

Tilburg University

The sounds of memory

Haas, Maximilian; Mehl, Matthias R; Ballhausen, Nicola; Zuber, Sascha; Kliegel, Matthias; Hering, Alexandra

Published in:

The Journals Of Gerontology. Series B, Psychological Sciences And Social Sciences

DOI:

[10.1093/geronb/gbac012](https://doi.org/10.1093/geronb/gbac012)

Publication date:

2022

Document Version

Publisher's PDF, also known as Version of record

[Link to publication in Tilburg University Research Portal](#)

Citation for published version (APA):

Haas, M., Mehl, M. R., Ballhausen, N., Zuber, S., Kliegel, M., & Hering, A. (2022). The sounds of memory: Extending the age-prospective memory paradox to everyday behavior and conversations. *The Journals Of Gerontology. Series B, Psychological Sciences And Social Sciences*, 77(4), 695-703.
<https://doi.org/10.1093/geronb/gbac012>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Research Article

The Sounds of Memory: Extending the Age–Prospective Memory Paradox to Everyday Behavior and Conversations

Maximilian Haas, MSc,^{1,2,*} Matthias R. Mehl, PhD,³ Nicola Ballhausen, PhD,^{4,5} Sascha Zuber, PhD,^{4,6} Matthias Kliegel, PhD,^{1,4} and Alexandra Hering, PhD^{1,5}

¹Faculty of Psychology and Educational Sciences, University of Geneva, Geneva, Switzerland. ²Faculty of Psychology, UniDistance Suisse, Brig, Switzerland. ³Department of Psychology, University of Arizona, Tucson, Arizona, USA. ⁴Centre for the Interdisciplinary Study of Gerontology and Vulnerability, University of Geneva, Geneva, Switzerland. ⁵Department of Developmental Psychology, Tilburg University, Tilburg, The Netherlands. ⁶Swiss National Center of Competences in Research LIVES—Overcoming Vulnerability: Life Course Perspectives, Lausanne & Geneva, Switzerland.

*Address correspondence to: Maximilian Haas, MSc, Faculty of Psychology and Educational Sciences, University of Geneva, Boulevard du Pont-d'Arve 28, 1205 Geneva, Switzerland. E-mail: Maximilian.Haas@unige.ch

Received: August 27, 2021; Editorial Decision Date: January 13, 2022

Decision Editor: Angela Gutches, PhD

Abstract

Objectives: Around the turn of the millennium, the “age–prospective memory (PM) paradox” challenged the classical assumption that older adults necessarily evidence a marked decline in PM functioning. As previous investigations highlighted ecological validity to be a potential explanation, the present study sought to extend established approaches by using novel real-world assessment technologies to examine PM unobtrusively in everyday-life conversations.

Method: Next to laboratory PM tasks, real-life PM performance of 53 younger adults (19–32 years) and 38 older adults (60–81 years) was assessed from three sources: Over 9 days, participants completed an experimenter-given naturalistic task, a diary-based approach assessing self-assigned intentions, as well as an ambulatory assessment with the Electronically Activated Recorder (EAR), a device that unobtrusively samples ambient sounds to detect spontaneous speech production related to (lapses in) everyday PM.

Results: Older adults showed lower performance in laboratory PM only for the time-based task and performed either equally well as or even better than younger adults in everyday PM. With regard to PM performance as captured in real-life ambient audio data, younger adults talked more frequently about PM than older adults, but no significant difference between younger and older adults was found for speech related to PM errors.

Discussion: Findings confirmed older adults' preserved PM performance in everyday life across different indicators with increasing ecological validity. Furthermore, as a novel method to assess conversational PM in everyday life, the EAR opens new insights about the awareness of PM lapses and the communication of intentions in real life.

Keywords: Ambulatory assessment, Ecological validity, Electronically activated recorder, Everyday cognition

Prospective memory (PM) refers to an individual's ability to remember carrying out planned intentions at an accurate moment in the future (Kliegel et al., 2016). Having to remember to take one's medication at a certain time of day or to post a letter when passing by the post office are examples that

illustrate PM in everyday life. With more than half of all daily memory and cognition errors being related to PM in both younger and older adults (Haas et al., 2020), PM plays a key role in daily living and functional independence, particularly in older adults (Haines et al., 2019; Hering et al., 2018).

The Age–Prospective Memory Paradox

Contrary to the established age-related decline in episodic memory (Park & Reuter-Lorenz, 2009), investigations of age differences in PM performance have revealed an intriguing pattern, which has been referred to as the age–PM paradox (Rendell & Craik, 2000; Rendell & Thomson, 1999; for a detailed outline, see Schnitzspahn et al., 2020). Specifically, results from classical laboratory PM paradigms, where the intention is embedded in a computerized attention-demanding ongoing activity (cf. Einstein & McDaniel, 2005), align with the marked age-related decline across the adult life span typically found for different cognitive domains (Kliegel et al., 2008; Tucker-Drob et al., 2019). However, the advantage of a highly controlled task setting comes at the potential cost of its ecological validity. Indeed, with regard to the dimensions defining ecological validity, a laboratory environment, lower familiarity with the PM stimuli, as well as their nonnatural occurrence are characteristics of a task setting that is rather artificial and less representative of everyday PM. Thus, research has aimed at increasing the ecological validity of PM measures by embedding PM tasks into participants' naturally occurring daily life (e.g., making a phone call, or sending postcards or text messages) or into the script of a laboratory session that is characterized by task properties that participants are familiar with from their daily routine (e.g., giving back a pencil to the experimenter; see Phillips et al., 2008, for a review on ecological validity of PM tasks). In fact, meta-analytic evidence has revealed an overall age benefit in these naturalistic tasks (Henry et al., 2004), thus challenging classical assumptions about cognitive aging.

To date, few studies have investigated the age–PM paradox in and outside of the laboratory for both time-based (i.e., to be performed after a certain time has elapsed) and event-based intentions (i.e., to be performed in response to a certain type of external stimuli) within the same sample of younger and older adults (Haines et al., 2020; Niedźwieńska & Barzykowski, 2012; Rendell & Craik, 2000; Schnitzspahn et al., 2020). Across these studies, results revealed the anticipated age deficit in the laboratory and an age benefit for the naturalistic tasks in some studies, while no difference between age groups across task environments was observed in others (see Haines et al., 2020, for a summary of results). Based on these discordant patterns, it has been suggested that one should assume no age-related differences rather than proclaiming a general age benefit in everyday PM (Schnitzspahn et al., 2020). However, one possible explanation for these inconsistencies may root in the nature of some of the laboratory tasks that were used (e.g., red pencil task, Dobbs & Rule, 1987; token task, Zeintl et al., 2007). Although these tasks were carried out in a laboratory setting, their interactional, game-like character contributed to a more naturalistic-seeming experience of the task, with a presumably higher familiarity and thus higher ecological validity than the one of commonly used but more artificial (i.e., computerized) laboratory PM

paradigms (cf. Phillips et al., 2008). Thus, we aimed to examine the age–PM paradox by comparing event-based and time-based PM tasks in the laboratory and different naturalistic settings in younger and older adults. Moreover, by systematically investigating (real-life) PM tasks with different degrees of ecological validity, the current study extends available approaches to assess everyday PM.

Real-Life Assessment of Intentions

So far, studies have mainly focused on experimenter-administered tasks or self-reports in naturalistic settings. The experimenter-given instructions allow assessment of PM in participants' daily lives; however, these tasks still do not capture naturally occurring PM performance limiting their ecological validity. In fact, instructing a task to participants may not only have an influence on the importance put into task completion, but also on the motivation with which that task is carried out for someone else, namely, the experimenter (Peter & Kliegel, 2018). Thus, increasing attention has been given to the assessment of real-life PM performance by focusing on participants' own intentions in diary studies. In this approach, participants either report memory lapses that occurred during the day in an evening assessment or are instructed to note their upcoming intentions and retrospectively report whether these intentions have been successfully performed. Studies using this paradigm have revealed either no age differences (Mogle et al., 2019) or an age advantage for everyday PM performance (Ihle et al., 2012; Schnitzspahn et al., 2016). However, these studies did not directly compare diary-reported everyday PM with performance in the laboratory.

Although experimenter-given tasks and intention diaries target PM performance that is everyday-related, their generalization to *actual* real-life PM performance is still limited by the potentially nonnaturalistic content of experimenter-given tasks or the retrospective bias of self-reports (i.e., many trivial daily PM tasks may simply be forgotten over the course of the day relative to the more salient nature of some isolated PM events). With the present study, we further explored the age–PM paradox by adding a novel level of unobtrusive assessment of everyday PM functioning. Importantly, by assessing PM performance in a passive observational way using the Electronically Activated Recorder (EAR; Mehl, 2017), we aimed to circumvent the above-mentioned pitfalls of experimenter-assigned tasks and intention diaries, reaching a higher level of ecological validity.

PM in Real-Life Conversations

The EAR is a smartphone-based application that records brief samples of ambient sound intermittently throughout the day. That way, it constitutes a passive ecological assessment method that allows for an unobtrusive measurement of acoustically encoded aspects of daily life. The EAR method has been employed with good acceptance and

adherence in both younger and older participants (Manson & Robbins, 2017; Ng et al., 2021). Studies using this methodology have covered a wide range of domains, such as personality, emotional expression, or behavioral sociability (Farrell et al., 2018; Holtzman et al., 2010; Ramírez-Esparza et al., 2009). In relation to cognition, recent studies have applied the EAR to investigate executive functions, autobiographical memory, and conversational time travel, that is, utterances related to mental time travel in everyday conversations (Demiray et al., 2018, 2019; Polsinelli et al., 2020; Wank et al., 2020). Results from the latter showed that, in their everyday conversations, individuals tend to talk more often about their past than about their future, which stands in contrast to established findings that foreground a prospective bias in mental time travel (Demiray et al., 2018). Considering the novel insights that the EAR method has afforded in different areas, we see potential for its application to a broader range of psychological phenomena, and particularly phenomena related to cognitive aging as it can be observed in everyday life, such as the age-PM paradox.

Study Goals and Hypotheses

The present study was the first to apply the EAR to PM research to assess naturally occurring PM passively, without experimenter interference. Analyzing everyday speech with regard to PM unfolds the potential to capture (less salient) daily intentions that might remain uncovered with other approaches and thus offers the opportunity to draw an even more detailed image of real-life PM. In the current study, we compared the performance of laboratory-based PM tasks to three different sources of everyday PM varying in their degree of ecological validity (i.e., experimenter-given task, self-assigned intentions, and EAR-observed). According to the directions suggested by the age-PM paradox, we expected to observe an age deficit on laboratory PM tasks and an age benefit (or at least no age deficit) for the naturalistic and self-assigned intentions. Moreover, and in line with the increase in the degree of ecological validity, we hypothesized that fewer PM errors will be observed for older adults relative to their younger peers on real-life PM assessed with the EAR.

Method

Participants

The sample consisted of 53 younger adults ($M_{\text{age}} = 23.29$, $SD = 2.27$, age range 19–32, 41 women) and 38 older adults ($M_{\text{age}} = 68.20$, $SD = 5.77$, age range 60–81, 25 women). The two age groups did not differ with respect to gender distribution ($\chi^2(1) = 1.48$, $p = .22$), but showed significant differences in measures of verbal abilities (Mill Hill; Deltour, 1993: $F(1, 89) = 14.35$, $p < .001$) and fluid reasoning (Matrix Reasoning; Wechsler, 2008: $F(1, 89) = 42.99$, $p < .001$), with

older adults ($M = 25.50$, $SD = 4.67$) attaining higher scores on verbal abilities compared to younger adults ($M = 22.00$, $SD = 4.07$), but younger adults ($M = 21.17$, $SD = 2.76$) outperforming older adults ($M = 15.84$, $SD = 4.95$) on fluid reasoning.

All participants were volunteers and provided informed consent to participate in this study, which was approved by the ethics committee of the University of Geneva (decision protocol: PSE.20170803.27). Younger adults were mostly undergraduate psychology students who received course credits and 20 Swiss francs ($\approx$ USD 20) for their participation. All other nonstudent younger adults as well as all older adults received 30 Swiss francs ($\approx$ USD 30) for their participation.

Material

Laboratory PM

A paradigm embedding PM cues in a two-back picture task adapted from Jäger and Kliegel (2008) was used to assess PM in the laboratory. The two-back ongoing task involved judging whether the picture displayed on the screen is identical (right arrow key) or not (left arrow key) to the one shown two trials before. Participants were given a short practice of the two-back task (14 trials) out of which 10 trials had to be correct before advancing and had to perform this task once at the beginning and once at the end of the laboratory PM task (48 trials each block). In between these two ongoing task-only blocks, participants were presented with additional PM instructions including a time-based and an event-based block. In the *time-based* block, they had to press the Enter key every 2 min (i.e., 02:00, 04:00, 06:00, 08:00, and 10:00 min) and were able to monitor the time elapsed since the beginning of the block by pressing the space bar displaying a digital clock. In the *event-based* block, participants had to remember to press the Enter key every time the image displayed on the screen showed a piece of clothing (e.g., a pair of trousers, a dress, a shirt). Both PM blocks consisted of 120 trials each including 40 two-back targets and 80 nontargets, including the five PM cues in the event-based condition. The stimuli were presented for 4,000 ms, followed by a 1,000 ms interstimulus interval. The order of the time-based and event-based blocks was counterbalanced across participants. Performance on the event-based task was scored as correctly answered PM cues out of the total number of PM stimuli; performance on the time-based task was given for any correct response within a time window of 6 s or less around the target times (Schnitzspahn et al., 2011).

Real-life PM

Experimenter-given.—The time-based naturalistic PM task was adapted from Schnitzspahn et al. (2011) and required sending six text messages at specific times, one in the morning and one in the afternoon for the duration of three consecutive days. At the end of the laboratory session,

participants were given a sheet containing task instructions and six blank spaces to define those times in the presence of the experimenter. They could set the times according to their own schedule with the constraint to choose different hours and minutes for every message (i.e., if they had chosen 14:10 on one day, the options 14:xx or xx:10 were no longer possible for any of the other days). Times were not constrained to be either always regular (i.e., xx:00/15/30/45) or always irregular (e.g., xx:18/33/51 or never on the hour). Messages that were sent within a time window of 6 min or less around the target time were considered as correct PM responses.

Additionally, we included an event-based task for which participants were asked to take a photo of their pillow each time they would go to sleep and each time they woke up. They were explicitly instructed that these moments could also include short naps during daytime. To parallel the number of events to the time-based task, we only looked at events occurring at regular bedtimes (i.e., the last photo taken at night, and the first photo taken in the morning). An activity tracker bracelet (*MiBand 2*, Xiaomi) allowed confirmation of whether the photo was taken within a reasonable delay preceding/following the time at which the participant fell asleep/woke up. For both tasks, participants were instructed not to use any external aids such as alarms or visual reminders. Task performance was based on the number of photos taken or messages sent out of the six possible events.

Intention diary.—For three consecutive days, participants were asked to fill out a smartphone-based intention diary, adapted from Ihle et al. (2012). Every evening, participants were asked to note their personal intentions for the following day. A push notification reminded them to list all their personal intentions and tasks for the next day in a questionnaire in the smartphone application *movisensXS* (movisens GmbH, Karlsruhe, Germany). On the following day, participants were presented with their reported intentions and were asked to state whether the intention had been completed. The diary form was sent every evening at 08:00 p.m. automatically and arrival was signaled by a phone alarm. Hence, participants did not have to remember by themselves to fill out the diary. During the day, participants had no possibility of referring to the intentions they had noted. PM performance was based on the percentage of intentions reported as being completed.

EAR.—Using the Android version of the EAR application (Mehl, 2017), 30 seconds of ambient sound were recorded at random times, on average every 12 min, from 07:00 a.m. to 09:00 p.m., thus covering about 5% of participants' wake times. At any time, participants could request 30 min of guaranteed nonrecording by pressing the privacy button on the app's lock screen (and were instructed about how to do so at the end of the laboratory session). Additionally, all participants were given the opportunity to

listen to their recordings before "handing them over" to the investigators so that they could request for recordings to be permanently deleted from the study before the research staff listens to them. All EAR sound files containing speech were transcribed by trained research assistants. They were instructed to transcribe the participant's voice only (based on a short speech sample recorded at the end of the laboratory session). All transcripts were then coded by two trained coders for various contextual psychological variables, among which we also coded for PM and PM errors. Recordings were coded as containing PM when the conversation was referring (directly or indirectly) to PM (i.e., any kind of future-directed action or intention, such as "I will have to call my brother to ask him" or "I must not forget to give you 20 francs"). A PM error was coded for when conversations indicated a problem related to an intention (e.g., "I forgot to call my mother").

Procedure

After signing the informed consent, the laboratory session started with the background cognitive testing, including vocabulary and reasoning tests. Participants then had to work on the laboratory PM tasks. At the end of the laboratory session, and before the start of the real-life phase, participants were provided with a smartphone for the study (Motorola g5 for younger adults; Motorola g4 Plus with a larger screen for older adults) and necessary instructions for completing the tasks. All participants were given time to test the smartphone and its functionalities until they felt comfortable to use the phone and the applications. During the 9-day home phase, participants sequentially completed the experimenter-given task, the intention diary, and the EAR assessment for 3 days each, with task orders being fully counterbalanced across participants. To ensure equal task conditions and retention delays, participants received a smartphone-based reminder to consult the instruction sheet with times on the evening prior to the start of the experimenter-given tasks.

Results

Performance on Laboratory and Experimenter-Given Naturalistic PM Tasks

Firstly, we analyzed age-related differences of PM performance for laboratory-based and experimenter-given naturalistic PM tasks to examine the age-PM paradox in the same sample of younger and older adults. Results of a 2 (age group) $\times$ 2 (task setting) $\times$ 2 (cue type) repeated-measures analysis of variance (ANOVA) on laboratory and naturalistic PM task performances are shown in Figure 1. Analyses did not reveal a significant main effect of age ($F(1, 89) = 0.01, p = .92, \eta_p^2 = 0.00$), indicating that, overall, older adults ($M = .74, SD = .16$) performed comparably to younger adults ($M = .74, SD = .16$). Furthermore, a significant main effect of task setting ($F(1, 89) = 34.60$,

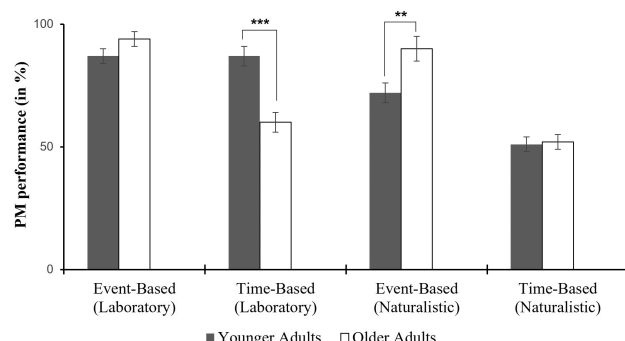


Figure 1. Mean performances on laboratory and naturalistic experimenter-given prospective memory (PM) tasks depending on task type in younger and older adults. *Note:* Performances are depicted as accuracies for the laboratory tasks (left) and naturalistic experimenter-given tasks (right). Higher values indicate better performance. Error bars represent the standard error of mean (SE). For reasons of clarity, only significant differences between age groups within the individual tasks are presented. Refer to the text for all other main effects and interactions between age, task setting, or/and cue type. *** $p < .001$, ** $p < .01$.

$p < .001$, $\eta_p^2 = 0.28$) was obtained, showing that participants performed better on the laboratory tasks ($M = .82$, $SD = .19$) than on the naturalistic tasks ($M = .66$, $SD = .22$). A significant main effect of cue type was observed ($F(1, 89) = 73.72$, $p < .001$, $\eta_p^2 = 0.45$). Overall, participants performed better on the event-based tasks ($M = .86$, $SD = .17$) than on the time-based tasks ($M = .62$, $SD = .24$).

Moreover, a significant interaction between age group and task setting revealed that younger adults ($M = .87$, $SD = .19$) performed better on laboratory tasks than older adults ($M = .77$, $SD = .19$), but that older adults ($M = .71$, $SD = .22$) outperformed younger adults ($M = .62$, $SD = .22$) on the naturalistic PM tasks ($F(1, 89) = 12.86$, $p = .001$, $\eta_p^2 = 0.12$).

A significant interaction between age group and cue type was obtained ($F(1, 89) = 19.75$, $p < .001$, $\eta_p^2 = 0.18$). Younger adults ($M = .69$, $SD = .23$) showed a higher performance on time-based PM (see Author Note 1) than their older peers ($M = .56$, $SD = .23$); however, older adults ($M = .92$, $SD = .17$) performed better on the event-based PM tasks than younger adults ($M = .80$, $SD = .17$).

Furthermore, a significant interaction between task setting and cue type was observed ($F(1, 89) = 5.24$, $p = .02$, $\eta_p^2 = 0.06$). Overall, participants showed a higher mean difference between cue types in the naturalistic setting ($\Delta M = -.30$, $SE = .04$; Cohen's $d = 0.49$) than in the laboratory ($\Delta M = -.18$, $SE = .04$; Cohen's $d = 0.29$). Additionally, the difference in mean performance between task settings was higher for time-based tasks ($\Delta M = .22$, $SE = .04$; Cohen's $d = .37$) than for event-based tasks ($\Delta M = .10$, $SE = .04$; Cohen's $d = 0.17$). The triple interaction between age group, task setting, and cue type only showed a trend toward significance ($F(1, 89) = 3.19$, $p = .07$, $\eta_p^2 = 0.04$). Bonferroni correction was applied to all analyses for adjustment of multiple comparisons by adapting the alpha level to the number of comparisons.

Performance on Self-Assigned Intentions

Next, we analyzed age differences for self-assigned PM tasks from the intention diaries. For each participant, the sum of all self-assigned intentions was calculated for the three subsequent days of the diary task. Overall, younger ($M = 6.40$, $SD = 2.76$) and older adults ($M = 7.52$, $SD = 2.99$) did not differ significantly in the number of real-life planned intentions they reported ($F(1, 84) = 3.21$, $p = .08$). With regard to content, participants mentioned a broad range of intentions such as meal arrangements with friends or relatives, plans for grocery shopping, or intentions related to work (papers to hand in, classes to attend) or household chores (laundry, ironing). Like Schnitzspahn et al. (2020), we did not include cue type as a factor in the analyses as reported intentions were almost exclusively referring to time-based activities.

In terms of PM performance, results of a one-way ANOVA showed a significant main effect of age on the completion of real-life intentions, $F(1, 587) = 8.20$, $p = .004$, $\eta_p^2 = 0.02$, revealing that older adults ($M = .86$, $SD = .18$) completed significantly more self-assigned intentions than younger adults ($M = .75$, $SD = .23$).

PM in EAR-Sampled Real-Life Ambient Audio Data

Finally, we analyzed the occurrence of PM tasks and failures in the EAR ambient audio recordings in younger and older adults. The EAR-sampled real-life ambient audio data were analyzed from 43 younger adults and 38 older adults only, as some younger participants had to be excluded due to technical problems with the application ($n = 8$) and non-compliance with the study protocol ($n = 2$). Table 1 presents the descriptive statistics of the EAR data. Participants from both age groups made use of the privacy button similarly often ($t(79) = 0.691$, $p = .56$), and younger adults and older adults did not differ significantly in the overall number of sound files recorded ($t(79) = 0.60$, $p = .55$).

Recordings were excluded if the audio information was bad or insufficient to perform coding, as well as if participants were sleeping or deemed not wearing the EAR. Overall, about 25.1% of all valid recordings contained speech. Relative to their number of valid waking sound files, younger adults ($M = .24$, $SD = .12$) and older adults ($M = .27$, $SD = .10$) were similarly talkative ($t(79) = 1.04$, $p = .30$).

The interrater reliability for the coding of PM in speech files showed strong agreement between the two coders (Cohen's $\kappa = .84$, $p < .001$). Discrepancies in coding were settled by discussions among coders and the authors to get to a mutual agreement. In summary, speech files of younger adults had more PM-related content than those of older adults ($\chi^2(1, N = 3,629) = 12.73$, $p < .001$).

With regard to PM errors, 14 recordings have been identified. Interrater reliability showed strong agreement between coders (Cohen's $\kappa = .88$, $p < .001$). Younger ($n = 6$)

Table 1. Descriptive Statistics of the Real-Life Ambient Audio Data in the Overall Sample and by Age Group

	Overall sample		Younger adults ^a			Older adults ^b		
	<i>n</i>	<i>M</i> (<i>SD</i>)	<i>n</i>	Range	<i>M</i> (<i>SD</i>)	<i>n</i>	Range	<i>M</i> (<i>SD</i>)
Privacy button use	72	0.89 (2.71)	31	0–10	0.72 (1.71)	41	0–20	1.08 (3.54)
Recorded sound files	16,799	207.40 (21.49)	8,936	133–247	196.33 (15.77)	7,863	116–273	198.55 (17.63)
Excluded sound files	2,327	28.73 (30.01)	1,530	133–214	35.58 (31.26)	797	116–214	20.94 (26.84)
Insufficient quality	812	10.02 (12.26)	494	0–42	11.49 (12.04)	318	0–66	8.37 (12.45)
Noncompliance	391	4.83 (14.54)	231	0–91	5.37 (15.20)	160	0–63	4.21 (13.93)
Sleeping	1,124	13.88 (18.61)	805	0–86	18.72 (22.69)	319	0–40	8.39 (10.31)
Valid waking files	14,472	178.67 (29.23)	7,406	63–205	172.23 (32.89)	7,066	116–209	185.95 (22.72)
Speech present	3,629	44.90 (20.21)	1,707	9–82	39.70 (17.79)	1,930	7–110	50.79 (21.37)
PM	267	3.30 (3.42)	153	0–21	3.56 (3.79)	114	0–13	3.00 (2.96)
PM errors	14	0.17 (0.41)	6	0–1	0.14 (0.35)	8	0–2	0.21 (0.47)

Notes: Mean number of sound files and mean number of privacy button use are presented with their respective standard deviations. Recorded sound files = total number of recordings obtained per participant. Sound files have been excluded if the audio quality was insufficient to perform any transcription, when participants were deemed not compliant with the study protocol (i.e., not carrying the EAR), or when participants were sleeping. Valid waking files = all recording that transcription where transcription could be obtained and where participants are deemed compliant (i.e., carrying the telephone and not sleeping). PM = transcription was referring to prospective memory; PM errors = transcription was containing a reference to a prospective memory error.

^a*n* = 43.

^b*n* = 38.

and older adults (*n* = 8) did not differ significantly in the frequency of real-life PM errors ($\chi^2(1, N = 267) = 1.26$, $p = .26$). Examples of transcribed EAR-sampled real-life audio data for PM (Supplementary Table 1) and PM errors (Supplementary Table 2) are given in the Supplementary Material.

Discussion

The present study is the first to investigate naturally occurring PM in sampled real-life conversations using the EAR method. Furthermore, it contributes to research on the age-PM paradox by comparing laboratory tasks and three sources of everyday PM differing in their degree of ecological validity in the same sample of younger and older adults. We hypothesized that younger adults would outperform older adults on the laboratory PM tasks, but that older adults would excel in everyday PM compared to younger adults.

In terms of the naturalistic side of the paradox, our findings are mostly in line with previous studies suggesting no age differences or even an age benefit (Schnitzspahn et al., 2020). First, in the experimenter-given task, we obtained an age benefit for the event-based, but no age difference for the time-based task. Second, in line with, for instance, Haas et al. (2020) and Schnitzspahn et al. (2020), an age benefit was revealed for the intention diary as older adults completed significantly more self-assigned intentions than younger adults. Third, the EAR analyses demonstrated that younger and older adults did not differ in the number of PM errors mentioned in their everyday conversations.

In terms of the laboratory side of the paradox, our results are only partially in line with previous findings as they showed the expected age deficit only for the time-based

task, but not for the event-based task. However, this may be explained by the ongoing task performances. In fact, exploratory analyses showed that older adults were slower and less accurate in performing the ongoing task of the PM blocks compared to the ongoing task-only blocks (Supplementary Table 3). This may indicate that older adults have prioritized the PM intention over the ongoing task performance, leading to a trade-off between the ongoing task and the PM task. Furthermore, although older adults had a significantly lower performance than younger adults, both age groups were descriptively still at a very high level of performance, leaving thus sufficient cognitive resources to perform well on both the ongoing task and the PM intention. A possible explanation may be that the relatively long presentation intervals allowed participants to increase their target checking. Future studies could investigate the impact of adapted ongoing task performance on the paradox (i.e., whether the paradox is replicated when task performance is matched among age groups; see Voigt et al., 2014, for a similar approach used in schoolchildren).

Conceptually, the present study reveals convergent evidence across multiple data sources that in situations where individuals have to deploy natural PM performance (in contrast to laboratory maximum performance settings), older adults do not show any sign of an age-related deficit in PM functioning—this holds for experimenter-given everyday tasks, self-reports of daily intentional behavior, and the analyses of PM-related daily talk. Interestingly, however, the increase in ecological validity of the everyday PM measures did not seem to be directly linked to an increase in performance as well, as the pattern of results does not support the idea that older adults will have the most advantage for measures at the highest level of ecological validity (i.e., the assessment with the EAR). Especially, with regard

to the latter approach, the current study may contribute to launching a novel methodology that might help understanding how PM functions in everyday life and where age-related (or possibly any other—clinical—population-related) differences may occur. Present analyses revealed that younger adults talked more about PM in their everyday lives than older adults, but participants from both age groups talked equally often about PM errors. As younger adults talked more about PM in their everyday lives, one may think that they have a high awareness of tasks that need to be accomplished. However, this did not reflect in their self-reported performance. Indeed, findings from the EAR seem diametrically opposed to the results from participants' diary performance, where no difference in the number of reported intentions was observed, but younger adults showed a lower completion rate than older adults. Differences between these findings can (to some extent) be explained by the different focus of these two methods, in the sense that 3 days of EAR observation only covering 5% of a participant's day do not fully map the information grasped by a retrospective diary approach. In fact, some intentions may possibly not be noted by participants, either because they would not consider it to be related to PM, or because it may seem too insignificant to be reported. Moreover, other intentions may only come up during the day as a consequence of some events, and performance on these tasks would therefore not appear in the diary, but they might be captured by the EAR instead.

Yet, these differences in results raise a more fundamental question, notably about the functional purpose of communicating certain intentions (or errors) and not communicating others. Previous research has already pointed to PM being tightly linked to interpersonal relations (Meacham, 1988). Thus, communicating PM and PM errors to others could fulfill different functional goals, such as rehearsal, offloading, or shared cognition (for general communication on PM), or emotion regulation and coping due to irritation, interference, and consequences of PM errors (Mogle et al., 2019). Analogous to the terminology of conversational time travel used for functional communication of mental time travel (Demiray et al., 2018), we propose the concept of conversational PM to describe PM in everyday conversations. Future studies may further investigate the functional purposes of conversational PM and potential age differences therein.

Limitations and Methodological Considerations

Although we used a standard laboratory procedure, one limitation concerns the laboratory task, as performance levels were high for both task types (Einstein & McDaniel, 2005; Jäger & Kliegel, 2008). Yet, our analyses suggest that older adults may have prioritized PM over ongoing task performance, therefore masking the expected PM deficit. Another limitation is the difference in cognitive load between tasks in the laboratory and the naturalistic setting, particularly for time-based PM (i.e., short vs. appointment-style intervals).

Future studies need to bear in mind that these performances are not obviously comparable and that the interpretation of results may depend on the analytic approach used (see Author Note 1). Furthermore, results concerning the real-life audio data should be interpreted with caution, as a direct comparison to performances from the other tasks is not possible as we do not have any information about the total number of daily intentions. As discussed above, some intentions may just not be communicated at all for some reason and would thus not be detected by the EAR either. Employing self-reports on everyday intentions complementary to the EAR would allow capture of important information in order to approach a more complete measure of everyday PM.

A final possible limitation concerns the sampling of study days and whether they represent an average day in terms of real-life PM load. The use of self-reports may provide complementary information to compare the daily PM load across participants. Additionally, differences in lifestyles between younger and older adults may play a role. For example, students spending time at university or studying usually have a lower number of interaction partners during the day, but may be very social and have a tendency to meet more people in their free time. Older adults, if already retired, can spend time more freely and possibly have a higher number of contacts, but they also spend a lot of time at home (alone) with socially passive activities (e.g., TV, reading, household). However, all those factors are likely also involved in the entire literature on the age-PM paradox. Future studies could directly address these factors, as well as the impact of partnerships or whether living on one's own or not may influence the results by combining different methodologies (i.e., ambulatory assessment and experience sampling) to expand the range of possibilities for studying daily life in more detail.

Conclusion

The present study provided robust evidence that older adults may perform equally well as, if not even better than younger adults, consolidating these patterns across laboratory tasks and three independent everyday PM indicators (i.e., experimenter-given, self-assigned, and real-life conversations). More importantly, the finding supports the urge for more ecologically valid measures, as maximum performance (laboratory-based) approaches may underestimate older adults' actual cognitive capacities. Beyond that, results from the EAR open a novel understanding of how people communicate about their intentions and how they are aware of their own PM failures in everyday life.

Supplementary Material

Supplementary data are available at *The Journals of Gerontology, Series B: Psychological Sciences and Social Sciences* online.

Author Note

1. When considering the absolute deviation from target times in the time-based PM tasks instead, younger adults ($M = 35.60$, $SD = 51.40$) and older adults ($M = 22.97$, $SD = 26.67$) showed no difference (i.e., deviation in minutes) on the naturalistic PM task ($F(1, 87) = 1.83$, $p = .18$), but descriptively younger adults ($M = 3.84$, $SD = 9.04$) tended toward a more precise response (i.e., deviations in seconds) compared to their older peers ($M = 7.48$, $SD = 7.69$) on the laboratory PM task ($F(1, 85) = 3.75$, $p = .05$). No interaction between age group and task setting was obtained that way ($F(1, 82) = 2.38$, $p = .12$).

Funding

This work was supported by the Swiss National Science Foundation (grant number 100019_165572).

Conflict of Interest

None declared.

Acknowledgments

We thank Nora Geppert, António Manuel Fernandes, and Chloé Ruche for their assistance in study preparation and data collection. Furthermore, we want to thank Mélissa Nobile, Clarisse Cao, Diana Paineira Martinez, and Pauline Chatton for their great help with transcription and coding of the audio data files. We thank Burcu Demiray for her precious help and advice with regard to the processing of the audio data. This study has not been preregistered. A minimal data set including the data underlying the findings described here will be made available to other researchers if requested.

References

- Deltour, J. J. (1993). Echelle de vocabulaire de Mill Hill de JC Raven. *Adaptation Française et Normes Européennes Du Mill Hill et Du Standard Progressive Matrices de Raven (PM38)*. Braine-le-Château: Editions l'application des techniques modernes.
- Demiray, B., Mehl, M. R., & Martin, M. (2018). Conversational time travel: Evidence of a retrospective bias in real life conversations. *Frontiers in Psychology*, 9, 1–18. doi:10.3389/fpsyg.2018.02160
- Demiray, B., Mischler, M., & Martin, M. (2019). Reminiscence in everyday conversations: A naturalistic observation study of older adults. *The Journals of Gerontology, Series B: Psychological Sciences and Social Sciences*, 74(5), 745–755. doi:10.1093/geronb/gbx141
- Dobbs, A. R., & Rule, B. G. (1987). Prospective memory and self-reports of memory abilities in older adults. *Canadian Journal of Psychology/Revue Canadienne de Psychologie*, 41(2), 209–222. doi:10.1037/h0084152
- Einstein, G. O., & McDaniel, M. A. (2005). Prospective memory. *Current Directions in Psychological Science*, 14(6), 286–290. doi:10.1111/j.0963-7214.2005.00382.x
- Farrell, A. K., Slatcher, R. B., Tobin, E. T., Imami, L., Wildman, D. E., Luca, F., & Zilioli, S. (2018). Socioeconomic status, family negative emotional climate, and anti-inflammatory gene expression among youth with asthma. *Psychoneuroendocrinology*, 91, 62–67. doi:10.1016/j.psyneuen.2018.02.011
- Haas, M., Zuber, S., Kliegel, M., & Ballhausen, N. (2020). Prospective memory errors in everyday life: Does instruction matter? *Memory*, 28(2), 196–203. doi:10.1080/09658211.2019.1707227
- Haines, S. J., Randall, S. E., Terrett, G., Busija, L., Tatangelo, G., McLennan, S. N., Rose, N. S., Kliegel, M., Henry, J. D., & Rendell, P. G. (2020). Differences in time-based task characteristics help to explain the age-prospective memory paradox. *Cognition*, 202, 104305. doi:10.1016/j.cognition.2020.104305
- Haines, S. J., Shelton, J. T., Henry, J. D., Terrett, G., Vorwerk, T., & Rendell, P. G. (2019). Prospective memory and cognitive aging. In B. Knight (Ed.), *Oxford research encyclopedia of psychology*. Oxford University Press. doi:10.1093/acrefore/9780190236557.013.381
- Henry, J. D., MacLeod, M. S., Phillips, L. H., & Crawford, J. R. (2004). A meta-analytic review of prospective memory and aging. *Psychology and Aging*, 19(1), 27–39. doi:10.1037/0882-7974.19.1.27
- Hering, A., Kliegel, M., Rendell, P. G., Craik, F. I. M., & Rose, N. S. (2018). Prospective memory is a key predictor of functional independence in older adults. *Journal of the International Neuropsychological Society*, 24(6), 640–645. doi:10.1017/S1355617718000152
- Holtzman, N. S., Vazire, S., & Mehl, M. R. (2010). Sounds like a narcissist: Behavioral manifestations of narcissism in everyday life. *Journal of Research in Personality*, 44(4), 478–484. doi:10.1016/j.jrp.2010.06.001
- Ihle, A., Schnitzspahn, K., Rendell, P. G., Luong, C., & Kliegel, M. (2012). Age benefits in everyday prospective memory: The influence of personal task importance, use of reminders and everyday stress. *Aging, Neuropsychology, and Cognition*, 19(1–2), 84–101. doi:10.1080/13825585.2011.629288
- Jäger, T., & Kliegel, M. (2008). Time-based and event-based prospective memory across adulthood: Underlying mechanisms and differential costs on the ongoing task. *The Journal of General Psychology*, 135(1), 4–22. doi:10.3200/GENP.135.1.4-22
- Kliegel, M., Ballhausen, N., Hering, A., Ihle, A., Schnitzspahn, K. M., & Zuber, S. (2016). Prospective memory in older adults: Where we are now and what is next. *Gerontology*, 62(4), 459–466. doi:10.1159/000443698
- Kliegel, M., Mackinlay, R., & Jäger, T. (2008). Complex prospective memory: Development across the lifespan and the role of task interruption. *Developmental Psychology*, 44(2), 612–7. doi:10.1037/0012-1649.44.2.612
- Manson, J. H., & Robbins, M. L. (2017). New evaluation of the Electronically Activated Recorder (EAR): Obtrusiveness, compliance, and participant self-selection effects. *Frontiers in Psychology*, 8, 658–658. doi:10.3389/fpsyg.2017.00658
- Meacham, J. A. (1988). Interpersonal relations and prospective memory. In M. M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.), *Practical aspects of memory: Memory of everyday life* (pp. 354–359). Wiley.
- Mehl, M. R. (2017). The Electronically Activated Recorder (EAR). *Current Directions in Psychological Science*, 26(2), 184–190. doi:10.1177/0963721416680611

- Mogle, J., Muñoz, E., Hill, N. L., Smyth, J. M., & Sliwinski, M. J. (2019). Daily memory lapses in adults: Characterization and influence on affect. *The Journals of Gerontology, Series B: Psychological Sciences and Social Sciences*, 74(1), 59–68. doi:10.1093/geronb/gbx012
- Ng, Y. T., Huo, M., Han, S. H., Birditt, K., & Fingerma, K. (2021). Older adult's marital status, conversation frequency, and well-being in everyday life. *The Journals of Gerontology, Series B: Psychological Sciences and Social Sciences*, 77(3), 499–512. doi:10.1093/geronb/gbab112
- Niedźwieńska, A., & Barzykowski, K. (2012). The age prospective memory paradox within the same sample in time-based and event-based tasks. *Aging, Neuropsychology, and Cognition*, 19(1–2), 58–83. doi:10.1080/13825585.2011.628374
- Park, D. C., & Reuter-Lorenz, P. (2009). The adaptive brain: Aging and neurocognitive scaffolding. *Annual Review of Psychology*, 60(1), 173–196. doi:10.1146/annurev.psych.59.103006.093656
- Peter, J., & Kliegel, M. (2018). The age-prospective memory paradox: Is it about motivation? *Clinical and Translational Neuroscience*, 2(2), 1–18. doi:10.1177/2514183x18807103
- Phillips, L. H., Henry, J. D., & Martin, M. (2008). Adult aging and prospective memory: The importance of ecological validity. In M. Kliegel, M. A. McDaniel, & G. O. Einstein (Eds.), *Prospective memory: Cognitive, neuroscience, developmental, and applied perspectives* (pp. 161–185). Taylor & Francis Group/Lawrence Erlbaum Associates.
- Polsinelli, A. J., Moseley, S. A., Grilli, M. D., Glisky, E. L., & Mehl, M. R. (2020). Natural, everyday language use provides a window into the integrity of older adults' executive functioning. *The Journals of Gerontology, Series B: Psychological Sciences and Social Sciences*, 75(9), 215–220. doi:10.1093/geronb/gbaa055
- Ramírez-Esparza, N., Mehl, M. R., Álvarez-Bermúdez, J., & Pennebaker, J. W. (2009). Are Mexicans more or less sociable than Americans? Insights from a naturalistic observation study. *Journal of Research in Personality*, 43(1), 1–7. doi:10.1016/j.jrp.2008.09.002
- Rendell, P. G., & Craik, F. I. M. (2000). Virtual week and actual week: Age-related differences in prospective memory. *Applied Cognitive Psychology*, 14(7), 43–62. doi:10.1002/acp.770
- Rendell, P. G., & Thomson, D. M. (1999). Aging and prospective memory: Differences between naturalistic and laboratory tasks. *The Journals of Gerontology, Series B: Psychological Sciences and Social Sciences*, 54(4), 256–269. doi:10.1093/geronb/54b.4.p256
- Schnitzspahn, K. M., Ihle, A., Henry, J. D., Rendell, P. G., & Kliegel, M. (2011). The age-prospective memory-paradox: An exploration of possible mechanisms. *International Psychogeriatrics*, 23(4), 583–592. doi:10.1017/S1041610210001651
- Schnitzspahn, K. M., Kvavilashvili, L., & Altgassen, M. (2020). Redefining the pattern of age-prospective memory-paradox: New insights on age effects in lab-based, naturalistic, and self-assigned tasks. *Psychological Research*, 84(5), 1370–1386. doi:10.1007/s00426-018-1140-2
- Schnitzspahn, K. M., Scholz, U., Ballhausen, N., Hering, A., Ihle, A., Lagner, P., & Kliegel, M. (2016). Age differences in prospective memory for everyday life intentions: A diary approach. *Memory*, 24(4), 444–454. doi:10.1080/09658211.2015.1018276
- Tucker-Drob, E. M., Brandmaier, A. M., & Lindenberger, U. (2019). Coupled cognitive changes in adulthood: A meta-analysis. *Psychological Bulletin*, 145(3), 273–301. doi:10.1037/bul0000179
- Voigt, B., Mahy, C. E., Ellis, J., Schnitzspahn, K., Krause, I., Altgassen, M., & Kliegel, M. (2014). The development of time-based prospective memory in childhood: The role of working memory updating. *Developmental Psychology*, 50(10), 2393–2404. doi:10.1037/a0037491
- Wank, A. A., Mehl, M. R., Andrews-Hanna, J. R., Polsinelli, A. J., Moseley, S., Glisky, E. L., & Grilli, M. D. (2020). Eavesdropping on autobiographical memory: A naturalistic observation study of older adults' memory sharing in daily conversations. *Frontiers in Human Neuroscience*, 14, 238. doi:10.3389/fnhum.2020.00238
- Wechsler, D. (2008). *Wechsler adult intelligence scale—Fourth edition (WAIS-IV)*. Pearson.
- Zeintl, M., Kliegel, M., & Hofer, S. M. (2007). The role of processing resources in age-related prospective and retrospective memory within old age. *Psychology and Aging*, 22(4), 826–834. doi:10.1037/0882-7974.22.4.826