationships, both personal and professional. Communication related failure in PM can result in misunderstandings. Compared to retrospective memory, PM carries a moral weight; failures are often perceived as signs of unreliability, potentially damaging an individual's credibility and self-esteem. This highlights the profound social importance of PM in maintaining trust and meaningful interactions⁽⁶⁾.

PM is particularly sensitive to age-related cognitive decline, and several studies have demonstrated that cognitive performance deteriorates with age⁽⁷⁾. These deficits are not only evident in healthy aging, but also well documented in cognitive-communication conditions, like mild cognitive impairment (MCI), dementia, and traumatic brain injury (TBI)⁽⁸⁾. Given its functional relevance, PM has increasingly been advocated as an essential component in cognitive-communication assessments much like other cognitive zones such as working memory, short-term memory, and executive function.

Neuroanatomically, age-related declines in PM performance have been connected to structural and functional modifications in key brain regions responsible for the formation, retention, initiation and execution of intentions⁽⁹⁾. Two distinct frontoparietal networks are considered critical for PM. The dorsal frontoparietal network is primarily involved in the formation and retention of intention by allowing the individual to keep the future tasks in mind while engaging in ongoing activities, primarily through strategic monitoring mechanisms⁽¹⁰⁾. While ventral frontoparietal network facilitates the initiation and execution of intentions by enabling spontaneous retrieval of intended actions in response to relevant internal or external cues⁽¹⁰⁾. Another key region implicated in PM is the anterior prefrontal cortex (aPFC),

which plays an important role in PM function. Age-related structural decline in the aPFC has been shown to negatively impact PM functioning⁽¹⁰⁾. These neuroanatomical changes help explain the observed behavioural impairments in PM among older individuals and highlight the need for assessment tools sensitive to such age-related shifts in brain function.

A wide array of tools is currently utilized to assess and measure PM, including experimental tasks, single-trial approaches and self-report questionnaires as well as paper-and-pencil test batteries⁽¹¹⁾. Test batteries include the Rivermead Behavioural Memory Test (RBMT)⁽¹²⁾, the Cambridge Test of Prospective Memory (CAMPROPT)⁽¹³⁾, the Memory for Intentions Screening Test (MIST)⁽⁶⁾, and the Royal Prince Alfred Prospective Memory Test (RPA-Pro-Mem)⁽¹⁴⁾. Single-execution protocols include the Envelope task⁽¹⁵⁾, the Prompt card task, the Telephone test⁽¹¹⁾, and the Key test⁽¹⁶⁾. Questionnaires include the Prospective Memory Questionnaire (PMQ)⁽¹⁷⁾, the Prospective and Retrospective Memory Questionnaire (PRMQ)⁽¹⁸⁾, the Comprehensive Assessment of Prospective Memory Questionnaire (CAPM)⁽¹⁹⁾, and the Brief Assessment of Prospective Memory Questionnaire (BAPM)⁽²⁰⁾. Analytical and investigative paradigms consist of the⁽²¹⁾, the Test Écologique de Mémoire Prospective (TEMP)⁽²²⁾, the Virtual Week, and Actual Week⁽²³⁾.

Despite the variety of tools available to assess PM, traditional test batteries like RBMT and MIST, often lack the ecological validity as they are administered in controlled environments that do not accurately reflect real-world cognitive demands. In everyday life, PM tasks are associated with specific environmental and situational cues. Since contextual relevance is most effective when aligned with cultural norms, incorporating culturally appropriate contexts during PM assessment may reduce cognitive load and improve PM performance⁽²⁴⁾. Despite the fact that PM capacities are an essential component in many cognitive testing batteries, there exists a notable lack of culturally relevant, context-based PM tests with adequate reliability and validity in India.

Acknowledging these limitations, a novel context-based PM task (CBPM)⁽²⁵⁾ was developed to provide a culturally relevant and ecologically valid tool for assessing PM in an Indian context, building on the framework established by MIST. It was validated among young adults in accordance with the assessment parameters outlined in MIST. This innovative task incorporates a virtual shopping mall environment, simulating real-world scenarios to evaluate PM abilities effectively. It comprises both an ongoing activity and a PM component, designed to mirror everyday cognitive demands. Its psychometric properties are being systematically evaluated through a multiphase approach. The first phase involved the development of the task and its initial validation in a young adult population⁽²⁵⁾. To further establish CBPM task as a robust assessment tool, it is crucial to evaluate its psychometric properties in middle-aged and older adults, as well as to examine its sensitivity to age-related cognitive changes, prior to its application in clinical populations or individuals with cognitive-communication disorders. Hence, as a second phase, the current study aims to establish the psychometric properties of the CBPM task, specifically assessing its internal consistency, test-retest reliability, inter-rater reliability and convergent validity among typically aging middle-aged adults, old-aged adults, and to determine its utility in detecting age-related changes in PM.

The PM task is non-focal in essence, requiring individuals to perform a designated action, either verbally or nonverbally, at a specific instant or place within the mall. A total of ten PM tasks are administered, categorized across three dichotomies: the cue of the stimuli (event-driven or time-driven, i.e., implementing an aimed activity at a pre-established time or place within the mall), response format (action or spoken, i.e., performing a nonverbal or verbal intention), and the length of the delay (short or lengthy duration, i.e., less than or more than 4 minutes between the delivery of the intention and its performance). The cognitive load is maintained between 1 and 5 PM targets. Each PM trial is described along with its presentation sequence, associated cognitive load, and classification under the three dichotomies, as depicted in Table 1.

The administration of CBPM task, in the present research, followed the same protocol outlined in D'Souza et al.⁽²⁵⁾, with a practice trial done prior to the main task. Participants received the following instructions:

You will be shown various scenes from a mall, each containing random numbers and words. Your task is to identify numbers that include the combination '81' (for example, 7814, 481, 812). While doing this, I will assign you additional tasks that you must remember and perform at the correct time, without any reminders. A digital clock is provided to help you track time for these tasks.

The investigator recorded the outcome of each of the subjects on both the ongoing tasks and PM trials, noting any errors made.

The scoring criteria for the CBPM task, in the present work, is identical to that defined in D'Souza et al.⁽²⁵⁾. Accuracy for PM targets was assessed on a scale of '0,' '1,' or '2,' depending on task type. For time-based tasks, a timely execution was scored as '2,' while a response with a slight timing error (within ± 30 seconds) was scored as '1,' and missed responses received a score of '0.' For event-based tasks, accurate execution was scored as '2,' while missed responses were marked as '0.' Six subscale measurements were calculated with respect to the predefined dichotomies: short-duration tasks, long-duration tasks, time-based tasks,

event-based tasks, action responses, and verbal responses, and a total overall score of PM was derived by summing the scores of subscales with a maximum possible score of 60.

Additionally, the type of errors displayed by participants were documented and classified based on the MIST error taxonomy. The first type of error, termed Prospective Memory Failure (PF), was recorded when a member was completely unsuccessful to respond to a PM cue. The second error type, known as a Task Substitution (TS) error, was noted when the participant performed an action suitably but executed a task intended for a different cue. The third error, Loss of Content (LC), was assigned when the participant recognized the PM cue but was unable to recall part or all of the intended action. The fourth, Loss of Time (LT) error, was marked when the correct task was performed, but at an incorrect time point. Finally, a Random Error (RE) was coded for any responses that did not fit into the previous error categories. Each error type is associated with a specific score.

Finally, performance of the ongoing task was evaluated independently, where each correct detection of the target ('81') was awarded a score of '1,' yielding a maximum possible score of 138. PM performance, error characteristics, and ongoing task outcomes were collectively used to assess the participant's cognitive functioning within the CBPM task.

MIST

Consistent with the protocol followed by D'Souza et al.⁽²⁵⁾ MIST was employed in the present study to establish the convergent validity of CBPM task in middle aged and older adults. MIST was developed by Raskin⁽⁸⁾ and is considered a standardized measure to assess PM. It comprises time-based and event-based PM tasks embedded within an ongoing word search activity, requiring participants to remember and execute specific intentions at designated times or in response to specific cues. It demonstrates good psychometric properties, with established internal consistency, inter-rater reliability, and construct validity, and has been shown to effectively differentiate between healthy adults and clinical populations with memory impairments.

Table 1. PM trials, the presentation sequel, and the corresponding cognitive load

Trails	Instructions	Cue	Response modality	Time duration	Order of execution	Cognitive load
1	"When you come across seafood, count and tell the me the number of salesmen present."	Event-based	Verbal	Short	1	4
2	"Ask me if I need anything from the mall, 10 minutes from now"	Time-based	Verbal	Long	4	5
3	"When you are at the section where you can buy lollipops, point out any object that is green in colour"	Event-based	Action	Short	2	5
4	"After a duration of 11 minutes from now, let me know the name of the store you are situated in"	Time-based	Verbal	Long	6	4
5	"Write the word 'GOLDFISH' when you reach the section where you see T shirts"	Event-based	Action	Long	5	4
6	"In the next 3 minutes, write down your phone number on the sheet/notepad next to you"	Time-based	Action	Short	3	5
7	"When you reach the section where you can order pizza, remember to show the me 'J' key on the keyboard"	Event-based	Action	Long	8	3
8	"After 8 minutes from now, show me a thumbs up"	Time-based	Action	Long	9	2
9	"In 2 minutes from now, name 3 items you can see in the room"	Time-based	Verbal	Short	7	4
10	"Tell the name of your bank, when you see a white car"	Event-based	Verbal	Short	10	1

Note. Adapted from the earlier version of the CBPM task, used in a study among young adults⁽²⁵⁾, with a balanced representation of PM tasks across dichotomies

Procedure

The stimulus of the shopping mall scenes was presented through a LENOVO IdeaPad 3 with a 15-inch display. The test was conducted in a quiet environment. Each participant had to complete two tasks: the CBPM task and the MIST. The sequence of task administration was counterbalanced, with one half of participants beginning with the CBPM task and the other half with the MIST, ensuring consistency across both the study groups. The overall time taken for the sessions was approximately 45 minutes, with the CBPM task taking 25 minutes and MIST around 20 minutes. A digital clock was positioned in front of each participant to aid in time management for time-based tasks, while pen and paper were provided for tasks necessitating written responses. The examiner delivered instructions verbally, and participants' responses were recorded in real time.

Data analysis

A statistical analysis was done using the Jamovi software (2.3.28) and IBM SPSS (version 29). Mean and Standard deviation were obtained through descriptive statistics for each subscale of MIST and the CBPM task. The reliability of the subscales was assessed by examining internal consistency using both Cronbach's alpha and Spearman-Brown's split-half reliability analysis. Additionally, Spearman's correlation was applied to examine the connection between individual trials and subscales of the CBPM task. Convergent validity was assessed across middle-aged adults and older adults by comparing their performance on CBPM task and MIST scores using McNemar's test. To further support convergent validity, an error code correlation analysis was conducted between the MIST and CBPM task using Spearman's correlation coefficient. The reliability of the PM scores from the CBPM task were evaluated using both test-retest and inter-rater reliability measures for the total performance scores as well as error scores. These reliability metrics were quantified using intraclass correlation coefficients (ICCs).

To evaluate the utility of the CBPM task in detecting age-related changes in PM, data from young adults reported by D'Souza et al.⁽²⁵⁾ were compared with middle-aged and older adult data obtained in the present study. Shapiro-Wilk's test was done to examine the normality of the data. Further, Kruskal-Wallis's test was done to study the age-linked effects

between the three groups. Post hoc analysis using Dwass-Steel-Critchlow-Fligner was done to identify the differences between the different age groups. Additionally, ongoing task accuracy was analysed using one-way ANOVA, followed by Games-Howell post hoc comparisons.

RESULT

Psychometric properties of CBPM task

Descriptive performance of the middle-aged and older adults on CBPM task across scales and the total performance scores has been provided in Table 2. Among the middle-aged adults, the short-duration task (M=6.13), time-based task (M=5.13) and the verbal response task (M=5.15) were observed to have the highest scores, while the long-duration task (M=3.98), event-based task (M=4.97) and the action response task (M=4.95) showed comparatively lower scores. A similar pattern was seen in older adults, where performance of short-duration task (M= 7.55), time-based task (M= 6.67) and verbal response task (M= 6.72) was relatively better than long-duration task (M= 3.15), event-based task (M=4.03) and action response task (M=3.98).

The percentage distribution of error types in the CBPM task across middle-aged and older adults is presented in Tables 3 and 4. Among the error types, PF and LC were the most frequent across both age groups, though PF was notably higher in older adults. LT errors were also very evident in older adults as compared to middle-aged adults. TS, PL, and RE were relatively infrequent in both groups. Overall, the error distribution reflects age-related trends in PM performance, with older adults exhibiting more cue detection and timing difficulties compared to middle-aged adults.

Consistency and Reliability

To determine the internal consistency, analyses were conducted using Cronbach's alpha and the Spearman-Brown split-half method. The analysis demonstrated a robust internal stability among the six subscales, as reflected by a Cronbach's alpha coefficient of 0.875 for middle-aged adults and 0.715 for older adults. Furthermore, the Spearman-Brown coefficient was found to be 0.784 and 0.952 for middle-aged and older aged adults respectively, indicating a good split-half reliability.

Table 2. Mean, median, and standard deviation of the subscales wise and total performance on CBPM task of middle-aged and older adults' groups

Subscale	Max	Mean		Median		SD	
		MA	OA	MA	OA	MA	OA
Short-duration task	10	6.13	7.55	6	8	2.10	1.28
Long-duration task	10	3.98	3.15	4	3	2.14	1.27
Time-based task	10	5.13	6.67	5	7	1.60	1.1
Event-based task	10	4.97	4.03	5	4	2.58	1.17
Verbal-response task	10	5.15	6.72	5	7	1.85	1.4
Action-response task	10	4.95	3.98	5	4	2.19	1.31
Total score	60	30.3	32.1	30	33	9.89	3.2
Ongoing Accuracy	138	95.8	86.1	94	87	17.3	9.93

MA-Middle Age; OA-Old Age; SD-Standard Deviation

Table 3. Percentage of occurrence of each error type in the PM trials of the context-based PM task in middle-aged adults

TRIALS	NE	LC	LT	PF	TS	RE	PL
1	16	7	2	12	2	0	1
2	11	10	0	11	1	0	7
3	24	5	6	3	2	0	0
4	11	4	4	21	0	0	0
5	12	11	1	13	2	0	1
6	3	4	3	29	1	0	0
7	29	2	3	5	1	0	0
8	29	5	0	2	2	0	2
9	17	3	5	13	2	0	0
10	28	5	0	5	2	0	0

NE- No Error; PF-Prospective memory failure; TS-Task substitution; LC-Loss of content; LT-Loss of time; RE- Random error; PL-Place Losing

Table 4. Percentage of occurrence of each error type in the PM trials of the context-based PM task in old age adults

TRIALS	NE	LC	LT	PF	TS	RE	PL
1	27	11	0	2	0	0	0
2	14	9	0	14	0	0	3
3	34	2	2	2	0	0	0
4	18	8	5	9	0	0	0
5	2	16	0	21	0	1	0
6	2	7	6	25	0	0	0
7	40	0	0	0	0	0	0
8	2	17	1	20	0	0	0
9	17	4	13	6	0	0	0
10	35	0	3	1	0	1	0

NE- No Error; PF-Prospective memory failure; TS-Task substitution; LC-Loss of content; LT-Loss of time; RE- Random error; PL-Place Losing

Table 5. Spearman's rank correlation between the performance on each trial and the subscale scores

Trials	SD	LD	TB	EB	V	A	PMT
1	0.645*	-0.104	0.051	0.506*	0.539*	-0.039	0.447*
2	0.494*	0.039	0.016	0.496*	-0.011	0.567*	0.431*
3	0.461*	-0.108	0.426*	-0.035	0.199	0.128	0.284*
4	0.123	0.400*	0.590*	0.035	0.591*	-0.059	0.410*
5	-0.055	0.294*	-0.174	0.335*	-0.119	0.372*	0.190
6	-0.027	0.263*	0.329*	-0.022	0.360*	-0.125	0.157
7	0.430*	-0.262	0.189	0.073	0.290*	-0.110	0.188
8	-0.299	0.453*	-0.341	0.372*	-0.347	0.486*	0.042
9	-0.070	0.534*	0.442*	-0.028	-0.023	0.447*	0.310
10	0.295*	0.049	0.210	0.213	0.572*	-0.196	0.288*

*p<0.05

Spearman's correlation analysis examined the relationship between individual trial performance and the corresponding subscale scores. The results, as presented in Table 5, revealed a statistically significant correlation in most cases, with a few exceptions (action response in trial 3, time cue in trial 7, and event cue in trial 10).

Test-Retest Reliability and Inter-Rater Reliability

To assess the reliability, 20% of the total sample (n=8 middle-aged adults; n=8 older adults) was selected. Test-retest reliability was assessed by re-administering the task after a two-week interval, while inter-rater reliability was determined based on independent ratings provided by two trained assessors.

Test-retest reliability was examined using intraclass correlation coefficients (ICCs). For the total performance scores, ICCs were 0.931 [95% CI: 0.69–0.98; F(7, 7.1) = 24.8, p < 0.001] in middle-aged adults and 0.813 [95% CI: 0.36–0.95; F(7, 7.5) = 10.8, p < 0.005] in older adults, indicating good test-retest reliability. For error classification, ICCs were 0.977 [95% CI: 0.98–0.99; F(6, 7) = 692, p < 0.001] in middle-aged adults and 0.997 [95% CI: 0.98–0.99; F(6, 7) = 573, p < 0.001] in older adults, reflecting excellent test-retest reliability.

Inter-rater reliability for total performance scores yielded ICCs of 0.917 [95% CI: 0.67–0.98; F(7, 8) = 23.1, p < 0.001] for middle-aged adults and 0.846 [95% CI: 0.45–0.96; F(7, 7.9) = 12, p < 0.001] for older adults, indicating strong inter-rater reliability.

For error classification, ICCs were 0.977 [95% CI: 0.88–0.99; $F(6, 7) = 84.9, p < 0.001$] and 0.997 [95% CI: 0.98–0.99; $F(6, 7) = 505, p < 0.001$] in middle-aged and older adults, respectively, demonstrating excellent inter-rater reliability.

Convergent validity

To assess convergent validity, participants' performance on the CBPM task was compared with their scores on MIST. McNemar's test was conducted to evaluate the alignment of the two measures in assessing PM abilities. To distinguish between high and low performance on each task, scores were categorized into two groups based on the median values for both tasks. The results of the test indicated no statistically significant difference between the proportion of high and low performers across MIST and CBPM task scores in the middle-aged individuals ($\chi^2=0.6, p=0.439$) and old -aged adults ($\chi^2=2.33, p=0.127$), thus supporting convergent validity.

Additionally, Spearman's correlation investigation was conducted to examine the relationship among various categories of errors made by participants in the MIST and CBPM tasks across the two groups. The results showed a high positive correlation between the totals of different errors in both tasks, with the strongest correlation observed in the older adult group ($\rho = 0.937, p = 0.002$), followed by middle-aged adults ($\rho = 0.919, p = 0.003$).

Utility of CBPM task in assessing age-related changes in PM abilities

To evaluate age-related changes in PM, the total performance scores of CBPM task from young adults⁽²⁵⁾ were compared with the total performance scores of middle-aged and older adults in the present study, with Shapiro-Wilk tests indicating non-normality across all groups ($p < 0.001$).

A Kruskal-Wallis test revealed a statistically difference in PM performance among the three ages ($\chi^2 = 30.7, p < 0.001$). Post hoc analysis using the Dwass-Steel-Critchlow-Fligner test showed that young-adults performed significantly better than middle-aged ($W = 6.17, p < .001$) and older-aged adults ($W = 7.02, p < 0.001$), however, no significant difference was found between middle-aged and older adults ($p = 0.28$) (Figure 2).

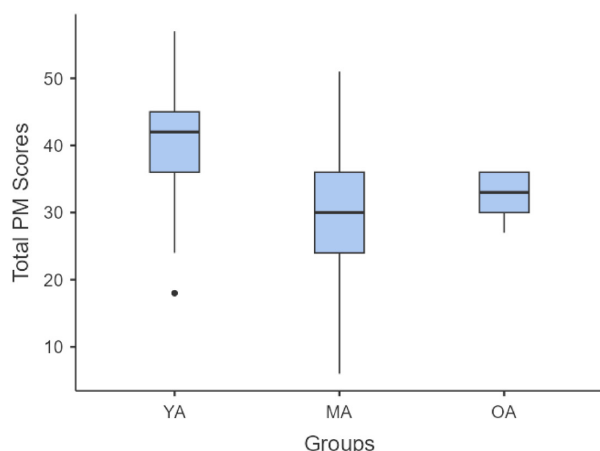


Figure 2. Box plot depicting the distribution of the total PM scores across young adults, middle-aged adults and older adults

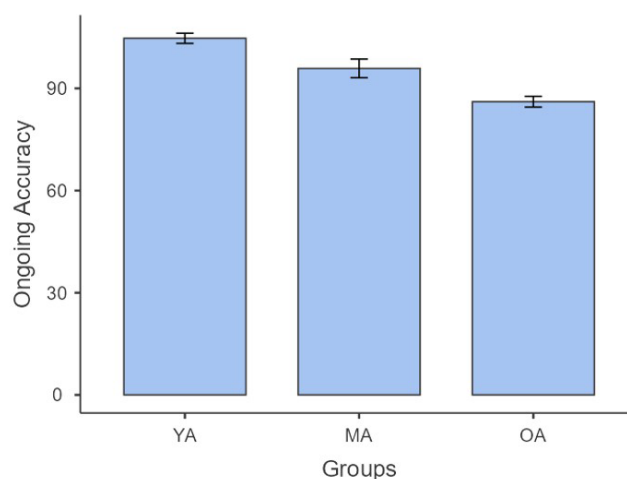


Figure 3. Bar graph depicting the distribution of the ongoing accuracy scores across young adults, middle-aged adults and older adults

For the ongoing task scores, the Shapiro-Wilk test indicated normal distribution across groups. Welch's one-way ANOVA was performed indicating a notable distinction among the groups, $F(2, 78) = 36.8, p < .001$. Games-Howell post hoc comparisons indicated that young adults outperformed both middle-aged adults ($M = 8.85, p = 0.017$) and older adults ($M = 18.63, p < 0.001$). Additionally, middle-aged adults scored significantly higher than older adults ($M = 9.77, p = 0.008$), demonstrating a clear age-related decline in task performance across the three groups (Figure 3).

DISCUSSION

Psychometric properties of CBPM task

The CBPM task was developed to address the lack of culturally and contextually relevant PM assessment tools suitable for the Indian population. Existing PM tasks often overlook cultural and environmental factors that influence memory performance. Drawing inspiration from studies like Ouellet⁽³⁰⁾, which emphasized the effectiveness of virtual shop-based assessments for evaluating memory domains, the CBPM task was designed using a familiar shopping mall scenario featuring diverse locations. Initially validated among young adults⁽²⁵⁾, the current research aims to establish the psychometric properties of the CBPM task in neurotypical middle-aged adults, older adults and explore its utility in detecting age related PM changes. By a methodical examination of internal consistency, reliability, validity, and group differences in performance, the current research adds to the emerging PM test literature, addressing the lack of contextually relevant PM tasks.

The validation of this measure and the establishment of its psychometric properties in middle-aged and older adults who are healthy cognitively were required to confirm its reliability and broader applicability.

Descriptive performance of CBPM task performance revealed that middle-aged and older adults each demonstrated strengths in different task. Interestingly, the mean total performance score was slightly higher in older adults ($M = 32.1$) compared to middle-aged adults ($M = 30.3$). Older adults also seemed to

perform better on short duration tasks, time-based tasks and tasks that required verbal responses. This marginal advantage may reflect increased motivation among older adults, potentially driven by heightened awareness of age-related cognitive changes⁽⁶¹⁾. Conversely, middle-aged adults showed superior performance on long-duration, event-based and action-response tasks, which are more dependent on self-initiated retrieval processes.

Assessing the internal consistency and reliability of the CBPM task is essential for determining the effectiveness of the assessment tool⁽⁶²⁾. In this study, reliability was measured through Spearman-Brown split-half method and Cronbach's alpha among middle-aged and older adults. The CBPM task established high internal consistency and reliability with Cronbach alpha of 0.875 among middle-aged and 0.715 among older adults, confirming that the instrument was effective in the measurement of PM performance among the two groups. The high internal consistency observed across subscales suggests that the test components measure a cohesive construct, making CBPM task a reliable assessment tool. This also corresponds with earlier findings of a strong internal consistency among young adults using CBPM task⁽²⁵⁾. Additionally, split-half reliability analysis further supports the robustness of the test, indicating that both halves of the assessment produce comparable results, consistent with the findings stated by D'Souza et al.⁽²⁵⁾. Further, the relationship between individual trial performance and subscale scores was examined through Spearman's correlation analysis to validate the CBPM task's ability to measure PM performance consistently. The strong correlations between individuals trial performance and overall subscale scores reinforce the validity of the CBPM task in capturing PM abilities in neurotypical adults of this study. These findings again resonate with the previous research on young adults using this tool⁽²⁵⁾.

The test-retest reliability results support the stability of the CBPM task over time, reinforcing its utility in assessing PM performance. The high intra-class correlation coefficients observed for total PM performance and error classification indicate strong test-retest between initial and follow-up assessments. This consistency suggests that CBPM task reliably captures individual differences in PM abilities across repeated administrations. Similar conclusions have been drawn in past studies^(8,25). Notably, the error classification component exhibited exceptionally high reliability, highlighting its robustness in distinguishing different error types. These findings confirm that the task yields stable and reliable results across repeated administrations, demonstrating good test-retest reliability for both total PM performance and error classification.

Inter-rater reliability for the total PM performance and error classification was high, suggesting consistent ratings across observers. These findings suggest that the rating criteria were well-defined and consistently applied, bolstering the credibility of the observational data. The high reliability across both the total PM performance and error metrics enhances confidence in the validity of the CBPM task and supports the robustness of the results derived from these measures. The strong inter-rater reliability observed in the present study mirrors findings by D'Souza et al.⁽²⁵⁾ among younger adults, suggesting that the rating procedures are robust across age groups.

Inspired by MIST in terms of its dichotomies, ongoing task component, and established psychometric properties, the CBPM task was developed to provide a comprehensive assessment of PM. Convergent validity analysis further supports its robustness by confirming statistical alignment with these measures. The comparison between CBPM task and MIST scores using McNemar's test indicated no significant differences across middle-aged adults and older adults, suggesting that both measures assess PM abilities similarly. This alignment, therefore, reinforces the construct validity of CBPM task as a reliable PM assessment tool. Furthermore, Spearman's correlation analysis revealed a strong positive relationship between error patterns in CBPM task and MIST, with the high correlation observed in older adults, followed by middle-aged adults. The significant correlations across all groups indicate that CBPM task effectively captures PM deficits in a manner consistent with established measures. The earlier study among the young adults using this tool confirms this pattern⁽²⁵⁾, making it a valid tool among the different age groups.

Error analysis also revealed that LC errors and PF errors were the most frequent errors seen in the two groups. The high occurrence of LC errors may be attributed to the heavy reliance of the content component of PM on retrospective memory⁽³³⁾. Whereas PF errors- representing a complete failure to remember and execute the intended task, can be linked to the overall cognitive deficits, affecting multiple cognitive domains essential for successful PM outcomes⁽³³⁾.

Utility of CBPM task in assessing age-related changes in PM abilities

The comparative analysis of CBPM task performance across the three groups provides critical insights into the trajectory of PM decline. Comparing the results of the total scores of PM with younger adults, it showed that middle-aged and older adults performed significantly lower than younger adults, reflecting the pronounced age-linked decline in PM abilities. However, the change between middle-aged and older adults was not very significant. This pattern suggests that age-related declines in PM may begin to manifest during midlife and follow a gradual trajectory⁽³⁴⁾. This decline is consistent with age-related changes in brain regions involved in PM, particularly the aPFC and the frontoparietal networks. These areas are responsible for maintaining and executing future intentions and atrophy of these areas can impair the strategic monitoring and cue detection required for successful PM resulting in poor scores with aging⁽¹⁰⁾.

Similarly, the analysis of ongoing task scores revealed significant distinctions across the three groups. However, when considering ongoing task accuracy, middle-aged adults outperformed older adults ($M = 95.8$ vs. 86.1) unlike the mean total performance scores, suggesting that while older adults may prioritize the prospective component of the task, their ability to maintain accuracy on the ongoing task may be compromised owing to age-related declines in attention and processing speed^(35,36). Age-associated degeneration of cortical areas like posterior parietal cortex and dorsolateral prefrontal cortex can lead to difficulties in dividing and sustaining attention across different tasks⁽³⁷⁾. This is also seen in the present study where

the performance of young adults was found to be higher than that of both middle-aged and older adults indicating a gradual deterioration in cognitive function across adult life. As these brain regions decline with age, the ability to coordinate between ongoing tasks and future intentions also weakens, contributing to poorer PM performance in older adults⁽³⁸⁾.

In conclusion, the CBPM task developed in this research proves to be a reliable and valid tool for assessing PM abilities comprehensively. Its excellent psychometric properties, as in high reliability, validity, and well-designed error analysis, establish it as a robust assessment measure and the task design incorporated different cognitive loads, retrieval demands, and response formats, allowing for a comprehensive evaluation of PM performance. The integration of an ongoing task, along with PM accuracy evaluation, ensures that the tool closely reflects real-world PM demands. Furthermore, the shopping mall context enhances ecological validity, making the task engaging and relatable for participants. Given its holistic approach and strong methodological foundation, this tool holds significant potential for cognitive-communicative assessments and could be further adapted for diverse populations and clinical applications.

Limitations and future directions

A key limitation of the current study lies in the generalizability of the CBPM task, which was examined only in neurotypical middle-aged and older adults, as a continuation of earlier validation conducted among young adults. While this is a critical step in establishing normative data and understanding age-related changes in PM performance, the findings do not extend to individuals from diverse socioeconomic and educational backgrounds or to clinical populations. Consequently, the results cannot yet inform the severity of performance impairments in populations with cognitive-communication disorders. Future studies are planned to validate the CBPM task across broader demographic profiles and clinical groups, such as individuals with MCI, to enhance understanding of PM performance and expand the tool's utility for early detection and intervention. Considering that the CBPM task is currently in the formative stage of development and its construct validity is still being established through a series of studies, methodological details, visual references, and illustrative examples are provided in this manuscript to ensure transparency and support critical evaluation of the task design. Although studies on its psychometric properties in clinical populations are underway to enhance its clinical utility, the CBPM task is presently available for research purposes and can be obtained upon reasonable request from the corresponding author.

CONCLUSION

This study aimed to extend the validation of CBPM task to middle aged and older adults, establishing strong psychometric properties, including high internal consistency, test-retest reliability and construct validity, as with prior established findings in young adults. The results support that CBPM task is an effective and reliable instrument to measure PM across the adult life span.

Its capacity to differentiate age related performance patterns and its strong correlation with established measures like MIST highlight its clinical utility for normative assessment. The utilization of real-world context and varied cognitive demands makes CBPM task representative of functional memory capabilities, and therefore a useful addition to neuropsychological batteries.

REFERENCES

1. Kliegel M, Ballhausen N, Hering A, Ihle A, Schnitzspahn KM, Zuber S. Prospective memory in older adults: where we are now and what is next. *Gerontology*. 2016;62(4):459-66. <https://doi.org/10.1159/000443698>. PMID:26950339.
2. McDaniel MA, Einstein GO. *Prospective memory: an overview and synthesis of an emerging field*. Thousand Oaks: Sage Academic Books; 2007.
3. Henry JD. Prospective memory impairment in neurological disorders: implications and management. *Nat Rev Neurol*. 2021;17(5):297-307. <https://doi.org/10.1038/s41582-021-00472-1>. PMID:33686303.
4. Anderson FT, McDaniel MA. Retrieval in prospective memory: multiple processes or just delay? *Q J Exp Psychol (Hove)*. 2019;72(9):2197-207. <https://doi.org/10.1177/1747021819845622>. PMID:30957661.
5. Penningroth SL, Scott WD. Prospective memory tasks related to goals and concerns are rated as more important by both young and older adults. *Eur J Ageing*. 2013;10(3):211-21. <https://doi.org/10.1007/s10433-013-0265-9>. PMID:28804296.
6. D'Souza DF, Bajaj G, Kotian H, Moolambally SR, Bhat JS. Communication-related aspects of prospective memory: an exploratory factor analysis of prospective memory questionnaires. *CoDAS*. 2024;36(4):e20230233. <https://doi.org/10.1590/2317-1782/20242023233en>. PMID:38775527.
7. Vallesi A, Cona G, Guo Y, Koo YW, Neumann DL, Ownsworth T, et al. Revisiting the age-prospective memory paradox using laboratory and ecological tasks. *Front Psychol*. 2021;12:691752. <https://doi.org/10.3389/fpsyg.2021.691752>. PMID:34220653.
8. Raskin SA. Memory for intentions screening test: psychometric properties and clinical evidence. *Brain Impair*. 2009;10(1):23-33. <https://doi.org/10.1375/brim.10.1.23>.
9. Craik FIM, Henry JD. Age-related changes in everyday prospective memory. In: Logie R, Cowan N, Gathercole S, Engle RW, Wen Z, editors. *Memory in science for society: there is nothing as practical as a good theory*. Oxford: Oxford Academic; 2023. p. 325-53.
10. Henry JD, Coundouris SP, Labuschagne I, Liu K, Haines SJ, Grainger SA, et al. Age-related differences in neural integrity are unrelated to prospective memory age effects. *Brain Cogn*. 2025;187:106301. <https://doi.org/10.1016/j.bandc.2025.106301>. PMID:40311230.
11. Blondelle G, Sugden N, Hainselin M. Prospective memory assessment: scientific advances and future directions. *Front Psychol*. 2022;13:958458. <https://doi.org/10.3389/fpsyg.2022.958458>. PMID:35967621.
12. Wilson BA, Cockburn J, Baddeley A. *Rivermead behavioural memory test* [Internet]. London: Thames Valley Test Company; 1985 [cited 2024 Feb 4]. Available from: <https://www.scirp.org/reference/referencespapers?referenceid=1219054>
13. Thöne-Otto AIT, Streubel B. The Cambridge Prospective Memory Test (CAMPROMPT). *Z Neuropsychol*. 2006;17(3):201-3. <https://doi.org/10.1024/1016-264X.17.3.201>.
14. Radford KA, Lah S, Say MJ, Miller LA. Validation of a new measure of prospective memory: the Royal Prince Alfred Prospective Memory Test. *Clin Neuropsychol*. 2011;25(1):127-40. <https://doi.org/10.1080/13854046.2010.529463>. PMID:21108144.
15. Huppert FA, Johnson T, Nickson J. High prevalence of prospective memory impairment in the elderly and in early-stage dementia: findings from a population-based study. *Appl Cogn Psychol*. 2000;14(7):14. <https://doi.org/10.1002/acp.771>.
16. Babicz MA, Sullivan KL, Weinborn M, Bucks RS, Ng A, Hodgson E, et al. Does the key task measure prospective memory? Cautionary findings from parallel studies in HIV disease and older adults. *Arch Clin Neuropsychol*. 2019;34(8):1438-44. <https://doi.org/10.1093/arclin/acz009>. PMID:30844064.

17. Hannon R, Adams P, Harrington S, Fries-Dias C, Gipson MT. Effects of brain injury and age on prospective memory self-rating and performance. *Rehabil Psychol*. 1995;40(4):289-98. <https://doi.org/10.1037/0090-5550.40.4.289>.
18. Smith G, Della Sala S, Logie RH, Maylor EA. Prospective and retrospective memory in normal ageing and dementia: a questionnaire study. *Memory*. 2000;8(5):311-21. <https://doi.org/10.1080/09658210050117735>. PMID:11045239.
19. Chau LT, Lee JB, Fleming J, Roche N, Shum D. Reliability and normative data for the comprehensive assessment of prospective memory (CAPM). *Neuropsychol Rehabil*. 2007;17(6):707-22. <https://doi.org/10.1080/09602010600923926>. PMID:17852758.
20. Man DWK, Fleming J, Hohaus L, Shum D. Development of the Brief Assessment of Prospective Memory (BAPM) for use with traumatic brain injury populations. *Neuropsychol Rehabil*. 2011;21(6):884-98. <https://doi.org/10.1080/09602011.2011.627270>. PMID:22150453.
21. Titov N, Knight RG. A video-based procedure for the assessment of prospective memory. *Appl Cogn Psychol*. [Internet]. 2020 [cited 2025 Feb 21];15(1):61-83. Available from: https://www.researchgate.net/publication/246857147_A_Video-based_Procedure_for_the_Assessment_of_Prospective_Memory
22. Potvin MJ, Rouleau I, Audy J, Charbonneau S, Giguère JF. Test écologique de mémoire prospective. *PsycTESTS Dataset*. Washington: APA Psyc Tests; 2021. <https://doi.apa.org/doi/10.1037/t79705-000>.
23. Rendell PG, Craik FIM. Virtual week and actual week: age-related differences in prospective memory. *Appl Cogn Psychol*. 2000;14(7):14. <https://doi.org/10.1002/acp.770>.
24. Smith RE, Hunt RR, Murray AE. Prospective memory in context: moving through a familiar space. *J Exp Psychol Learn Mem Cogn*. 2017;43(2):189-204. <https://doi.org/10.1037/xlm0000303>. PMID:27504679.
25. D'Souza DF, Ashley S, Bajaj G, Moolambally SR, Bhat JS. Development and validation of a novel Context-Based Prospective Memory Task among neurotypical adults. *CoDAS*. 2024;36(3):e20230180. <https://doi.org/10.1590/2317-1782/20242023180en>. PMID:38747747.
26. Vandembroucke JP, Von Elm E, Altman DG, Gøtzsche PC, Mulrow CD, Pocock SJ, et al. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and Elaboration. *PLoS Med*. 2007;4(10):e297. <https://doi.org/10.1371/journal.pmed.0040297>. PMID:17941715.
27. Marian V, Blumenfeld HK, Kaushanskaya M. The Language Experience and Proficiency Questionnaire (LEAP-Q): assessing language profiles in bilinguals and multilinguals. *J Speech Lang Hear Res*. 2007;50(4):940-67. [https://doi.org/10.1044/1092-4388\(2007\)067](https://doi.org/10.1044/1092-4388(2007)067). PMID:17675598.
28. Radhakrishnan M, Nagaraja SB. Modified Kuppusswamy socioeconomic scale 2023: stratification and updates. *Int J Community Med Public Health*. 2023;10(11):4415-8. <https://doi.org/10.18203/2394-6040.ijcmph20233487>.
29. Ustün TB, Chatterji S, Kostanjsek N, Rehm J, Kennedy C, Epping-Jordan J, et al. Developing the World Health Organization Disability Assessment Schedule 2.0. *Bull World Health Organ*. 2010;88(11):815-23. <https://doi.org/10.2471/BLT.09.067231>. PMID:21076562.
30. Ouellet É, Boller B, Corriveau-Lecavalier N, Cloutier S, Belleville S. The Virtual Shop: A new immersive virtual reality environment and scenario for the assessment of everyday memory. *J Neurosci Methods*. 2018;303:126-35. <https://doi.org/10.1016/j.jneumeth.2018.03.010>. PMID:29581009.
31. Ennis GE, Hess TM, Smith BT. The impact of age and motivation on cognitive effort: implications for cognitive engagement in older adulthood. *Psychol Aging*. 2013;28(2):495-504. <https://doi.org/10.1037/a0031255>. PMID:23421325.
32. Souza AC, Alexandre NMC, Guirardello EB. Psychometric properties in instruments evaluation of reliability and validity. *Epidemiol Serv Saude*. 2017;26(3):649-59. PMID:28977189.
33. Costa A, Caltagirone C, Carlesimo GA. Prospective memory impairment in mild cognitive impairment: an analytical review. *Neuropsychol Rev*. 2011;21(4):390-404. <https://doi.org/10.1007/s11065-011-9172-z>. PMID:21773811.
34. Kirova AM, Bays RB, Lagalwar S. Working memory and executive function decline across normal aging, mild cognitive impairment, and Alzheimer's Disease. *BioMed Res Int*. 2015;2015(1):748212. <https://doi.org/10.1155/2015/748212>. PMID:26550575.
35. Salthouse T. Consequences of age-related cognitive declines. *Annu Rev Psychol*. 2012;63(1):201-26. <https://doi.org/10.1146/annurev-psych-120710-100328>. PMID:21740223.
36. Henderson SE, Lockhart HA, Davis EE, Emrich SM, Campbell KL. Reduced attentional control in older adults leads to deficits in flexible prioritization of visual working memory. *Brain Sci*. 2020;10(8):542. <https://doi.org/10.3390/brainsci10080542>. PMID:32796655.
37. Saikia B, Tripathi R. Executive functions, processing speed, and memory performance: untangling the age-related effects. *J Psychiatry Spectr*. 2023;3(1):12-9. https://doi.org/10.4103/jopsys.jopsys_13_23.
38. Crook-Rumsey M, Howard CJ, Hadjiefthymou F, Sumich A. Neurophysiological markers of prospective memory and working memory in typical ageing and mild cognitive impairment. *Clin Neurophysiol*. 2022;133:111-25. <https://doi.org/10.1016/j.clinph.2021.09.019>. PMID:34839236.

Author contributions

SV: conceptualization, data curation, methodology, investigation, formal analysis, visualization, writing- original draft;

SA: formal analysis, methodology, visualization, writing- review and editing;

GB: conceptualization, project administration, validation, supervision, writing- review and editing; DFD: conceptualization, methodology, project administration, validation, writing- review and editing.