

Prospective and Retrospective Metacognitive Judgments of Prospective Memory Performance Across the Lifespan

Steven Stelz  and Mareike Altgassen 



Johannes Gutenberg University Mainz, Germany


ABSTRACT

Prospective Memory (PM), the ability to carry out future intentions at a specific time or event, develops in an inverted U-shaped function across the lifespan. Most studies have examined age-related changes in PM with regard to the role of executive functions. Although more recent studies have emphasized the importance of metacognition for successful prospective remembering, very little is known about the development of metacognitive abilities across the lifespan and their impact on PM performance. Therefore, the current study examined developmental differences in metacognitive monitoring and PM in five age groups in Germany ($N = 151$): children (9–10 years), adolescents (14–15 years), young adults (20–25 years), middle-aged adults (30–50 years) and old adults (60+ years). Metacognitive monitoring was assessed by asking participants to judge their performance before (prediction) and after (postdiction) working on a lab-based PM task (Virtual Week). PM performance increased from childhood into adolescence, remained stable until middle-adulthood and declined again in older adults. In terms of metacognitive monitoring, PM performance prediction was least accurate in children, and accuracy increased across the lifespan into old adulthood; interestingly, older adults' more accurate metacognitive ability did not enable them to show better PM. Overall, individuals tended to overestimate their upcoming PM performance and to underestimate their past performance. Possible strategies to improve especially children and adolescents' metacognitive skills and PM performance are discussed.

Introduction

Being able to remember to carry out a future intention (prospective memory [PM]; Einstein & McDaniel, 1990), such as remembering to call the mechanic to make an appointment or to hand in schoolwork at the appropriate time, is of crucial importance for developing and maintaining day-to-day functioning and an independent life across the lifespan (Hering et al., 2018; Mahy et al., 2014). Research categorizes PM tasks based on the type of cue that indicates the appropriate moment to initiate the intended action. Time-based PM tasks require the individual to remember to execute an intention after a delay on one's own initiative at a certain future time point (e.g., taking medication at 8:00 am) or after a certain time period elapsed (e.g., taking the cake out of the oven after 60 minutes). Event-based PM

CONTACT Steven Stelz  ststelz@uni-mainz.de  Department of Psychology, Johannes Gutenberg University of Mainz, Binger Straße 14-16, Mainz D-55122, Germany

 Supplemental data for this article can be accessed online at <https://doi.org/10.1080/15248372.2024.2441689>.

© 2024 The Author(s). Published with license by Taylor & Francis Group, LLC.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

tasks need to be initiated when an event occurs (e.g., buying milk at the grocery store when passing by after work; McDaniel & Einstein, 2000). PM comprises four successive stages (Process model, Kliegel et al., 2002): intention formation, retention, initiation, and execution. PM tasks have to be initiated and executed while being engaged in a so-called “Ongoing Task.” The development of PM across the lifespan follows an inverted U-shaped function, with PM beginning to develop around the age of two to 3 years with still very limited abilities (e.g., Guajardo & Best, 2000; Somerville et al., 1983), PM performance increasing steadily throughout childhood (Zimmermann & Meier, 2006), peaking at the end of adolescence or young adulthood (e.g., Altgassen et al., 2017; Wang et al., 2011), and declining again in older adulthood (Kliegel et al., 2008; Zimmermann & Meier, 2006). To date, only a few studies have examined age-related differences in PM performance across the lifespan using the same paradigm (Kliegel et al., 2008; Kretschmer-Trendowicz & Altgassen, 2016; Mattli et al., 2014; Zimmermann & Meier, 2006; for an overview see; Zuber & Kliegel, 2020), and they all found support for an inverted U-shaped development of PM across lifespan.

Regarding the cognitive functions underlying the observed increase and later decrease in PM performance across the lifespan, most studies to date have looked at the role of executive functions (e.g., Guajardo & Best, 2000; Kerns, 2000; Kliegel & Jäger, 2007; Kliegel et al., 2002; Mahy & Moses, 2011). Executive functions are high-level prefrontal cognitive processes that regulate and control lower-order cognitive processes in the service of goal-directed behavior (Ardila, 2008). Similar to PM development (Best & Miller, 2010; Zelazo et al., 2004), their development is characterized by an inverted U-shaped function across the lifespan. Executive functions are assumed to be involved in planning the intended action, in monitoring for the appropriate moment to execute the intention as well as in inhibiting performing the ongoing task and switching to executing the PM task (Kliegel et al., 2002). Accordingly, numerous studies across the lifespan have found executive functions to be critically involved in prospective remembering (e.g., Executive Framework of PM development, Mahy et al., 2014; Martin et al., 2003; Schnitzspahn et al., 2013). Recently, the impact of another cognitive function, namely metacognition, on successful PM performance has gained some interest (e.g., Cottini et al., 2019, 2021; Lavis & Mahy, 2021; Rummel & Meiser, 2013).

Metacognition refers to the ability to reflect about one’s own thinking and learning process and to choose the appropriate strategies that lead us to achieve our goals based on this reflection (Flavell, 1979). Metacognition is an umbrella term comprising two metamemory components: on the one hand *declarative* metamemory which refers to the explicit knowledge about memory and how memory performance is affected by person (e.g., awareness of one’s own cognitive limitations), task (e.g., difficulty level) and strategy variables (e.g., rehearsal is a good strategie for learning vocabulary) and on the other hand *procedural* metamemory which describes the strategic application of that knowledge in the service of individuals own memory performances (Tauber & Dunlosky, 2016). Procedural metamemory can be further divided into *metacognitive monitoring* (i.e., the ability to evaluate one’s own mental states, knowledge, and performance) and *metacognitive control* (i.e., the ability to adjust one’s behavior to monitoring feedback; Koriat et al., 2006; Kuhlmann, 2019).

Expectations on children’s ability to independently monitor their memory and strategy use to successfully remember and execute their future intentions increases with age as the

negative consequences of failure of remembering (e.g., forgetting to take essential medication) may increase across childhood and adolescence. For example, accurately assessing one's progress in learning new vocabulary will yield better results on a test, while overconfidence may lead to inadequate study and lower grades. Metacognition has been extensively studied in various fields of memory (e.g., retrospective memory; Dunlosky & Metcalfe, 2008 or short-term memory; Bertrand et al., 2017); however, only recently have PM researchers began to investigate the interplay between metacognition and PM (e.g., Cottini et al., 2019, 2021; Gilbert, 2015). To date, the developmental course of metacognitive abilities across the lifespan is still somewhat unknown (O'Leary & Sloutsky, 2019). It has been suggested that the development of metacognition might follow the age-related trajectory of executive functions given the great overlap of shared processes of metacognitive abilities and executive functioning (O'Leary & Sloutsky, 2019) and reflecting the development of prefrontal cortex (Morton et al., 2009). In order to successfully complete a PM task, it is of crucial importance to select the right strategy for the respective task demands (e.g., setting a reminder at the appropriate time). Highlighting the overlap between metacognition and PM, R. E. Smith (2016) stated in her meta-intentional research framework that metacognitive processes (e.g., choosing an appropriate strategy based on past experience in the intention formation phase to successfully execute the planned action in the execution phase) are inherently part of PM.

Typically, *metacognitive monitoring* (in PM) is investigated by asking participants to predict their upcoming (predictions) or to rate their past PM performance (postdictions; Lavis & Mahy, 2021). There are two primary methods for monitoring performance that are commonly used: 1) *judgements of confidence* which entail determining the extent of certainty an individual possesses regarding their task performance (e.g. Cottini et al., 2021; Gilbert, 2015), and 2) *accuracy estimations*, where participants are required to indicate the proportion of correct answers they expect to provide or have provided, respectively (e.g., S. J. Smith et al., 2011). Predictions are based on the premise that the task has either never been executed or has only been performed in a trial run which implies the (almost) absence of past experience and solely relies on assumptions about future performance. In addition, prediction accuracy may be influenced by task familiarity, task difficulty, and motivation (Schneider et al., 1998). Conversely, postdictions are made immediately following completing the task and are based on the recent experience of performing the task which might allow for a more accurate reflection of the actual performance of the task (Kuhlmann, 2019).

To date only a handful of studies have investigated the relationship between metacognitive processes and PM performance. The majority of studies conducted in children have focused on very narrow age ranges, such as 4–6 year-olds (Lavis & Mahy, 2021), 5 year-olds (Kvavilashvili & Ford, 2014), 7–8 year-olds (Cottini et al., 2018, 2019), or compared specific child age groups (e.g., 5–6 vs. 8–10 years; Cottini et al., 2021). Similarly, studies in adults have primarily examined metacognitive abilities within specific age groups, like young adults (Meeks et al., 2007; Rummel & Meiser, 2013; Schnitzspahn et al., 2011), middle-aged adults (Knight et al., 2005), and older adults (S. J. Smith et al., 2011). Gilbert's (2015) study was an exception, as it included a wide age range (18–68 years). Hence, it is unclear if age-related differences in metacognitive monitoring across the lifespan affect PM performance.

In terms of metacognitive development in children, Kvavilashvili and Ford (2014) showed that two-thirds of the examined 5-year-old children were able to accurately

predict their upcoming PM performance within a one-trial-PM-task. Lavis and Mahy (2021) asked 4–6-year-old children to predict their performance on either an easy (salient PM cue) or a difficult PM task (nonsalient PM cue) and to rate their PM performance afterward. Regardless of their age, children tended to overestimate their upcoming performance but showed more accuracy in their postdictions and were even able to consider task-difficulty in their ratings. Cottini et al. (2019) asked 7–8-year-old children to rate their PM performance in a picture classification task with five embedded event-based PM trials, both before and after performing the task. Prediction accuracy and actual PM performance were not related, but metacognitive monitoring improved when it was given retrospectively (postdictions).

In another study, Cottini et al. (2021) compared both metacognitive monitoring (predictions and postdictions) and metacognitive control in two groups of children (5–6 year-old and 8–10 year-old) also using a picture classification task with three or four embedded event-based PM cues. Both age groups rated their PM performance better than it actually was, both before and after performing the task, but the accuracy of predictions increased with age, indicating better metacognitive abilities in the older age group. Both groups were more accurate at rating their PM performance after completing the task as compared to predicting their performance, suggesting that already for children task experience improves the accuracy of their judgment. So far, no study has investigated PM and metacognition in adolescents. Even with regards to the more extensively studied field of retrospective memory, there seems to be a research gap for adolescents older than 14 years of age. Overall, the limited existing evidence points to adolescents still overestimating their predicted performance, though less strongly than younger age groups (Hoffmann-Biencourt et al., 2010; for an overview see; Godfrey et al., 2023).

Regarding metacognitive performance in adults, Meeks et al. (2007) compared two groups of undergraduate students in their metacognitive monitoring abilities (prediction and postdiction) in a lexical decision (high vs. low cognitive load) task with an embedded event-based task. Overall, students in the high and low cognitive load conditions were underconfident in predicting their upcoming performance. Postdictions were more accurate for the low cognitive load condition; while underconfidence remained high in the high cognitive demand condition even after completing the task. The authors concluded that young adults are not well calibrated in their metacognitive awareness of laboratory PM tasks, neither before nor after completing a task. Gilbert (2015) investigated the influence of metacognitive abilities on PM performance across the adult lifespan (age range 18–66, mean age 33) in a web-based intention-offloading-task. Participants were underconfident about their upcoming performance (predictions); their postdictions were more accurate, though participants were still underconfident about their performance (c.f., Knight et al., 2005; Schnitzspahn et al., 2011). S. J. Smith et al. (2011) examined performance predictions in older adults (mean age 74 years). Sixteen older adults performed event- and time-based PM tasks and were asked to predict how many PM cues they expected to find or had found, respectively. In contrast to Gilbert's results, participants were clearly overconfident about their upcoming performance, but postdictions were more accurate than predictions in relation to actual performance.

Taken together, it seems that children and adolescents become more accurate in their assessments as they age, but often overestimate their upcoming performance, whereas young and middle-aged adults seem to underestimate their performance. In old adulthood,

there appears to be a shift toward overestimating one's upcoming performance in a laboratory task. With regard to postdictions, overall, all age groups show more accurate assessments after actual task performance as compared to pre-task predictions.

The goal of the current study was to investigate, for the first time, age-related differences in prospective metacognitive monitoring and PM performance across the lifespan. To this end, five age groups, namely children, adolescents, young, middle-aged and old adults completed the same PM task (Virtual Week; Rendell & Craik, 2000) and were requested to pre- and postdict their PM performance. In addition, they were asked to rate their everyday PM performance and their use of internal and external strategies to remember to execute intentions in everyday life (Rummel et al., 2019).

Given previous evidence of PM performance showing an inverted U-shaped pattern across the lifespan (Kliegel et al., 2008; Kretschmer-Trendowicz & Altgassen, 2016; Mattli et al., 2014; Zimmermann & Meier, 2006, for an overview see; Zuber & Kliegel, 2020), we expected children and old adults to show the poorest PM performance (Mattli et al., 2014; Zimmermann & Meier, 2006), while adolescents, young adults and middle-aged adults would outperform children and old adults (Kretschmer-Trendowicz & Altgassen, 2016; Mattli et al., 2014). We further expected no age-related differences between adolescents, young adults and middle-aged adults (Kretschmer-Trendowicz & Altgassen, 2016; Mattli et al., 2014).

Regarding developmental differences in prospective metacognitive monitoring abilities across the lifespan, for prospective judgments (predictions), we hypothesized that children (based on PM research; e.g. Cottini et al., 2019, 2021), adolescents (based on retrospective memory research; Godfrey et al., 2023) and old adults (based on PM research; e.g. S. J. Smith et al., 2011) would overestimate their upcoming PM performance, whereas young- and middle-aged adults would underestimate it (based on PM research, Gilbert, 2015; Meeks et al., 2007). In contrast, for retrospective judgment (postdictions), we expected all age groups to give more accurate judgments after performing the task (Cottini et al., 2019; Gilbert, 2015; Meeks et al., 2007; S. J. Smith et al., 2011). As the link between metacognitive awareness and the selection of encoding and retrieval strategies is largely unexplored in everyday life, but of high importance for successful PM performance, possible associations between self-reported everyday PM performance and laboratory PM measures across the lifespan will be explored.

Methods

Participants

A statistical power analysis based on a medium effect size ($f = .25$) was performed with G*Power 3.1 (Faul et al., 2007) and indicated a sample size of 150 was required to achieve 80% power ($\alpha = .05$), resulting in a group size of 30 subjects per age group. Data was collected from 155 participants in Germany; however, the data sets of four participants were excluded due to technical problems (2 young adults) and difficulties in PM task comprehension (1 child, 1 old adult). Thus, the final sample comprised 151 native German-speaking individuals between the ages of 9 and 75 years belonging to five different age groups (for details see Table 1). Participants were recruited from various sources, including local schools for children and adolescents, e-mail lists of the local university for young

Table 1. Sample characteristics.

	<i>N</i>	Age-range in years	Age <i>M</i> (<i>SD</i>) in years	Gender	Verbal abilities <i>M</i> (<i>SD</i>)
Children	30	9–10	9.7 (0.4)	16 F, 14 M	9.7 (2.5)
Adolescents	30	14–15	14.5 (0.5)	15 F, 15 M	10.5 (1.9)
Young adults	30	20–25	22.5 (1.9)	27 F, 3 M	9.7 (2.6)
Middle-aged adults	31	30–50	41.6 (4.1)	25 F, 5 M	9.5 (1.5)
Old adults	30	60+	66.6 (4.7)	14 F, 16 M	9.0 (1.4)

F = Female, M = male.

adults, neighborhood apps for middle-aged individuals, and e-mail lists of a study program for adults over 50 at a local university. Exclusion criteria for all groups were the presence of a psychiatric or neurological disorder, severe visual or hearing impairment, reading disorders, present or past drug or alcohol abuse.

Participants' verbal abilities were assessed using the German version of the vocabulary test of the Wechsler Intelligence Scale for Children – Fifth Edition (WISC-V, Petermann, 2017), and the Wechsler Adult Intelligence Scale Fourth Edition (WAIS-IV; Petermann, 2012), respectively. Age groups were parallel for age-related scaled scores ($F(4, 146) = 2.044$, $p = 0.09$). All children attended grade 4 (age-range 9–10) of local primary schools, while adolescents (aged between 14 and 15 years) attended grades 8 (20% of adolescents), 9 (50%) and 10 (30%) of different local secondary schools. All adult participants were highly educated. Young adults were all psychology students at the local university; 43.2% of the middle-aged adults and 62.7% of the old adults had at least a master's degree or a higher academic qualification (16.1% of the middle-aged adult group had a PhD). A differentiated presentation of participants' vocational and educational qualifications is provided in the appendix.

The study was conducted in accord with the Declaration of Helsinki and was approved by the local ethics committee. All participants gave written informed consent prior to study participation. Only children and adolescents whose parents had given written informed consent were allowed to participate. The study was not preregistered.

Materials

Prospective memory paradigm

The present study used a shortened computerized version of Virtual Week (Rendell & Craik, 2000) to assess participants' PM performance. Virtual Week was presented as a board game (see Figure 1) on a laptop and assessed four different types of PM tasks (time- and event-based PM tasks, regular and irregular PM tasks). Participants moved a token clockwise around the board after clicking on a dice sign in the middle of the screen, with each circuit representing one waking day. The game featured a virtual clock in the center, which was calibrated to the token's position on the board. Every two squares, 15 virtual minutes elapsed. While individuals moved around the board and landed on or passed an E sign (10 cards per day), they needed to virtually pick an event card and choose one of three options that were appropriate for the virtual time (e.g., breakfast at 8:30 a.m. and the participant had to select their breakfast out of three options). Depending on participants' choice, a dice roll consequence was revealed (e.g., to move on by rolling an even number), participants then had to roll the dice until they met the presented dice roll criterium. Rolling the dice, moving around the board, reading the event cards and making decision about the daily activities

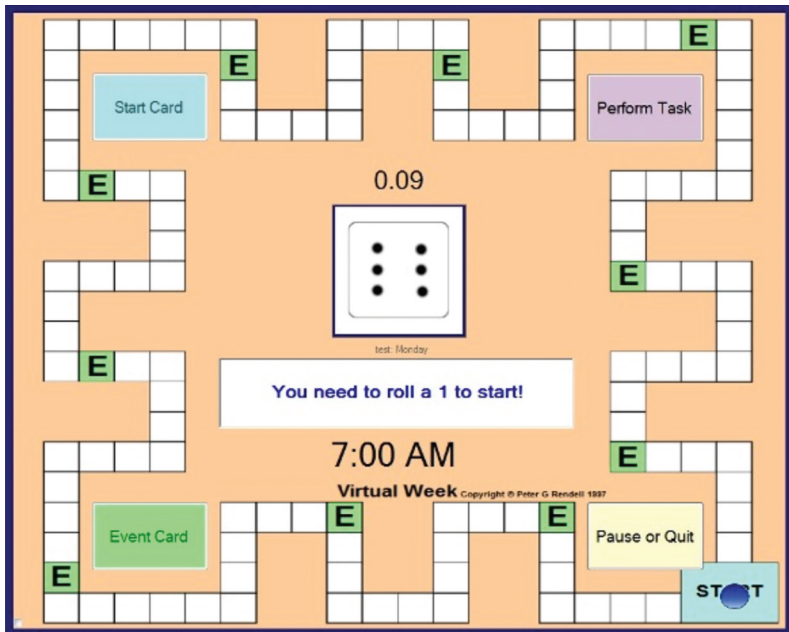


Figure 1. Screenshot of user interface of virtual week, English version.

represented the ongoing task. Into this ongoing task, several time- and event-based PM tasks as well as regular and irregular PM tasks were embedded.

Participants completed 2 days of Virtual Week and one trial day (comprising two irregular time-based and two irregular event-based PM tasks) beforehand to ensure that they understood all features of the game. After completion of the trial day, participants were instructed with the regular PM tasks.

Each Virtual Week testing day included four regular and four irregular PM tasks. The *regular* tasks were the same each day. Participants were informed that they were ill and therefore had to take medication regularly. Two regular PM tasks were time-based (i.e., remember to take asthma spray at 11:00 a.m. and 21:00 p.m.) and two were event-based (i.e., remember to take medication for breakfast and dinner at the appropriate event-cards). The *irregular* PM tasks also consisted of two time-based and two event-based PM tasks; whereby two irregular tasks (one time- and one event-based task) were given at the beginning of a new day and two irregular (again one time- and one event-based task) tasks later appeared during the game.

For the present study, the English adult-version (Rendell & Henry, 2009) was translated and adapted into a German adult-version (Aberle et al., 2010) and a German child/adolescent version (Henry et al., 2014). The adult and children versions differed regarding the content that was presented to make them more suitable to the everyday life of each age group. All task instructions were presented in writing within the game. For children and adolescents, written instructions were slightly shortened. If needed, the children received additional verbal instructions. All children and adolescents had to read out the instructions aloud. After the trial day all participants were asked to repeat the instructions, ensuring understanding of all instructions. The entire task

(trial day plus two testing days) took about half an hour with each circuit around the board lasting about 10 minutes. PM performance for time- and event-based PM tasks were measured as number of executed PM tasks over the two days (in total 8 time-based PM tasks; comprising 4 regular and 4 irregular tasks) and 8 event-based PM tasks (likewise, 4 regular and 4 irregular tasks). A time-based PM performance was considered to have been correct, if the task was remembered between passing the field of the correct time and realized before the next dice roll.

Metacognitive prospective memory inventory short version (mpmi-s)

The MPMI-s (Rummel et al., 2019) is a self-report assessment of individuals everyday PM abilities and everyday PM strategy use. The questionnaire consists of 22 items divided into 3 subscales: 1.) “Prospective Memory Ability Scale” (PMA, 8 items), 2.) “Prospective Memory Internal Scale” (PMSi, 7 items) and 3.) “Prospective Memory External Scale” (PMSe, 7 items). Participants rated their answer on a 5-point Likert scale ranging from (1) “rarely”, (2) “rather rarely”, (3) “sometimes”, (4) “rather often,” to (5) “often.” The original questionnaire was developed for adults (Rummel et al., 2019). To ensure that children and adolescents could comprehend and relate to the items of the MPMI-s, we adapted several items regarding wording and age-appropriate content. The amended questionnaire was piloted in a small sample of children ($N = 12$, age range: 8–12) and confirmed that it was comprehensible for the tested children. For example, we changed the original item: “I forget to cancel contracts on time, like trial subscriptions for newspaper” into “I forget to turn in announcements from school on time.” The reliability of the MPMI-s scales in the present study were high (Cronbach’s alpha from .88 to .91). As the three scales differ in the absolute number of items, z-standardized scores were computed to compare age groups across scales.

Metacognitive monitoring

To assess metacognitive monitoring, all participants were asked to answer questions about their upcoming PM performance (predictions) after going through the trial day of Virtual Week. The first question to answer was: *Overall, how well do you think you will do on the task?* Participants were asked to indicate their performance on a scale from 1 to 6 which was presented as reflecting the German school-grade system and is very familiar to German citizens. The lower the grade, the better the performance. A second question referred to the perceived personal challenge they would face completing the task: *How difficult do you think this task is going to be for you?* Participants rated on a 5-point Likert scale ranging from 1 – very easy to 5 very difficult.

Similarly after completing Virtual Week, participants were asked to rate their performance (postdictions). The first and second questions (1. *Overall, how well do you think you did on the task?*, 2. *How difficult do you think this task was for you?*) were similar to the one’s described above (predictions) only that performance had to be assessed retrospectively. As previous research has shown, that postdictions might be more accurate than predictions (Cottini et al., 2019; Gilbert, 2015; Lavis & Mahy, 2021; Meeks et al., 2007; S. J. Smith et al., 2011), we added three more questions to get a deeper understanding of retrospective performance predictions. We further asked the following: 1.) *How many tasks out of 16 do you think you have completed correctly?* 2.) *How confident are you about your answer?* 3.)

Did you use any strategies to remember the task? Participants were asked to rate their answers for questions 2 and 3 on a 5-point Likert scale ranging from 1 – not confident at all to 5 – completely confident. The final question asked participants to report if any strategies were used. The responses were then analyzed by two independent raters and categorized into four groups: no strategy reported, rehearsal, mnemonics, or episodic future thinking. The versions for children/adolescents and adults differed slightly in language use (e.g., more formal wording). To compare prediction and postdiction accuracy between the five age groups, the mean differences between metacognitive monitoring and actual PM performance were calculated (e.g., Cottini et al., 2019). Therefore, we converted the amount of correct PM tasks into a corresponding grade from 1 to 6. The corresponding number of correct tasks to grade was the following: 16–14 = 1, 13–11 = 2, 10–8 = 3, 7–5 = 4, 4–2 = 5, 1–0 = 6. A difference score of 0 indicated that the performance predictions/postdictions were exactly as estimated, a score below 0 indicated underconfidence, while a score above 0 indicated overestimation.

Procedure

The testing session, which lasted between 45 minutes and 1 hour, was administered individually to all participants. After giving written informed consent, participants filled in a sociodemographic questionnaire which was presented via SoSci-Survey (Leiner, 2019) and included questions such as age, gender, education, and occupation, followed by the MPMI-s. All participants were encouraged to ask questions at any time if something was unclear. Children were repeatedly offered help while filling out the questionnaires ensuring comprehension of all items. Then, participants were introduced to Virtual Week and completed a trial day of either the children/adolescent- or the adult-version of the task, depending on age-related differences in attentional processes and language comprehension. During the trial day, the examiner answered any questions about the procedure and rules to ensure that participants understood the task. After completing the trial day, participants were asked to predict their task performance on Virtual Week, whereupon they completed two more days of the task without receiving any feedback or assistance. Finally, after task completion, participants were asked to retrospectively assess their performance

Results

All statistical analyses were conducted with IBM SPSS Statistics 27.0.

Prospective memory performance (laboratory)

The mean accuracy for correctly performed PM responses overall and separately for correct time- and event-based PM trials through all five age groups are displayed in [Table 2](#).

A 2 (PM cue: PM time-based, PM event-based) by 5 (age groups: children, adolescents, young adults, middle-aged adults, old adults) mixed analysis of variance (ANOVA) was carried out to analyze differences in PM performance in time- versus event-based PM across the five different age groups. The results revealed significant main effects of PM cue, $F(1, 146) = 30.3, p < 0.001, \eta_p^2 = .172$ and age group, $F(4, 146) = 17.43, p < 0.001, \eta_p^2 = .323$. Post hoc Tukey tests indicated that overall, participants showed superior event-based as

compared to time-based PM performance (see Table 2). Furthermore, significant differences in PM performance between children and adolescents, children and young adults, as well as children and middle-aged adults (all p s < .01), but no significant differences between children and old adults ($p = 0.99$). The performance of adolescents, young adults and middle-aged adults did not differ (all p s > .95). The interaction between PM cue \times age group was not significant, $F(4,146) = 1.25$, $p = .29$, $\eta_p^2 = .033$, indicating no differential age effects on time- vs. event-based PM performance.

Everyday prospective memory performance

To assess differences in self-rated everyday PM performance (using z values, see Table 3) a second mixed ANOVA with the within-subject factor scale (PM Ability, Internal, External) and between-subject factor age group (children, adolescents, young-, middle-aged-, old adults) was conducted. There was no significant main effect of scale, $F < 1$, but a significant main effect of age group, $F(4, 146) = 13.0$, $p < 0.001$, $\eta_p^2 = .263$. Post hoc Tukey tests revealed that there was no significant difference between children and adolescents ($p = 0.99$), but both children and adolescents differed significantly from young -, middle-aged - and old adults (all p s < .01). The interaction between scale \times age group was significant, $F(8,292) = 8.8$, $p < 0.001$, $\eta_p^2 = .194$.

To better understand the two-way interaction, follow-up analyses were conducted. Analyses of simple effects of age groups separately for each scale revealed the following: There was a significant main effect on "PM ability scale," $F(4,146) = 14.29$, $p < .001$, $\eta_p^2 = 0.28$. Children reported significantly better PM performance than adolescents ($p < .009$) but significantly poorer performance than old adults ($p < .001$). There was no difference between children and young adults and middle-aged adults on this scale (all p s > .195). Adolescents reported the lowest PM performance compared to all other age groups (all p s < .009). In addition, young adults showed significantly lower PM performance compared to

Table 2. Prospective memory performance in the virtual week.

	PM total	time-based PM	event-based PM
	<i>M (SD)</i>	<i>M (SD)</i>	<i>M (SD)</i>
Children	7.27 (3.9)	2.93 (2.05)	4.33 (2.23)
Adolescents	11.37 (3.0)	5.4 (1.98)	5.97 (1.67)
Young adults	11.97 (2.46)	5.27 (1.82)	6.7 (1.27)
Middle-aged A,	11.35 (3.31)	5.19 (2.12)	6.16 (1.81)
Old adults	6.9 (2.39)	3.2 (2.19)	3.7 (2.18)

Table 3. Summary scores and z -standardized scores for all three mpmi-s scales.

	PM Ability Scale		Internal Scale		External Scale	
	<i>M (SD)</i>	<i>z (SD)</i>	<i>M (SD)</i>	<i>z (SD)</i>	<i>M (SD)</i>	<i>z (SD)</i>
Children	29.07 (6.8)	-.15 (1.23)	18.6 (6.95)	-.64 (1.16)	19.07 (6.05)	-.77 (.90)
Adolescents	25.8 (3.01)	-.74 (.54)	22.4 (5.83)	-.00 (.97)	20.5 (6.04)	-.57 (.90)
Young adults	30.67 (4.4)	.13 (.79)	24.33 (4.52)	.32 (.75)	27.3 (5.03)	.44 (.75)
Middle-aged A.	29.23 (5.3)	-.13 (.96)	24.19 (5.57)	.30 (.93)	26.45 (6.29)	.32 (.94)
Old adults	34.87 (3.24)	.89 (.59)	22.53 (5.36)	.02 (.89)	28.07 (4.58)	.56 (.68)

old adults ($p < .001$) but no difference to middle-aged adults ($p = 0.239$). Old adults reported the best everyday PM performance compared to all other groups (all $ps < .001$).

A significant main effect was also shown for the “Internal strategy scale,” $F(4,146) = 4.97$, $p < .001$, $\eta_p^2 = 0.12$. Children reported the lowest use of internal strategies compared to all other age groups (all $ps < .01$), while adolescents, young, middle-aged and old adults did not differ in their internal strategy use (all $ps > .191$). Finally, there was also a significant main effect for “External strategy scale,” $F(4,146) = 16.44$, $p < .001$, $\eta_p^2 = 0.31$. Children and adolescents reported significantly lower external strategy use compared to young, middle-aged and old adults (all $ps < .001$). There was no difference between children and adolescents ($p = 0.327$). In addition, all adults groups did not differ in their usage of external strategies (all $ps > .266$).

We further analyzed each scale separately for all age groups and found a significant main effect for children, $F(2,145) = 6.107$, $p < .003$, $\eta_p^2 = 0.08$, showing that the report of everyday PM performance was significantly higher than the report of both, internal and external strategy use (all $ps < .009$). The report of internal and external strategy use from children did not differ ($p = 0.384$). We also found a significant main effect for adolescents, $F(2,145) = 9.932$, $p < .001$, $\eta_p^2 = 0.12$, indicating a significantly lower report of everyday PM performance compared to internal strategy use ($p < .001$) but not to external strategy use ($p = 0.331$). Adolescents reported a significantly higher internal strategy use than external strategy use ($p < .001$). There was no main effect for the young adult group, $F(2,145) = 1.474$, $p = 0.232$, $\eta_p^2 = 0.02$ but for middle-aged adults, $F(2,145) = 3.646$, $p = 0.029$, $\eta_p^2 = 0.05$. Middle-aged adults reported significantly lower everyday PM performance than internal and external strategy use (all $ps < .02$) but reported use of internal and external strategies ($p = 0.866$) did not differ. Finally, there was also a significant main effect for old adults, $F(2,145) = 12.421$, $p < .001$, $\eta_p^2 = 0.15$. Old adults reported significantly higher everyday PM performance than the use of internal strategies ($p < .001$) while external strategies use did not differ ($p = 0.073$). In addition, old adults reported significantly higher use of external strategies compared to internal strategy use ($p < .001$).

Prediction and postdiction accuracy

A 2 (time: prediction accuracy, postdiction accuracy) \times 5 (age group: children, adolescents, young, middle-aged, old adults) mixed ANOVA was conducted to explore the effects of time of judgment on age groups’ assumed PM performance (see Figure 2). There was a significant effect of time of judgment, $F(1, 146) = 42.9$, $p < 0.001$, $\eta_p^2 = .227$. Overall, participants overestimated their upcoming performance, while they underestimated their past performance. There was also a significant main effect of age group, $F(4, 146)$, $p < 0.001$, $\eta_p^2 = .322$. Post hoc Tukey test revealed significant differences between children and all other age groups (all $ps < .02$), as well as between old adults and all other age groups (all $ps < .02$) for predictions and postdictions. Except for young adults who tended to underestimate their upcoming performance, all other age groups overestimated their predicted PM performance. In contrast, only children and old adults overestimated their past performance, while adolescents, young adult and middle-aged adults underestimated their performance in postdictions. The interaction between time of judgment \times age group was not significant, $F(4, 146) = 0.201$, $p < 0.94$, $\eta_p^2 = .005$.

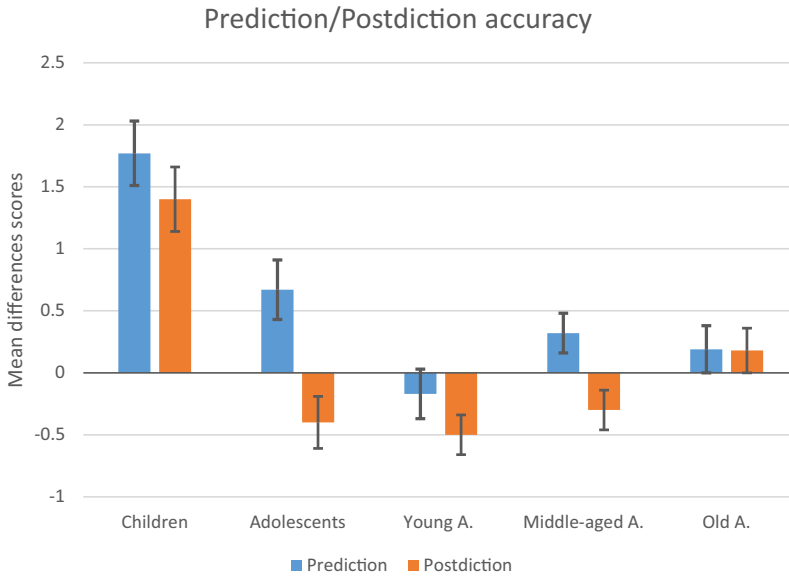


Figure 2. Prediction and postdiction accuracy scores across the five age groups. Note: Error bars are presented as standard errors.

Confidence judgements of postdictions

We analyzed the confidence of judgments for postdictions across the five age groups. The mean confidence ratings are displayed in Table 4. A one-way ANOVA showed a significant effect of age groups, $F(4, 146) = 4.709$, $p < 0.001$, $\eta_p^2 = .114$. Post hoc analyses showed significant differences between children and old adults ($p = 0.02$), middle-aged adults and young adults ($p = 0.04$), and between middle-aged adults and old adults ($p = 0.01$). Overall, middle-aged adults demonstrated the highest level of confidence in their retrospective judgments, closely followed by children, while old adults trusted their judgments the least.

Table 4. Descriptive statistics for pre- and postdictions across the age groups.

	Children	Adolescents	Young Adults	Middle-aged Adults	Old Adults
	<i>M</i> (<i>SE</i>)	<i>M</i> (<i>SE</i>)	<i>M</i> (<i>SE</i>)	<i>M</i> (<i>SE</i>)	<i>M</i> (<i>SE</i>)
Predictions					
(1) Overall, how well do you think you will do the task? (max. 6*)	1.77 (0.73)	2.07 (0.69)	2.17 (0.74)	2.19 (0.65)	(0.82)
(2) How difficult do you think this task is going to be for you?(max 5)	2.10 (0.92)	2.33 (0.76)	2.47 (0.86)	2.23 (0.56)	2.87 (0.86)
Postdictions					
(1) Overall, how well do you think you completed the task? (max. 6)	2.13 (0.82)	2.53 (0.73)	2.5 (0.77)	2.52 (0.63)	3.23 (0.79)
(2) How difficult did you experienced this task?(max.5)	2.37 (1.0)	2.83 (0.87)	2.97 (0.89)	2.52 (0.96)	3.5 (0.78)
(3) How many tasks out of 16 do you think you've completed correctly?	11.97 (2.37)	12 (2.27)	12.17 (2.26)	11.52 (3.23)	10.23 (2.46)
(4) How confident are you about your answer? (max. 5)	3.77 (0.68)	3.57 (0.82)	3.37 (0.81)	3.87 (0.62)	3.2 (0.55)

* We used the German grading system; a lower number indicates a better performance..

Correlation between metacognitive monitoring and prospective memory

Separately for each age group, correlational analyses were carried out to analyze the relationships between PM performance (total), PM prediction accuracy, PM postdiction accuracy, postdiction confidence judgment and MPMI-s scales (see Table 5). For all age groups, predictions accuracy correlated negatively with PM performance, indicating that better PM performance is associated with better PM prediction accuracy (CAVE: the lower the accuracy value, the greater the agreement between prediction and actual performance). Similarly, for postdiction accuracy across all age groups, better PM performance was associated with better judgment of PM performance after completing the PM task. In addition, there was a highly significant positive correlation between PM prediction accuracy and PM postdiction accuracy across all age groups, indicating that individuals that had

Table 5. Correlational analyses.

	prediction accuracy	postdiction accuracy	postdiction confidence	PM Ability Scale	Internal Scale	External Scale
Children						
PM total	-.85**	-.80**	-.07	.21	-.10	-.12
prediction accuracy		.91**				
postdiction accuracy			.43*			
PM Ability Scale					.69**	.70**
Internal Scale						.76**
Adolescents						
PM total	-.83**	-.73**	.17	.11	.17	.33
prediction accuracy		.85**				
postdiction accuracy			.25			
PM Ability Scale					.02	-.11
Internal Scale						.38*
Young adults						
PM total	-.66**	-.52*	.42*	.10	-.14	-.07
prediction accuracy		.74**				
postdiction accuracy			.02			
PM Ability Scale					.50**	.06
Internal Scale						.27
Middle-aged adults						
PM total	-.78**	-.81**	.28	-.05	-.03	.22
prediction accuracy		.67*				
postdiction accuracy			-.07			
PM Ability Scale					.27	.27
Internal Scale						.62**
Old adults						
PM total	-.66**	-.76**	.38*	.01	-.08	.05
prediction accuracy		.74**				
postdiction accuracy			-.21			
PM Ability Scale					.14	.32
Internal Scale						.40*

*= $p < 0.05$, **= $p < 0.01$.

Table 6. Spontaneous strategy use while performing virtual week.

	No strategy reported	Rehearsal	Mnemonics	Episodic Future Thinking
	<i>n</i> (%)	<i>n</i> (%)	<i>n</i> (%)	<i>n</i> (%)
Children	24 (80%)	2 (6.7%)	5 (16.7%)	-
Adolescents	14 (46.7%)	10 (33.3%)	5 (16.7%)	2 (6.7%)
Young Adults	-	18 (52.4%)	16 (47.1%)	-
Middle-aged Adults	19 (66.3%)	5 (16.7%)	8 (26.7%)	-
Old adults	9 (30%)	9 (30%)	10 (33.3%)	2 (6.7%)

Multiple answers per person were possible.

a higher accuracy in predicting their upcoming performance also had a better accuracy in their retrospective assessment.

PM performance in the laboratory and everyday PM performance did not correlate in any age group. Similarly, there was no correlation between laboratory PM performance and the reported use of internal or external strategies in everyday life, measured with the Internal and External Scale of the MPMI-s. For children self-reported everyday PM performance and the use of internal or external strategies was significantly related. Similarly, young adults showed a significant correlation between everyday PM performance and internal strategy use ($r = .50, p < .01$). For all other age groups, there were no significant relations between reported everyday PM performance and the use of internal or external strategies.

Except for young adults, there were significant correlation between the reported use of internal and external strategies.

Strategy use

While only 20% of children ($n = 6$) reported to have spontaneously applied strategies while completing the Virtual Week task, strategy use increased with age [adolescents: 53,3% ($n = 16$); young adults: 100% ($n = 30$); middle-aged adults: 35,5% ($n = 11$); old adults: 70% ($n = 21$)]. Overall, children and middle-aged adults reported the least use of spontaneous strategies whereas all young adults indicated to have used strategy. Answers are displayed in Table 6. The responses were assigned into three categories: rehearsal, mnemonics, and episodic future thinking, whereby rehearsal and mnemonics were selected most frequently across all age groups.

Discussion

The current study set out to investigate the impact of metacognitive monitoring on the development of PM performance across the lifespan. To this end, we examined age-related differences in PM performance in five age groups ranging from childhood to old adulthood using the same paradigm (Virtual Week; Rendell & Craik, 2000) and investigated developmental differences in prospective and retrospective metacognitive monitoring on PM performance.

In line with our expectations and previous studies that found PM to develop in an inverted U-shaped function across the lifespan (Kliegel et al., 2008; Kretschmer-Trendowicz & Altgassen, 2016; Mattli et al., 2014; Zimmermann & Meier, 2006), our results revealed

children and old adults to show the poorest PM performance, while adolescents, young adults and middle-aged adults all performed at a similar level and outperformed children and old adults. The increase and later decrease of PM performance across the lifespan has been linked to the parallel development of executive functions and retrospective memory which are both critically involved in prospective remembering (Mahy et al., 2014; Martin et al., 2003).

Regarding the assessment of developmental differences in *prospective* metacognitive monitoring, in line with our hypothesis, children overestimated their upcoming PM performance but contrary to our expectations, only young adults underestimated their upcoming PM performance, while adolescents, middle-aged adults and old adults also overestimated their upcoming performance. Regarding participants' *retrospective* PM performance judgments (postdictions), overall, as hypothesized, accuracy judgments given after task performance were more accurate than those given before performing the task (cf. similar results, Cottini et al., 2019; Gilbert, 2015; Lavis & Mahy, 2021; Meeks et al., 2007). Overall, children and old adults overestimated their past PM performance, while adolescents, young adults and middle-aged adults underestimated it.

Regarding the children group, the observed overestimation of their future and past PM performance was in line with the current literature (Cottini et al., 2021; Lavis & Mahy, 2021). Cottini and colleagues compared younger (5–6 years) and older (8–10 years) children in their metacognitive monitoring in an event-based PM task (three targets for young children and four targets for older children). Both groups overestimated their upcoming performance, but older children were significantly better at predicting their performance compared to younger children. Cottini's results (2021) concur with developmental studies in the area of retrospective memory research which indicate that 9- to 10-year-old children become more accurate and less overconfident than younger children aged 6–7 years (Schneider et al., 2000). The children group of the current study was of a similar age as the older children group of Cottini's study. Descriptively, children in our study were more overconfident in predicting their performance than Cottini's sample, possibly due to the different tasks used and the larger amount of PM targets. The current study comprised 16 PM tasks (both, time- and event-based tasks, including both regular and irregular tasks) which represents a significant increase compared to the three, respectively, four PM tasks in the Cottini study, and may have made it harder for children to accurately predict their performance. Future studies should vary the number of PM tasks to investigate its impact on children's ability to assess their future performance. Consistent with Cottini et al. (2021) results, we found children's performance judgments given retrospectively (postdictions) to be more accurate than those given prospectively (predictions).

This was the first study to investigate metacognitive monitoring in PM in adolescents. Notably, studies on retrospective memory are also scarce (Godfrey et al., 2023). The limited research available indicates that children aged 10–13 years tend to become more accurate in predicting their past memory performance, approaching the accuracy of adults. Consistent with these findings, our study-observed adolescents were significantly more accurate in predicting their performance than children (while adolescents were on average about 4.5 PM tasks more positive in their prediction than they actually performed, children overestimated their performance by about 5.5 PM tasks); though they were still rather optimistic about their future performance compared to all adult groups (adults overestimated or underestimated themselves in their predictions by about 1 to 1.5 PM tasks). Given that

the used PM task Virtual Week tries to mirror everyday like demands – and despite using the adapted child version of the task -, it is possible that children and adolescents had more difficulties with assessing their abilities as both children and adolescents may not have not been fully exposed yet to the various demands of life. Possibly, children tend to be less accurate in their assessments than adolescents who already show more responsibility for different aspects of their lives and may therefore be better equipped to provide more accurate assessments. In addition, during the testing session children often reported that they were generally very good at computer games, so possibly the presentation of the task as a board game may have led children to assume that their performance in Virtual Week will be as good as their perceived everyday game experience. Another possibility is that children did not realize that they should predict their memory performance and not their game performance, resulting in their assessment being influenced by their gaming experience rather than their memory performance.

Regarding the adult groups, our predictions were confirmed for young and old adults, but not for middle-aged adults who slightly overestimated their upcoming performance. Interestingly, among all age groups, young adults showed the most accurate prospective assessment of their upcoming performance, but were even more underconfident in their retrospective assessment than they were initially. On the other hand, middle-aged and older adults showed a comparable level of accuracy in their performance assessments, although the direction of their assessment in middle-aged adults changed from overestimation (predictions) to underestimation (postdictions). A possible explanation for individuals' tendency to underestimate their retrospective performance could be participants' modesty or their attempt to give answers that are socially desirable. Adults appear to exhibit greater caution and reservation in their retrospective performance assessments while demonstrating slightly more confidence before completing the task, these differences in performance pre- and postdictions could potentially be due to the absence of prior experience with the task and resulting difficulties to assess their upcoming task performance. Another possible explanation for this observation could be participants' modesty or social desirability given that typically in society it is better to downplay one's performance than to boast about it.

Overall, all adult groups were generally rather accurate in their performance assessments, whether made prospectively or retrospectively. This contrasts with earlier studies that suggested that adults are not good at predicting their upcoming performance (Kuhlmann, 2019). One reason for this contrasting result could be the design of the task used. The PM tasks of Virtual Week aim to reflect everyday tasks that should be familiar to most adults (e.g., remembering to call a friend at a certain time or to buy a certain item when passing the supermarket), making it easier for them to predict their performance. This would also explain why the retrospective judgments were only slightly better than those given prospectively.

Taken together, our results suggest that – with the exception of children – individuals seem to be able to assess their PM performance in the laboratory, thus showing accurate metacognitive abilities. Interestingly though, being able to accurately predict one's success and failures does not automatically lead to better PM performance as seen in our older adult group who accurately predicted a poor PM performance. A possible explanation for this could be that older adults are aware of their deficits, but in order to compensate for them, the use of more external strategies is needed (e.g., setting reminders or writing notes) which they were prevented from using in this laboratory task setting. Possibly, a greater use of

compensatory strategies in everyday life by older adults may also explain the findings of the age-PM paradox (Henry et al., 2004; Rendell & Thomson, 1999), namely that older adults perform worse in laboratory-based PM measurements than young adults, but show spared or even better performance in more naturalistic settings. Similarly, it remains unclear if older adults did not use any internal strategies in the present study or if they could not use them efficiently to cope with the demands of the lab-based task. Future studies should address the impact of internal and external strategies on PM performance in the lab and in everyday life.

In addition to developmental differences in prediction and postdiction accuracy, we also investigated differences in retrospective confidence judgments across the lifespan. Middle-aged adults were most confident in their postdictions, followed by children, adolescents and young adults. Old adults were least confident; which is in line with previous evidence reporting a tendency of older adults to rely less on their cognitive abilities (e.g., Bouazzaoui et al., 2016; Lineweaver & Hertzog, 1998).

Two key findings emerged across all age groups. Firstly, it seems that the timing of the performance judgment may not be as critical as previously found; at least in PM tasks that are closer to everyday tasks. A high correlation was found between prediction accuracy and postdiction accuracy, indicating that individuals who made more accurate predictions of their future performance also made more accurate postdictions afterward. This could suggest that with regard to more naturalistic laboratory-based PM tasks individuals may have a good understanding of their PM abilities. Second, better PM performance was associated with higher levels of prediction and postdiction accuracy for all age groups. More research is needed to disentangle whether PM and metacognition develop in parallel or whether they are mutually dependent in their development.

The current results align with theoretical perspectives suggesting that metacognitive abilities, such as understanding one's cognitive strengths or selecting effective strategies, are essential for the success and development of PM (Kuhlmann, 2019; R. E. Smith, 2016). Effective metacognitive monitoring skills can aid individuals in completing a PM task by allowing them to monitor, reflect on, and evaluate their cognitive abilities, and select the most suitable strategy for implementing their intention. In contrast, poor metacognitive skills such as individuals' overconfidence about their performance (e.g., children and adolescents) may lead to poor prospective remembering as individuals may invest less effort (e.g., monitoring, strategy use) than is necessary for successful intention completion. Providing children and adolescents with information on internal and external metacognitive strategies could enhance their PM performance, leading to greater success and independence in various aspects of their lives (e.g. learning for school or preparing for a sports competition).

Finally, we explored self-reported everyday PM performance across all five age groups. Results revealed a different developmental pattern compared to the laboratory PM assessment. Adolescents who demonstrated descriptively the best performance in the lab-based PM task, reported the lowest everyday PM performance. On the other hand, old adults who showed the lowest performance in the laboratory PM task reported the best everyday PM performance. Interestingly, children reported comparable everyday PM to young adults and middle-aged adults, which is in strong contrast to their lab-based PM performance where children showed significantly lower performance compared to young and middle-aged adults. Possibly, children did not fully understand the content of the items of the

questionnaire which could have affected the validity of the results. However, it is worth noting that regarding the two scales assessing strategy use, both internal and external strategies, children reported the least use of these strategies. This is consistent with previous studies suggesting that children may be less adept at using strategies to achieve their goals (Cottini et al., 2019). Importantly, laboratory-measured PM performance may not be consistent with everyday PM performance (e.g. Rummel et al., 2019; Schnitzspahn et al., 2016).

Overall, the use of questionnaires to assess PM in everyday life is still under debate. While most studies do not find a correlation between self-report data and laboratory measures and thus question the validity of questionnaires for PM measurement (Kliegel & Jäger, 2007; Uttl & Kibreab, 2011), Rummel et al. (2019) emphasize the importance of self-report measures, especially for clinical settings when a quick assessment of self-perceived PM skills and strategy use is required. Importantly, our findings indicate that the self-perception of individuals of their accomplished PM performance (at least from adolescence onwards) seems to be well calibrated. This could indicate that retrospective self-assessments of one's own PM performance could be more accurate than sometimes assumed. However, future research is needed to investigate whether people are also accurate in their retrospective assessment of their everyday PM performance, not only just when assessing an everyday-like laboratory task.

This study set out to gain, for the first time, a better understanding of the development of prospective metacognitive monitoring and PM performance across the lifespan. By examining five age groups, we confirmed the previously observed inverted U-shaped trajectory of PM performance across the lifespan. With regard to metacognition, children tended to strongly overestimate their PM performance, whereas, the accuracy of metacognitive monitoring abilities increased than from adolescents into old adulthood, both prospectively and retrospectively. Overall, retrospective assessments of PM performance were more accurate than prospective assessments. It can therefore be concluded that from adolescence onwards, individuals are very well able to assess their laboratory PM performance. Nevertheless, for old adults accurate knowledge of future performance in the lab-based assessment does not appear to lead to the selection of appropriate internal strategies that result in an increase in performance. More research is needed to further understand the relationship between metacognitive monitoring and PM performance.

Disclosure statement


No potential conflict of interest was reported by the author(s).

Funding

The author(s) reported that there is no funding associated with the work featured in this article.

ORCID

Steven Stelz  <http://orcid.org/0009-0000-8835-2016>

Mareike Altgassen  <http://orcid.org/0000-0002-2644-8298>

Availability of data and material

This study was not preregistered. The datasets used and/or analyzed during the current study are available from the corresponding author on request.

Statement of ethics

The study was conducted in accord with the Declaration of Helsinki and was approved by the local ethics committee. All participants gave written informed consent prior to study participation. Only children and adolescents whose parents had given written informed consent were allowed to participate.

References

- Aberle, I., Rendell, P. G., Rose, N. S., McDaniel, M. A., & Kliegel, M. (2010). The age prospective memory paradox: Young adults may not give their best outside of the lab. *Developmental Psychology*, 46(6), 1444–1453. <https://doi.org/10.1037/a0020718>
- Altgassen, M., Kretschmer, A., & Schnitzspahn, K. M. (2017). Future thinking instructions improve prospective memory performance in adolescents. *Child Neuropsychology: A Journal on Normal & Abnormal Development in Childhood & Adolescence*, 23(5), 536–553. <https://doi.org/10.1080/09297049.2016.1158247>
- Ardila, A. (2008). On the evolutionary origins of executive functions. *Brain & Cognition*, 68(1), 92–99. <https://doi.org/10.1016/j.bandc.2008.03.003>
- Bertrand, J. M., Moulin, C. J. A., & Souchay, C. (2017). Short-term memory predictions across the lifespan: Monitoring span before and after conducting a task. *Memory*, 25(5), 607–618. <https://doi.org/10.1080/09658211.2016.1200625>
- Best, J. R., & Miller, P. H. (2010). A developmental perspective on executive function. *Child Development*, 81(6), 1641–1660. <https://doi.org/10.1111/j.1467-8624.2010.01499.x>
- Bouazzaoui, B., Follenfant, A., Ric, F., Fay, S., Croizet, J. C., Atzeni, T., & Taconnat, L. (2016). Ageing-related stereotypes in memory: When the beliefs come true. *Memory*, 24(5), 659–668. <https://doi.org/10.1080/09658211.2015.1040802>
- Cottini, M., Basso, D., & Palladino, P. (2018). The role of declarative and procedural metamemory in event-based prospective memory in school-aged children. *Journal of Experimental Child Psychology*, 166, 17–33. <https://doi.org/10.1016/j.jecp.2017.08.002>
- Cottini, M., Basso, D., Pieri, A., & Palladino, P. (2021). Metacognitive monitoring and control in children's prospective memory. *Journal of Cognition and Development*, 22(4), 619–639. <https://doi.org/10.1080/15248372.2021.1916500>
- Cottini, M., Basso, D., Saracini, C., & Palladino, P. (2019). Performance predictions and postdictions in prospective memory of school-aged children. *Journal of Experimental Child Psychology*, 179, 38–55. <https://doi.org/10.1016/j.jecp.2018.10.008>
- Dunlosky, J., & Metcalfe, J. (2008). *Metacognition*. Sage Publications.
- Einstein, G. O., & McDaniel, M. A. (1990). Normal aging and prospective memory. *Journal of Experimental Psychology: Learning, Memory & Cognition*, 16(4), 717–726. <https://doi.org/10.1037//0278-7393.16.4.717>
- Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* power, 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175–191. <https://doi.org/10.3758/bf03193146>
- Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *The American Psychologist*, 34(10), 906–911. <https://doi.org/10.1037/0003-066X.34.10.906>
- Gilbert, S. J. (2015). Strategic use of reminders: Influence of both domain-general and task-specific metacognitive confidence, independent of objective memory ability. *Consciousness and Cognition*, 33, 245–260. <https://doi.org/10.1016/j.concog.2015.01.006>

- Godfrey, M., Casnar, C., Stolz, E., Ailion, A., Moore, T., & Gioia, G. (2023). A review of procedural and declarative metamemory development across childhood. *Child Neuropsychology: A Journal on Normal and Abnormal Development in Childhood and Adolescence*, 29(2), 183–212. <https://doi.org/10.1080/09297049.2022.2055751>
- Guajardo, N. R., & Best, D. L. (2000). Do preschoolers remember what to do? Incentive and external cues in prospective memory. *Cognitive Development*, 15(1), 75–97. [https://doi.org/10.1016/S0885-2014\(00\)00016-2](https://doi.org/10.1016/S0885-2014(00)00016-2)
- 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. <https://doi.org/10.1037/0882-7974.19.1.27>
- Henry, J. D., Terrett, G., Altgassen, M., Raponi-Saunders, S., Ballhausen, N., Schnitzspahn, K. M., & Rendell, P. G. (2014). A virtual week study of prospective memory function in autism spectrum disorders. *Journal of Experimental Child Psychology*, 127, 110–125. <https://doi.org/10.1016/j.jecp.2014.01.011>
- Hering, A., Kliegel, M., Rendell, P. G., Craik, F. I., & 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. <https://doi.org/10.1017/S1355617718000152>
- Hoffmann-Biencourt, A., Lockl, K., Schneider, W., Ackerman, R., & Koriat, A. (2010). Self-paced study time as a cue for recall predictions across school age. *British Journal of Developmental Psychology*, 28(4), 767–784. <https://doi.org/10.1348/026151009X479042>
- Kerns, K. A. (2000). The CyberCruiser: An investigation of development of prospective memory in children. *Journal of the International Neuropsychological Society*, 6(1), 62–70. <https://doi.org/10.1017/S1355617700611074>
- Kliegel, M., & Jäger, T. (2007). The effects of age and cue-action reminders on event-based prospective memory performance in preschoolers. *Cognitive Development*, 22(1), 33–46. <https://doi.org/10.1016/j.cogdev.2006.08.003>
- Kliegel, M., Mackinlay, R., & Jäger, T. (2008). Complex prospective memory: Development across the lifespan and the role of task interruption. *Developmental Psychology*, 44(2), 612–617. <https://doi.org/10.1037/0012-1649.44.2.612>
- Kliegel, M., Martin, M., McDaniel, M. A., & Einstein, G. O. (2002). Complex prospective memory and executive control of working memory: A process model. *Psychologische Beiträge*, 44, 303–318.
- Knight, R. G., Harnett, M., & Titov, N. (2005). The effects of traumatic brain injury on the predicted and actual performance of a test of prospective remembering. *Brain Injury*, 19(1), 19–27. <https://doi.org/10.1080/02699050410001720022>
- Koriat, A., Ma'ayan, H., & Nussinson, R. (2006). The intricate relationships between monitoring and control in metacognition: Lessons for the cause-and-effect relation between subjective experience and behavior. *Journal of Experimental Psychology General*, 135(1), 36–69. <https://doi.org/10.1037/0096-3445.135.1.36>
- Kretschmer-Trendowicz, A., & Altgassen, M. (2016). Event-based prospective memory across the lifespan: Do all age groups benefit from salient prospective memory cues? *Cognitive Development*, 39, 103–112. <https://doi.org/10.1016/j.cogdev.2016.04.005>
- Kuhlmann, B. G. (2019). Metacognition of prospective memory: Will I remember to remember? In J. Rummel & M. A. McDaniel (Eds.), *Prospective memory* (pp. 60–77). Routledge.
- Kvavilashvili, L., & Ford, R. M. (2014). Metamemory prediction accuracy for simple prospective and retrospective memory tasks in 5-year-old children. *Journal of Experimental Child Psychology*, 127, 65–81. <https://doi.org/10.1016/j.jecp.2014.01.014>
- Lavis, L., & Mahy, C. E. V. (2021). “I’ll remember everything no matter what!”: The role of metacognitive abilities in the development of young children’s prospective memory. *Journal of Experimental Child Psychology*, 207, 105117. <https://doi.org/10.1016/j.jecp.2021.105117>
- Leiner, D. J. (2019). SoSci survey (Version 3.1. 06). *Computer software*.
- Lineweaver, T. T., & Hertzog, C. (1998). Adults’ efficacy and control beliefs regarding memory and aging: Separating general from personal beliefs. *Aging, Neuropsychology & Cognition*, 5(4), 264–296. <https://doi.org/10.1076/anec.5.4.264.771>

- Mahy, C. E. V., & Moses, L. J. (2011). Executive functioning and prospective memory in young children. *Cognitive Development*, 26(3), 269–281. <https://doi.org/10.1016/j.cogdev.2011.06.002>
- Mahy, C. E. V., Moses, L. J., & Kliegel, M. (2014). The development of prospective memory in children: An executive framework. *Developmental Review*, 34(4), 305–326. <https://doi.org/10.1016/j.dr.2014.08.001>
- Martin, M., Kliegel, M., & McDaniel, M. A. (2003). The involvement of executive functions in prospective memory performance of adults. *International Journal of Psychology*, 38(4), 195–206. <https://doi.org/10.1080/00207590344000123>
- Mattli, F., Schnitzspahn, K. M., Studerus-Germann, A., Brehmer, Y., & Zöllig, J. (2014). Prospective memory across the lifespan: Investigating the contribution of retrospective and prospective processes. *Aging, Neuropsychology & Cognition*, 21(5), 515–543. <https://doi.org/10.1080/13825585.2013.837860>
- McDaniel, M. A., & Einstein, G. O. (2000). Strategic and automatic processes in prospective memory retrieval: A multiprocess framework. *Applied Cognitive Psychology*, 14(7), S127–S144. <https://doi.org/10.1002/acp.775>
- Meeks, T. J., Hicks, J. L., & Marsh, R. L. (2007). Metacognitive awareness of event-based prospective memory. *Consciousness and Cognition*, 16(4), 997–1004. <https://doi.org/10.1016/j.concog.2006.09.005>
- Morton, J. B., Bosma, R., & Ansari, D. (2009). Age-related changes in brain activation associated with dimensional shifts of attention: An fMRI study. *Neuroimage*, 46(1), 249–256. <https://doi.org/10.1016/j.neuroimage.2009.01.037>
- O’Leary, A. P., & Sloutsky, V. M. (2019). Components of metacognition can function independently across development. *Developmental Psychology*, 55(2), 315–328. <https://doi.org/10.1037/dev0000645>
- Petermann, F. (2012). *WAIS-IV. Wechsler adult intelligence scale* (4th ed.). *Deutschsprachige Adaptation der WAIS-IV von D. Wechsler*. Pearson Assessment.
- Petermann, F. (2017). *WISC-V. Wechsler intelligence scale for children* (5th ed.). *Deutschsprachige Adaptation der WISC-V von David Wechsler*. Pearson Assessment.
- 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. <https://doi.org/10.1002/Acp.770>
- Rendell, P. G., & Henry, J. D. (2009). A review of virtual week for prospective memory assessment: Clinical implications. *Brain Impairment*, 10(1), 14–22. <https://doi.org/10.1375/brim.10.1.14>
- Rendell, P. G., & Thomson, D. M. (1999). Aging and prospective memory: Differences between naturalistic and laboratory tasks. *Journals of Gerontology, Series B: Psychological Sciences & Social Sciences*, 54B(4), PP256–P269. <https://doi.org/10.1093/geronb/54b.4.p256>
- Rummel, J., Danner, D., & Kuhlmann, B. G. (2019). The short version of the metacognitive prospective memory inventory (MPMI-s): Factor structure, reliability, validity, and reference data. *Measurement Instruments for the Social Sciences*, 1(1), 1–8. <https://doi.org/10.1186/s42409-019-0008-6>
- Rummel, J., & Meiser, T. (2013). The role of metacognition in prospective memory: Anticipated task demands influence attention allocation strategies. *Consciousness and Cognition: An International Journal*, 22(3), 931–943. <https://doi.org/10.1016/j.concog.2013.06.006>
- Schneider, W., Schlagmüller, M., & Visé, M. (1998). The impact of metamemory and domain-specific knowledge on memory performance. *European Journal of Psychology of Education*, 13(1), 91–103. <https://doi.org/10.1007/BF03172815>
- Schneider, W., Visé, M., Lockl, K., & Nelson, T. O. (2000). Developmental trends in children’s memory monitoring: Evidence from a judgment-of-learning task. *Cognitive Development*, 15(2), 115–134. [https://doi.org/10.1016/S0885-2014\(00\)00024-1](https://doi.org/10.1016/S0885-2014(00)00024-1)
- Schnitzspahn, K. M., Stahl, C., Zeintl, M., Kaller, C. P., & Kliegel, M. (2013). The role of shifting, updating, and inhibition in prospective memory performance in young and older adults. *Developmental Psychology*, 49(8), 1544–1553. <https://doi.org/10.1037/a0030579>
- Schnitzspahn, K. M., Zeintl, M., Jäger, T., & Kliegel, M. (2011). Metacognition in prospective memory: Are performance predictions accurate? *Canadian Journal of Experimental Psychology*, 65(1), 19–26. <https://doi.org/10.1037/a0022842>

- Smith, R. E. (2016). Prospective memory: A framework for research on metaintentions. In J. Dunlosky & S. U. K. Tauber (Eds.), *The oxford handbook of metamemory* (pp. 217–244). Oxford University Press.
- Smith, S. J., Souchay, C., & Moulin, C. J. A. (2011). Metamemory and prospective memory in Parkinson's disease. *Neuropsychology*, 25(6), 734–740. <https://doi.org/10.1037/a0025475>
- Somerville, S. C., Wellman, H. M., & Cultice, J. C. (1983). Young children's deliberate reminding. *Journal of Genetic Psychology*, 143(1), 87–96. <https://doi.org/10.1080/00221325.1983.10533537>
- Tauber, S. U. K., & Dunlosky, J. (2016). A brief history of metamemory research and handbook overview. In J. Dunlosky & S. Tauber (Eds.), *The oxford handbook of metamemory* (pp. 217–244). Oxford University Press.
- Uttl, B., & Kibreab, M. (2011). Self-report measures of prospective memory are reliable but not valid. *Canadian Journal of Experimental Psychology*, 65(1), 57. <https://doi.org/10.1037/a0022843>
- Wang, L., Altgassen, M., Liu, W., Xiong, W., Akgün, C., & Kliegel, M. (2011). Prospective memory across adolescence: The effects of age and cue focality. *Developmental Psychology*, 47(1), 226–232. <https://doi.org/10.1037/a0021306>
- Zelazo, P. D., Craik, F. I., & Booth, L. (2004). Executive function across the life span. *Acta Psychologica*, 115(2–3), 167–183. <https://doi.org/10.1016/j.actpsy.2003.12.005>
- Zimmermann, T. D., & Meier, B. (2006). The rise and decline of prospective memory performance across the lifespan. *The Quarterly Journal of Experimental Psychology*, 59(12), 2040–2046. <https://doi.org/10.1080/17470210600917835>
- Zuber, S., & Kliegel, M. (2020). Prospective memory development across the lifespan: An integrative framework. *European Psychologist*, 25(3), 162–173. <https://doi.org/10.1027/1016-9040/a000380>