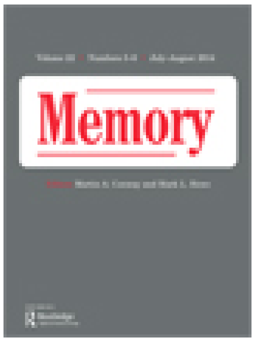

Memory

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/pmem20

Aging and time-based prospective memory in the laboratory: a meta-analysis on age-related differences and possible explanatory factors

G. Laera, F. Borghese, A. Hering, M. Kliegel & G. Mioni

To cite this article: G. Laera, F. Borghese, A. Hering, M. Kliegel & G. Mioni (2023): Aging and time-based prospective memory in the laboratory: a meta-analysis on age-related differences and possible explanatory factors, Memory, DOI: 10.1080/09658211.2023.2191901

To link to this article: https://doi.org/10.1080/09658211.2023.2191901

MEMORY
https://doi.org/10.1080/09658211.2023.2191901

Routledge
Taylor & Francis Group

REVIEW ARTICLE

Aging and time-based prospective memory in the laboratory: a meta-analysis on age-related differences and possible explanatory factors

G. Laera[a,b,c], F. Borghese [a,b,c], A. Hering [a,b,d], M. Kliegel [a,b,c] and G. Mioni [e]

[a]Cognitive Aging Lab (CAL), Faculty of Psychology and Educational Sciences, University of Geneva, Geneva, Switzerland; [b]Centre for the Interdisciplinary Study of Gerontology and Vulnerability, University of Geneva, Geneva, Switzerland; [c]LIVES, Overcoming Vulnerability: Life Course Perspectives, Swiss National Centre of Competence in Research, University of Geneva, Geneva, Switzerland; [d]Department of Developmental Psychology, Tilburg School for Social and Behavioral Sciences, Tilburg University, Tilburg, The Netherlands; [e]Department of General Psychology, University of Padua, Padua, Italy

ABSTRACT

In older adults' everyday life, time-based prospective memory (TBPM) is relevant as health-related intentions are often part of daily activities. Nonetheless, it is still unclear which task-related factors can potentially moderate the magnitude of age-related differences, such as duration of the PM target time (the time-window within which an individual must complete a given TBPM task), the frequency of the TBPM tasks, and the criterion chosen to compute PM accuracy. The present meta-analysis aimed to quantify age-related differences in laboratory TBPM tasks, and to investigate how specific task-related factors potentially moderate the magnitude of age effects. The results showed that age effects consistently emerged among the studies, with older adults showing lower TBPM performance and checking the clock less often than younger adults, especially for shorter intervals (e.g., ≤ 4 min). Furthermore, the results indicated that the duration of the PM target time interacted with the frequency of the PM task, suggesting that learning effects may attenuate the magnitude of age differences in TBPM performance. The results are discussed in terms of potential implications about the possible cognitive processes involved in TBPM and aging, as well as in terms of robustness of the TBPM laboratory paradigm in aging research.

ARTICLE HISTORY
Received 25 November 2022
Accepted 10 March 2023

KEYWORDS
Aging; delayed intentions; time monitoring; task features; duration

Time-based prospective memory (TBPM) is the ability to perform an intended action at a precise moment in the future, such as going to the office at 9 am for a meeting (Einstein et al., 1995), or at a specific time interval, such as taking the pizza out of the oven after 20 min. In the classic laboratory TBPM paradigm (Park et al., 1997), people have to remember to perform a specific action after a certain amount of time (e.g., to press Enter every 3 min) while they are engaged in a background activity (usually referred to as ongoing task: OT); moreover, people are usually free to check a clock on the computer screen by pressing another key. Several studies showed that controlling the clock (i.e., time monitoring behaviour) is positively correlated to TBPM accuracy (Ceci & Bronfenbrenner, 1985; Harris & Wilkins, 1982; Mäntylä et al., 2006; Mioni et al., 2019; Mioni & Stablum, 2014; Vanneste et al., 2016); yet, it is unclear which cognitive processes underlie age differences in time monitoring and TBPM (Varley et al., 2021).

Age-related differences in TBPM

In older adults' everyday life, TBPM is very relevant as health-related intentions are often part of daily activities, such as taking medication regularly, or going to appointments with the doctor (Haas et al., 2020; Hering et al., 2018; Woods et al., 2015); nonetheless, it is not clear how TBPM is affected by aging. Indeed, although most of the studies examining the age-related differences in TBPM showed that younger adults outperform older adults in laboratory TBPM tasks (Einstein et al., 1995; Mioni et al., 2019; Mioni & Stablum, 2014; Park et al., 1997; Vanneste et al., 2016), there is an ongoing debate on how large this difference between younger and older adults indeed is. In fact, meta-analytic evidence on TBPM and aging has been provided by Henry and colleagues (2004), which showed that, in laboratory settings, younger participants outperformed older participants at TBPM tasks; however, this meta-analysis included only six studies that assessed TBPM in the

CONTACT G. Laera Gianvito.Laera@unige.ch Cognitive Aging Lab (CAL), Faculty of Psychology and Educational Sciences, University of Geneva, 28 Boulevard du Pont d'Arve, 1205 Geneva, Switzerland

Supplemental data for this article can be accessed online at https://doi.org/10.1080/09658211.2023.2191901.

laboratory (Henry et al., 2004). In the last 15 years, there have been an increasing number of studies on TBPM, but so far, a systematic quantification of the age-related differences in the laboratory TBPM paradigm is still missing.

Moreover, it is unclear whether time monitoring differs between age groups: although most of the studies found that younger adults checked the clock more often than older adults (e.g., Mioni et al., 2019; Mioni & Stablum, 2014; Vanneste et al., 2016), other authors found the opposite pattern (Mäntylä et al., 2009), or even no differences between age groups (McFarland & Glisky, 2009). The cognitive processes underlying age differences in TBPM are also not well understood, with some authors arguing that age differences are due mainly in time estimation (Labelle et al., 2009; Mioni & Stablum, 2014; Vanneste et al., 2016) and others suggesting rather that attentional processes is the main source of age differences in TBPM (Lecouvey et al., 2017; Varley et al., 2021; Zuber & Kliegel, 2020). A recent study by Varley and colleagues (2021) examined experimentally the impact of attentional and temporal processes on TBPM in younger and older adults. In this study, participants were randomly assigned at three conditions: the "visible" condition, a timer was constantly present; in the "monitored" condition, the timer was only made visible after button press (i.e.,: as in any traditional TBPM task); in the "hidden" condition, access to the timer was not possible, eliminating the need to disengage focus from the OT, but adding the requirement to engage internal time estimation processes. The authors hypothesised that age-related declines in time estimation would lead to worse performance in the hidden condition, while age-related declines in attentional processes would lead to worse performance in the visible and monitored conditions. The study found that age-related impairments in TBPM performance were only present when participants had access to the timer ("visible" and "monitored" conditions), but not in the condition where the timer was not available ("hidden" condition), suggesting that age differences in TBPM accuracy were due to impairments in attentional processes, rather than time estimation abilities (Varley et al., 2021).

The role of specific task-related factors

The cognitive processes responsible for the age effects in time monitoring and TBPM performance could be affected by other task-related factors, which could potentially moderate the magnitude of the age-related effect associated with TBPM (Bastin & Meulemans, 2002; D'Ydewalle et al., 2001; Einstein et al., 1992; McBride et al., 2011; Meier et al., 2006). For example, some authors highlighted the importance of the duration of the PM target time (i.e.,: the time-point indicating that a given action needs to be performed) and the task frequency (i.e.,: how many PM task in a TBPM task block). In the literature, many studies used different durations of the PM target time, ranging from 30 s to 10 min (Bastin & Meulemans, 2002; Gonneaud et al., 2017; Mioni & Stablum, 2014; Vanneste et al., 2016; Waldum & McDaniel, 2016), as well as different paradigms in which the PM task needed to be carried out once (Waldum & McDaniel, 2016) or multiple times (Mioni et al., 2019; Vanneste et al., 2016); however, only few studies investigated the impact of the PM task's duration and frequency on the age-related differences in TBPM (Bastin & Meulemans, 2002; Einstein et al., 1995; McBride et al., 2011; Meier et al., 2006). Yet, it can be very informative to investigate the effect of these task's parameters as they can have direct implications for internal timing as well as executive and memory processes (Bastin & Meulemans, 2002; Block & Zakay, 2006; Conte & McBride, 2018; Gan & Guo, 2019; Guo & Huang, 2019; Lecouvey et al., 2017; McBride et al., 2011; Varley et al., 2021).

For example, Einstein and colleagues (1995) found that younger adults consistently outperformed older adults on PM tasks, regardless of the duration of the PM target time (Einstein et al., 1995); other studies have also found similar age-related differences in performance with different PM target times, ranging from 1 to 6 min (Bastin & Meulemans, 2002; Conte & McBride, 2018; Park et al., 1997), although one study reported that both younger and older adults had lower accuracy when the PM target time was 1 min compared to when it was 2 min (Bastin & Meulemans, 2002). Interestingly, the authors of this study suggested that a PM target time of 1 min required more attentional control processes than a target time of 2 min, leading to lower accuracy for both younger and older adults. The effects of PM task frequency on age-related differences in TBPM are less clear. The first study that systematically investigated the effect of PM task frequency on aging in TBPM found no significant differences in accuracy and time monitoring between 6- and 12-event PM tasks (Park et al., 1997). However, a recent study has shown that repeating the same PM task with the same target time can lead to learning effects in younger adults (Gan & Guo, 2019); nonetheless, it is still unclear which processes are responsible for such learning effect, as it could be due either to better distribution of the attentional resources between OT and PM task, or to an improvement of time estimation abilities involved in the monitoring of the PM target time. Moreover, it is currently unknown whether and how the frequency of the TBPM task has similar effects on older adults' performance too. Another methodological aspect that can affect age-related TBPM differences is the criterion chosen to compute PM accuracy. Typically, PM accuracy is measured as a binary score based on whether participants completed the task within a specified interval around the PM target time. Some studies have used lenient criteria with larger intervals, such as 15% of the whole PM target time interval (e.g.,: Mioni & Stablum, 2014), while others used stricter criterion with smaller interval, such as 10% of the whole PM target time (e.g.,: Vanneste et al., 2016). There have been a few studies that have explored the impact of different criterion on age-related differences in TBPM. One study found that older adults had more difficulty with TBPM regardless of whether a larger or smaller interval was used for accuracy (Park et al., 1997). A more recent study contrasted these findings, showing that a larger interval improved TBPM performance for older adults but not for younger adults (Yang et al., 2013). Apart from these few findings, there is currently no systematic investigation on the effect of the criterion of the PM accuracy on age-related differences in TBPM; yet this aspect can have methodological and analytical implications on how PM tasks are designed and on how PM accuracy is scored.

The present meta-analysis

Overall, the empirical evidence suggested that there are age-related differences in TBPM performance and time monitoring, as measured by laboratory tasks, but it is currently unknown how large the age effect is in time monitoring. Time monitoring has only been assessed in laboratory settings so far, whereas studies measuring time monitoring in naturalistic settings are still missing. Hence, this meta-analysis aimed to: (1) quantify age-related differences in TBPM and time monitoring assessed in the laboratory setting, (2) determine if there's a relationship between age effects in TBPM performance and time monitoring, and (3) measure how specific task-related factors (i.e., the duration of the PM target time, the frequency of the PM task, and the interval criterion for correct PM responses) affect age-related differences in TBPM performance and time monitoring. This meta-analysis is the first to quantify the relationship between time monitoring and TBPM performance and explore meta-analytically the potential role of PM task-related factors. It provides a conceptual understanding of the cognitive processes behind the age effect in time monitoring and TBPM performance and offers a methodological framework for future aging research.

Method

Search strategy

This systematic review follows the guidelines of Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA, Moher et al., 2015). We searched the articles using PsycInfo, PubMed, and Web of Science databases, from the earliest available date to the end of October 2022. We used the following descriptive verbal expressions: "prospective memory", "time-based", combined with "aging" or "ageing", and "monitoring". the meta-analysis has been registered before data coding (Open Science Framework pre-registration DOI: https://doi.org/10.17605/OSF.IO/9JW6X).

Eligibility criteria

For an outline of the search and screening steps, see the PRISMA flow chart (Figure 1). Included studies were required to meet the following criteria: (1) had experiments involving young and older adults[1], (2) used laboratory TBPM tasks, (3) tested PM performance (as sum or proportional accuracy) or time monitoring (as total clock checks), or both, as dependent variable(s), (4) were published in a peer-reviewed, English language journals. The following exclusion criteria were also applied: (a) studies, or single experiments within studies, that included any experimental manipulation of the OT and/or TBPM task, as they could affect OT and PM performance. From these studies, we kept only the data from the TBPM tasks that were not subjected to any experimental manipulation; this choice had been made to ensure that the studies were comparable without the risk of confounding effects related to different experimental manipulations (van Rhee et al., 2015), (b) studies that included clinical samples (e.g.,: Costa et al., 2015; Mioni et al., 2017; Smith-Spark et al., 2017), (c) studies that involved drug interventions and/or ingestion of substances (e.g.,: Behrendt et al., 2015; Costa et al., 2008; Platt et al., 2016), or that manipulated other factors including sleep (e.g.,: Bezdicek et al., 2018; Esposito et al., 2015), (d) experiments that included children, adolescents, and middle-age adults (Nigro et al., 2002; Zöllig et al., 2010), although we kept some studies still eligible, but only when it was possible to extract the data of both younger and older adults leaving out the middle-age group (Bozdemir & Cinan, 2021; Einstein et al., 1995: Experiment 3; Gonneaud et al., 2014; Mäntylä et al., 2009; Zuber et al., 2021). Finally, two studies from the same research group reported different neural measures on the same behavioural results (Morand et al., 2021, 2022), which are therefore redundant; thus, we decided to keep one of them (Morand et al., 2021).

Study selection

In total, 93 studies were screened; 36 were excluded because they did report only samples of younger (e.g., Huang et al., 2014; Khan et al., 2008) or older adults (e.g., Schnitzspahn & Kliegel, 2009; Sullivan et al., 2018), with no age comparisons, so it was not possible to calculate the (age) effect size for these studies. Hence, 56 studies were assessed; seven were excluded because they reported only samples of younger adults (Cona et al., 2012; Cruz et al., 2017; Gonneaud et al., 2014; Haines et al., 2020; Oksanen et al., 2014; Okuda et al., 2007; Tracy et al., 2000). We excluded four studies as they reported only naturalistic assessment of TBPM (Kvavilashvili & Fisher, 2007; Maylor, 1990; McBride et al., 2013; Rendell & Thomson, 1993), as well as single experiments (within 3 studies) that included only naturalistic assessment of TBPM (Aberle et al., 2010: Experiment 2; Niedźwieńska & Barzykowski, 2012: Experiment 2; Rendell & Thomson, 1999: Experiment 1 and 2). One further experiment was excluded as it reported only EBPM task (Einstein et al., 1995: Experiment 2). Concerning the studies with

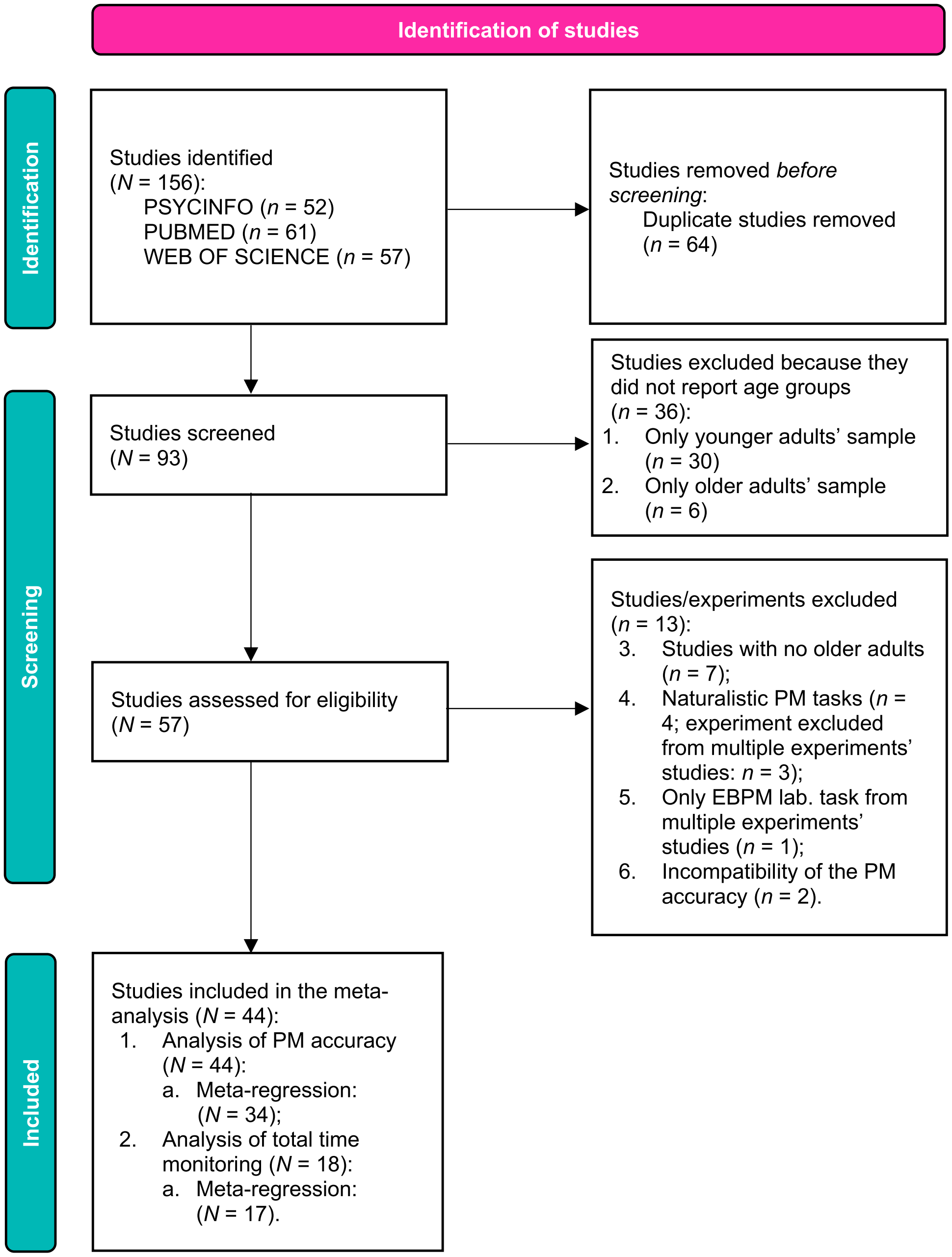

Figure 1. PRISMA flow diagram. *Note*. PRISMA flow diagram of the literature review process for the meta-analysis (up to October 2022); PM: prospective memory.

Virtual Week (e.g.,: Henry et al., 2012; Mioni et al., 2019; Rendell & Craik, 2000), we kept only the stop-clock sub-task, as it comprised comparable TBPM tasks (e.g.,: "check lung capacity at 2 and 4 min on the clock"), whereas the other sub-tasks, often referred as "regular" and "irregular", are based on a fictious time-week tasks (e.g.,: "remember to call your partner at 4 pm to collect photocopies") and, as such, cannot be compared with the outcomes provided by the traditional TBPM paradigm (Einstein et al., 1995).

Many studies did not report the measures necessary for the meta-analysis (for more information on the outcome measures, see the following section "statistical analyses"); for example, some studies reported a task in which monitoring was (or could have been) measured, but the behavioural data on monitoring were not reported (e.g.,: Lecouvey et al., 2017; Varley et al., 2021). Therefore, we used two alternative sources of information: figure digitalisation and author's contact. The figure digitalization has been applied to 16 articles that reported data in the figures, but not in the tables (Altgassen et al., 2010; Bastin & Meulemans, 2002; D'Ydewalle et al., 2001; Einstein et al., 1995; Gonneaud et al., 2011; Henry et al., 2012; Lecouvey et al., 2017; Martin & Schumann-Hengsteler, 2001; Mäntylä et al., 2009; Mioni et al., 2015; Mioni & Stablum, 2014; Rendell et al., 2011; Rendell & Thomson, 1999; Schnitzspahn et al., 2011, 2014; Vanneste et al., 2016); therefore, we extracted the data from the figure by digitalization using the software DigitizeIt (version 2.5); such software has been proved to be reliable for meta-analytical studies in psychology as well as in other disciplines, showing that the values obtained using the software do not differ from the real data (Rakap et al., 2016; Schild & Voracek, 2013; Wojtyniak et al., 2020).

Seven papers did not report data of TBPM accuracy and/or time monitoring neither in table nor depicted in figures (Aberle et al., 2010; Haines et al., 2020; Rendell et al., 2011; Rendell & Craik, 2000; Rendell & Thomson, 1999; Varley et al., 2021; Zuber et al., 2021). Thus, we contacted the corresponding authors of these articles; all authors replied to the email. However, only in three cases data were available or compatible with the goals of the present meta-analysis (Aberle et al., 2010; Varley et al., 2021; Zuber et al., 2021). Finally, we excluded two studies because they reported incompatible PM accuracy measures, such as time completion of the TBPM task (Waldum & McDaniel, 2016) or deviation of the subjective PM response from objective PM target time (Patton & Meit, 1993). Such choice was made because, in a meta-analysis, it is important to use a consistent measure of the construct of interest across studies to calculate effect sizes (Harrer et al., 2021; Hedges, 1981); the time of task completion and the PM response's deviation are purely temporal measures that (1) do not have a maximum score (as the PM accuracy score), and (2) it is not standardised across different PM target time durations (e.g., the minutes of completion can be problematic as they depend from the duration of the PM target time). The consistent use of a single measure of the construct across studies is essential in meta-analysis to calculate effect sizes; therefore, given that PM accuracy was way more common across studies than the other measures, it was better to use only studies that have PM accuracy to ensure consistency in effect size computation, and to facilitate interpretation of the results. For the analysis on age effects in TBPM performance, 52 unique effect sizes were included, nested in 44 studies; for the analysis on age effects in time monitoring, 20 unique effect sizes were included, nested in 18 studies. The selection from the search results has been executed by the first and the second author in advance; nonetheless, if a full-text review of an article did not result in a clear verdict, the decision on in- or exclusion was made by mutual agreement of all the authors.

Statistical analyses

All analyses were carried out in R (version 4.2.1) (R Core Team, 2022) using the packages metafor (Viechtbauer, 2010), meta (Balduzzi et al., 2019), and metaSEM (Cheung, 2015). Data, metadata, and R-code are available in the Open Science Framework (https://doi.org/10.17605/OSF.IO/EPBNK). All analyses were carried out using the standardised mean difference (Hedge's *g*) of TBPM performance (i.e.,: proportional accuracy, sum scores, z-values) and total time monitoring (i.e., number of clock checks) as outcome measures (formulas are reported in the Supplementary material). Among the studies included in the meta-analysis, few of them reported multiple effect sizes as a function of PM duration (Bastin & Meulemans, 2002), PM task frequency (Park et al., 1997), or criterion chosen for the PM accuracy (Yang et al., 2013). Even though these are not the majority of the studies, it is reasonable to assume that some kind of dependency is introduced within the reported data; such dependency was taken into account by integrating a third layer into the structure of the meta-analytic model, resulting in a three-level meta-analysis (Assink & Wibbelink, 2016; Cheung, 2014; Van den Noortgate et al., 2015) with participants (level 1) nested in the individual effect sizes (level 2), which were, in turn, part of a number of larger units, the studies (level 3). Wald-type tests was used to calculate the confidence interval around the pooled effects; the amount of heterogeneity (i.e., τ^2), was estimated using the restricted maximum-likelihood estimator (Viechtbauer, 2005) and it was de-composed into two partitions to account for within- and between-studies sources of heterogeneity simultaneously (Cheung, 2014). In addition to the estimate of τ^2, the *Q*-test for heterogeneity (Cochran, 1954) and the I^2 statistic (Higgins & Thompson, 2002) are reported. In case any amount of heterogeneity is detected (i.e., $\tau^2 > 0$, regardless of the results of the *Q*-test), a prediction interval for the true outcomes is also provided (Riley et al., 2011).

The meta-analysis was carried out in three steps. In the first step, two random-effect models on age effect were carried out separately on time monitoring (17 studies) and TBPM performance (43 studies); the aim of this first analytic step was to pool effect sizes and to quantify age effects and studies heterogeneity; at this level, publication bias analyses using Egger regression (Borenstein et al., 2009, Chapter 30), and outliers and sensitivity analyses were performed too (Viechtbauer & Cheung, 2010). In the second step, a multi-variate model was carried out jointly on time monitoring and TBPM performance; the aim of this model was to investigate the relationship between the age effect in time monitoring and TBPM performance. In the third and final step, we carried out the same multi-variate model as in step 2, but this time we included task-related features as predictors (i.e., duration and frequency of the PM target time, and standardised interval criterion for correct PM responses); the aim of this model was to investigate the relationship between the age effect time in monitoring and TBPM performance, as well as the effect of task-related features on age effects and studies heterogeneity. The duration of the PM target time was stored in a variable that comprised the duration of the PM target time in minutes; the frequency of the TBPM task was the number of times the TBPM task was performed within each task block. Interval for correct PM responses was standardized as ratio between the whole interval and the PM target time in seconds (e.g.,: if a study reported as correct answer any response falling within ± 6 s for a 2-minute TBPM task, we have computed the value as follows: 6 * 2 / 120 = .10).

Results

In Tables 1 and 2, we report the sample characteristics of the eligible studies included in the meta-analysis, predictors (i.e.,: duration of the PM target time, frequency of the PM tasks, and criterion for accuracy), as well as effects of age for TBPM accuracy (Table 1) and time monitoring (Table 2). Overall, the mean age for the samples of younger adults was 23 years (18–41), whereas the mean age for the samples of older adults was 69 (51–97).

Age effects & sensitivity analyses

For the random-effect model of TBPM performance, a total of $k = 52$ unique effect sizes were included in the analysis, which were nested into 44 unique studies. The observed standardized mean differences ranged from −0.324–2.675; most estimates were positive (96%; i.e.,: younger adults performed better than older adults at the TBPM task). The estimated average standardized mean difference based on the random-effects model was $g = 1.064$ (95% CI: 0.904–1.224); the average outcome differed significantly from zero ($z = 13.41$, $p < 0.001$). The forest plot shows the observed outcomes and the estimate based on the random-effects model (Figure 2(A)). According to the Q-test, the true outcomes appear to be heterogeneous ($Q(50) = 240.531$, $p < 0.001$); the source was related almost exclusively to the between-studies heterogeneity ($\tau^2_{\text{between-studies}} = 0.188$, $I^2_{\text{between-studies}} = 66.39\%$) rather than to within-studies heterogeneity ($\tau^2_{\text{within-studies}} = 0.042$, $I^2_{\text{within-studies}} = 14.90\%$). A 95% prediction interval for the true outcomes is given by 0.08–2.04. Hence, even though there may be some heterogeneity, the true outcomes of the studies are generally in the same direction as the estimated average outcome. However, the forest plot suggested that there are some extreme studies' values that contributed substantially to the heterogeneity; thus, a parallel analysis of the outliers is needed.

For the random-effect model of time monitoring, a total of $k = 20$ unique effect sizes were included in the analysis, which were nested into 18 unique studies. The observed standardized mean differences ranged from −0.869–1.974; most estimates were positive (94%; i.e.,: younger adults checked the clock more often than older adults). The estimated average standardized mean difference based on the random-effects model was $g = 0.587$ (95% CI: 0.333–0.842); the average outcome differed significantly from zero ($z = 4.85$, $p < 0.001$). The forest plot shows the observed outcomes and the estimate based on the random-effects model (Figure 2(B)). According to the Q-test, the true outcomes appear to be heterogeneous ($Q(19) = 76.068$, $p < 0.001$); and such heterogeneity was related exclusively to the between-studies differences ($\tau^2_{\text{between-studies}} = 0.213$, $I^2_{\text{between-studies}} = 78.86\%$) rather than to differences within the studies ($\tau^2_{\text{within-studies}} < 0.001$, $I^2_{\text{within-studies}} < 1\%$). A 95% prediction interval for the true outcomes is given by −0.41–1.59. Hence, although the average outcome is estimated to be positive, in some studies the true outcome may in fact be negative.

Egger statistic was not significant for TBPM accuracy ($z = 1.92$, $p = .055$) and time monitoring ($z = 1.43$, $p = .153$), indicating no publication bias. However, 11 studies were identified as outliers for TBPM performance (Bastin & Meulemans, 2002; D'Ydewalle et al., 1999; Henry et al., 2012, 2020; Mäntylä et al., 2009; McFarland & Glisky, 2009; Mioni et al., 2019; Niedźwieńska & Barzykowski, 2012; Rendell et al., 2011; Shum et al., 2013; Yang et al., 2013), while 2 studies were identified as outliers for time monitoring (Mäntylä et al., 2009; Mioni et al., 2019); thus, sensitivity analyses were carried out to investigate the contribution of outliers on either the combined effect size and/or on studies heterogeneity. To achieve this purpose, two separate three-level random-effect models were carried out excluding outliers, one for age effects in TBPM performance, and one for the age effects in time monitoring (Harrer et al., 2021). The model on the TBPM performance without outliers showed that the I^2 heterogeneity shrank considerably when outliers were excluded (from $I^2 = 66.39\%$ to $I^2 = 46.10\%$; $Q(41) = 69.601$, $\tau^2 < 0.01$, $p = 0.002$); the pooled age effect ($g = 1.055$) was very close to the age effect in the model with outliers ($g = 1.064$). Concerning the time monitoring, the analysis without outliers showed that the I^2 heterogeneity shrank considerably when the two outliers were excluded (from $I^2 = 78.86\%$ to 12.81%; $Q(17) = 19.03$, $\tau^2 < 0.01$, $p = 0.329$). The pooled age effect ($g = 0.573$) was not so different from the model with outliers ($g = 0.587$). In summary, it is possible to argue that removing outliers did not change the average age effect size of TBPM performance and time monitoring, but it affected the heterogeneity in the data substantially (see Supplementary material for detailed information about the publication bias, outliers and sensitivity analyses).

Association between age effects in time-based prospective memory performance & time monitoring

Multivariate meta-analytic approaches can quantify the relationship between age effects in time monitoring and TBPM performance by estimating the effect sizes for both outcomes jointly in one model; moreover, such approach can be used to determine if studies with a high effect size on one outcome also have higher effect sizes on the other outcome. To achieve this, the multivariate meta-analysis was carried out taking into account the correlation between the age effects (see Supplementary material for more detailed information); however, among the studies that were included in this meta-

Table 1. TBPM performance, sample and predictors' characteristics of the eligible studies.

	Participants								Predictors			Age effect in TBPM accuracy			
	Younger adults				Older adults							Confidence Intervals			
Study	*N*	Women	*Age M* (range)	*S.D.*	*N*	Women	*Age M* (range)	*S.D.*	Duration of PM target time	Frequency of TBPM task	Standardised criterion used for TBPM accuracy	Hedges' *g*	Lower	Upper	Weight
Aberle et al. (2010)	20	16	26.25	8.27	40	32	63.26	5.09		10		0.34	−0.28	0.97	1.45%
Altgassen et al. (2010)	40	21	24.73	3.5	40	19	68.7	4.5	2	4	0.05	1.70	1.19	2.21	1.61%
Bastin and Meulemans (2002)	48	24	23.17 (20–30)	2.55	48	24	64.44 (60–70)	3.17	2	6	0.05	1.05	0.62	1.47	3.03%
	48	24	23.17 (20–30)	2.55	48	24	64.44 (60–70)	3.17	1	12	0.1	1.69	1.22	2.15	3.13%
Bozdemir and Cinan (2021)	41	26	22.29 (19–30)	2.45	20	10	65.69 (57–75)	4.37				1.35	0.86	1.83	1.65%
Cona et al. (2012)	15	10	23.81 (21–28)	2.01	47	30	67.77 (60–67)	5.41	3	5	0.1	0.75	0.04	1.46	1.33%
Costermans and Desmette (1999)	20	8	22.35	2.52	20		66.15	3.17	7	6	0.02	0.74	0.09	1.38	1.43%
D'Ydewalle et al. (1999)															
(OT: quiz)	60	36	19.35 (18–22)		59	37	62.93 (55–84)		4	6	0.09	0.55	0.18	0.92	3.65%
(OT: face recognition)	60	36	19.35 (18–22)		48	20	62.93 (55–84)		4	6	0.09	0.49	0.12	0.85	3.65%
D'Ydewalle et al. (2001)	48	30	20 (18–25)		23	12	69 (60–86)		2	5	0.25	0.60	0.18	1.01	1.75%
Einstein et al. (1995) (Experiment 1)	12		(18–21)		12		66 (61–78)		10	1	0.1	0.94	0.09	1.78	1.50%
	12		(18–21)		12		66 (61–78)		10	1	0.05	0.69	−0.16	1.51	1.47%
Einstein et al. (1995) (Experiment 3)	36		20.2 (18–22)		26		66.3 (61–76)		5	6	0.25	0.84	0.31	1.36	1.59%
Gonneaud et al. (2011)	29	14	24.3 (18–35)	4.5	23	13	68.2 (60–84)	6.7	3	8	0.06	1.34	0.73	1.94	1.48%
Gonneaud et al. (2017)	20	9	25.15 (18–35)	5.14	18	12	62.1 (51–76)	2.7	0.5		0.23	1.63	0.91	2.34	1.33%
Guimond et al. (2006)	35	16	22 (15–29)	5.2	38	19	68 (60–85)	6.4		2		1.69	1.16	2.23	1.58%
Haas et al. (2022)	53	41	23.29 (19–32)	2.27	38	25	68.2 (60–81)	5.77	5	5	0.1	0.86	0.43	1.30	1.72%
Haines et al. (2020) (Experiment 1)	40	30	24.1 (19–30)	3.6	31	21	71.6 (65–86)	4.9	2	4	0.08	1.23	0.75	1.70	1.66%
Haines et al. (2020) (Experiment 3)	23	14	22.9 (18–34)	4.1	20	13	70.6 (60–83)	5.5	2	4	0.08	1.55	0.93	2.16	1.47%
Henry et al. (2020)	125	89	22.9	3.45	41	28	73.8	5.57		8		1.60	1.31	1.88	1.92%

(*Continued*)

Table 1. Continued.

	Participants								Predictors			Age effect in TBPM accuracy			
	Younger adults				Older adults							*Confidence Intervals*			
Study	*N*	Women	*Age M* (range)	*S.D.*	*N*	Women	*Age M* (range)	*S.D.*	Duration of PM target time	Frequency of TBPM task	Standardised criterion used for TBPM accuracy	Hedges' *g*	Lower	Upper	Weight
			(18–30)				(65–85)								
Henry et al. (2012)	48		20.4	2.9	30		73.3	5.48		4		2.30	1.78	2.82	1.60%
			(18–27)				(65–84)								
Hering et al. (2014)	30	6	20.87	4.15	30	12	67.7	4.72	3	2	0.33	0.80	0.27	1.32	1.59%
Ihle et al. (2014)	33		20.8	2.1	29		65.2	4.9	1	10	0.17	1.51	0.95	2.08	1.54%
			(18–26)				(54–74)								
Jäger and Kliegel (2008)	30	16	24	3	32	18	67.1	7.2	2	5	0.04	0.94	0.42	1.47	1.59%
			(18–30)				(58–91)								
Lecouvey et al. (2017)	35	12	24.8	5.7	40	30	65.28	7.49	4	3		1.69	1.12	2.26	1.53%
Logie et al. (2004)	40	19	21.5	2.4	40	24	65.6	6.7	3	5		0.62	0.18	1.07	1.71%
			(17–27)				(54–78)								
Mäntylä et al. (2009)	39	21	23.3	2.4	40	23	70.2	6.3	5	7	0.07	0.36	−0.09	0.80	1.71%
			(20–30)				(64–81)								
Martin et al. (2003)	40	21	24.8	2	20	8	69.3	5.6	2	4	0.08	0.56	0.11	1.00	1.71%
			(22–31)				(60–80)								
Martin & Schumann-Hengsteler (2001)	90	75	24	3.77	38	25	69	5.49	2	6	1	1.03	0.70	1.35	1.86%
Maylor et al. (2002)	30		25.40	4.96	30		67.27	4.24	3	5	0.02	1.10	0.56	1.65	1.57%
McFarland and Glisky (2009)	32				32		74.88 (65 +)	5.2	5	8	0.1	−0.32	−0.82	0.17	1.64%
Mioni et al. (2019)	30	26	22.6	4.23	30	23	74.33	5.54	2	8	0.17	2.68	1.98	3.37	1.35%
Mioni et al. (2015)	19	9	29.95	1.22	39	22	73.75	5.22	2	10	0.17	1.22	0.53	1.90	1.37%
			(22–27)				(65–84)								
Mioni and Stablum (2014)	76	45	23.11	2.58	76	44	70.05	7.47	5	4	0.07	1.14	0.80	1.49	1.84%
			(19–34)												
Morand et al. (2021)	22	12	25.4	5.19	22	10	62.5	6.05	0.5		0.23	1.19	0.55	1.82	1.44%
			(18–35)				(51–76)								
Niedźwieńska and Barzykowski (2012)	63	42	21.56	1.94	29	14	68.4	3.16	10	4		0.27	−0.11	0.65	1.80%
			(19–27)				(64–74)								
Park et al. (1997)	56	39	19.59	2.07	55	28	69.8	5.84	2	6	0.03	0.79	0.40	1.17	3.95%
	56	39	19.59	2.07	55	28	69.8	5.84	2	6	0.06	1.01	0.62	1.41	4.08%
	56	39	19.59	2.07	55	28	69.8	5.84	2	6	0.16	1.45	1.03	1.87	4.15%
Pupillo et al. (2021)	109		19.94		103		70.79		1	16	0.08	0.97	0.68	1.25	1.92%
			(18–27)				(59–85)								
Rendell et al. (2011)	30		21.9	3.28	20	16	75	5.72		12		1.88	1.27	2.48	1.48%
Rendell and Craik (2000)	20	16	21.3		47		73.34		2.5	14	0.13	1.05	0.48	1.61	1.53%
			(19–24)				(61–84)								
Rendell and Thomson (1999)	126		(18–28)		125	87	(60–80)		7	1	0.05	0.93	0.70	1.16	1.97%
Schnitzspahn et al. (2020)	31	19	23.71	3.07	67	70	67.09	4.66	10	2	0.03	0.62	0.06	1.18	1.55%
			(20–29)				(60–70)								
Schnitzspahn et al. (2014)	64		19.11		57		69.79		1	4	0.17	1.00	0.63	1.38	1.80%
			(18–25)				(59–84)								
Schnitzspahn et al. (2011)	20	16	21.5	2.26	59	37	68.55	4.66	2	6	0.1	0.83	0.19	1.48	1.42%
			(18–25)				(61–79)								

Shum et al. (2013)	79	65	21.44 (18–33)	4.53	50	23	68.23 (60–75)	4.13	5		0.07	0.28	−0.04	0.61	1.86%
Vanneste et al. (2016)	40	19	22.7	1.74	38	18	69.15	5.99	1	10	0.1	1.54	1.03	2.04	1.63%
Varley et al. (2021)	53		19.32 (17–29)	2.11	40		71.2 (60–97)	7.5	1	3	1	1.05	0.61	1.49	1.72%
Yang et al. (2013)	25		21.92 (20–24)	0.95	50	23	71.31 (60–80)	3.82	1	5	0.17	1.83	1.27	2.39	2.44%
	25		21.92 (20–24)	0.95	199		71.31 (60–80)	3.82	1	5	1	1.61	1.07	2.16	2.48%
Zuber et al. (2021)	86	67	28.26 (20–40)	6.06	47	0	67.81 (60–86)	7.08	1	6	0.17	0.78	0.41	1.15	1.81%
Combined effect size												**1.06**	**0.90**	**1.22**	

Note. Sample characteristics of the eligible studies included in the meta-analysis, and predictors variables (i.e.,: duration – in minutes – of the PM target time, frequency of the prospective memory tasks, criterion for accuracy), as well as the age effect (Hedges' *g*) on time-based prospective memory accuracy. The list of studies is sorted alphabetically by the name of the first author. Bastin and Meulemans (2002), Einstein et al. (1995, Experiment 1), Park et al. (1997), and Yang et al. (2013) reported distinct accuracy measures on the same sample computed using different intervals for accuracy; D'Ydewalle et al. (1999) reported two measures of time-based prospective memory accuracy: the first one was computed while people performed a quiz as ongoing task, and the second one was computed while people performed a face recognition test as ongoing task. TBPM: time-based prospective memory; OT: ongoing task.

analysis, only 3 reported correlations between time monitoring and TBPM performance, separately for younger (r = .42, .55, and .56) and older adults (r = .51, .69, and .71). Therefore, we have calculated the correlation between age effects in the two outcomes by transforming the effect sizes into Pearson's r coefficients (detailed formulas and procedures are reported in the Supplementary material). According to the multi-variate analysis, the age effects were $g_{\text{TBPM perf.}} = 1.064$ and $g_{\text{time monit.}} = 0.565$; both effect sizes were statistically significant ($p > 0.001$). According to the Q-test, the true outcomes appear to be heterogeneous ($Q(69) = 309.406$, $p < 0.001$), especially for TBPM performance ($\tau^2 = 0.214$, $p < 0.001$), and less for time monitoring ($\tau^2 = 0.164$, $p = 0.117$); moreover, the heterogeneity introduced by the relationship between the two outcomes was significant ($\tau^2 = 0.126$, $p = 0.071$). The values of I^2 indicated high between-study heterogeneity in both outcomes ($I^2_{\text{TBPM perf.}} = 80.17\%$; $I^2_{\text{time monit.}} = 74.16\%$). The correlation between age effects on TBPM performance and time monitoring is $r = 0.67$, suggesting that there was a positive association between the age effect on time monitoring and its effect on TBPM performance (Figure 3); in other words, studies that found higher age effects in time monitoring seem to find higher age effects in TBPM performance too.

Effect of predictors

The multi-variate meta-regression model was carried out with three task-related features as continuous predictors: (1) the duration of the PM target time (i.e.,: the delay of the PM cue in minutes), (2) the PM task frequency (i.e.,: how many PM task in a TBPM task block), and (3) arbitrary criterion chosen to compute PM accuracy. The aims of this model were (1) to investigate whether these predictors were linearly associated with the (age) effect size in both TBPM performance and time monitoring, and (2) to establish whether predictors accounted for (some of) the between-studies heterogeneity introduced by the presence of the outliers, which were included in this model. According to the multi-variate analysis, the age effect sizes when predictors were set to their means were $g_{\text{TBPM perf.}} = 0.883$ and $g_{\text{time monit.}} = 0.396$; both effect sizes were statistically significant ($p > 0.001$). According to the Q-test, the true outcomes appear to be heterogeneous ($Q(57) = 202.933$, $p < 0.001$); however, such heterogeneity was explained by between-studies variance introduced from the variance in the TBPM performance ($\tau^2 = 0.075$, $p = 0.038$), but not from the variance in time monitoring ($\tau^2 = 0.067$, $p = 0.246$), as well as from the variance in the relationship between the two outcomes ($\tau^2 = 0.043$, $p = 0.329$). Indeed, predictors explained the most of the variance of the age effect in TBPM performance ($r^2 = 64.79\%$); specifically, the model indicated that longer durations of target time were associated with a reduction of the age effect in TBPM performance ($\beta = -0.15$; $p < 0.001$; Figure 4(A), upper left panel); similarly, higher PM task

Table 2. Time monitoring, sample and predictors' characteristics of the eligible studies.

	Younger adults				Older adults						*Confidence Intervals*			
Study	*N*	Female	*Age* *M* (range)	*S.D.*	*N*	Female	*M* (range)	*S.D.*	Duration of the TBPM target time	Frequency of TBPM task	Hedges' *g*	Lower	Upper	Weight
Altgassen et al. (2010)	40	21	24.73	3.5	40	19	68.7	4.5	2	4	0.75	0.30	1.21	4.19%
Bastin and Meulemans (2002)	48	24	23.17 (20–30)	2.55	48	24	64.44 (60–70)	3.17	2	6	0.45	0.04	0.85	14.19%
	48	24	23.17 (20–30)	2.55	48	24	64.44 (60–70)	3.17	1	12	0.45	0.04	0.85	14.19%
Cona et al. (2012)	15	10	23.81 (21–28)	2.01	47	30	67.77 (60–67)	5.41	3	5	0.91	0.19	1.63	3.22%
Costermans and Desmette (1999)	20	8	22.35	2.52	20		66.15	3.17	7	6	0.14	−0.48	0.76	3.58%
Einstein et al. (1995) (Experiment 1)	12		(18–21)		12		66 (61–78)		10	2	0.76	−0.07	1.59	2.85%
Einstein et al. (1995) (Experiment 3)	36		20.2 (18–22)		26		66.3 (61–76)		5	6	0.49	−0.02	1.01	3.99%
Gonneaud et al. (2011)	29	14	24.3 (18–35)	4.5	23	13	68.2 (60–84)	6.7	3	8	0.38	−0.17	0.93	3.83%
Hering et al. (2014)	30	6	20.87	4.15	30	12	67.7	4.72	3	2	0.79	0.26	1.32	3.93%
Ihle et al. (2014)	33		20.8 (18–26)	2.1	29		65.2 (54–74)	4.9	1	10	1.17	0.62	1.71	3.87%
Logie et al. (2004)	40	19	21.5 (17–27)	2.4	40	24	65.6 (54–78)	6.7	3	5	1.04	0.57	1.51	4.14%
Mäntylä et al. (2009)	39	21	23.3 (20–30)	2.4	40	23	70.2 (64–81)	6.3	5	7	−0.87	−1.33	−0.41	4.18%
Maylor et al. (2002)	30		25.40	4.96	30		67.27	4.24	3	5	0.73	0.21	1.25	3.94%
McFarland and Glisky (2009)	32				32		74.88 (65+)	5.2	5	8	0.26	−0.23	0.75	4.06%
Mioni and Stablum (2014)	76	45	23.11 (19–34)	2.58	76	44	70.05	7.47	5	4	0.33	0.01	0.65	4.67%
Mioni et al. (2019)	30	26	22.6	4.23	30	23	74.33	5.54	2	8	1.97	1.36	2.59	3.59%
Schnitzspahn et al. (2014)	64		19.11 (18–25)		57		69.79 (59–84)		1	4	0.56	0.19	0.92	4.53%
Vanneste et al. (2016)	40	19	22.7	1.74	38	18	69.15	5.99	1	10	0.59	0.14	1.05	4.19%
Varley et al. (2021)	53		19.32 (17–29)	2.11	40		71.2 (60–97)	7.5	1	3	0.66	0.24	1.08	4.32%
Zuber et al. (2021)	86	67	28.26	6.06	47	0	67.81	7.08	1	6	0.41	0.05	0.77	4.53%
Combined effect size											**0.59**	**0.33**	**0.84**	

Note. Sample characteristics of the eligible studies included in the meta-analysis, and predictors variables (i.e.,: duration – in minutes – of the PM target time, and frequency of the prospective memory tasks), as well as age effect (Hedges' *g*) on time monitoring (as total clock checks). The list of studies is sorted alphabetically by the name of the first author. Bastin and Meulemans (2002) reported the same values of monitoring averaged for two different frequency of the prospective memory task (6 vs. 12), and different durations of the target time (1-minute vs. 2-minutes); hence, we decided to duplicate these values and assign each of them for each of the two values of the respective predictors (i.e., duration of the target time, frequency of the prospective memory task, and intervals for accuracy), thus measuring the moderating contribution of the predictors on the age effect in prospective memory performance (see Results for more information); TBPM: time-based prospective memory.

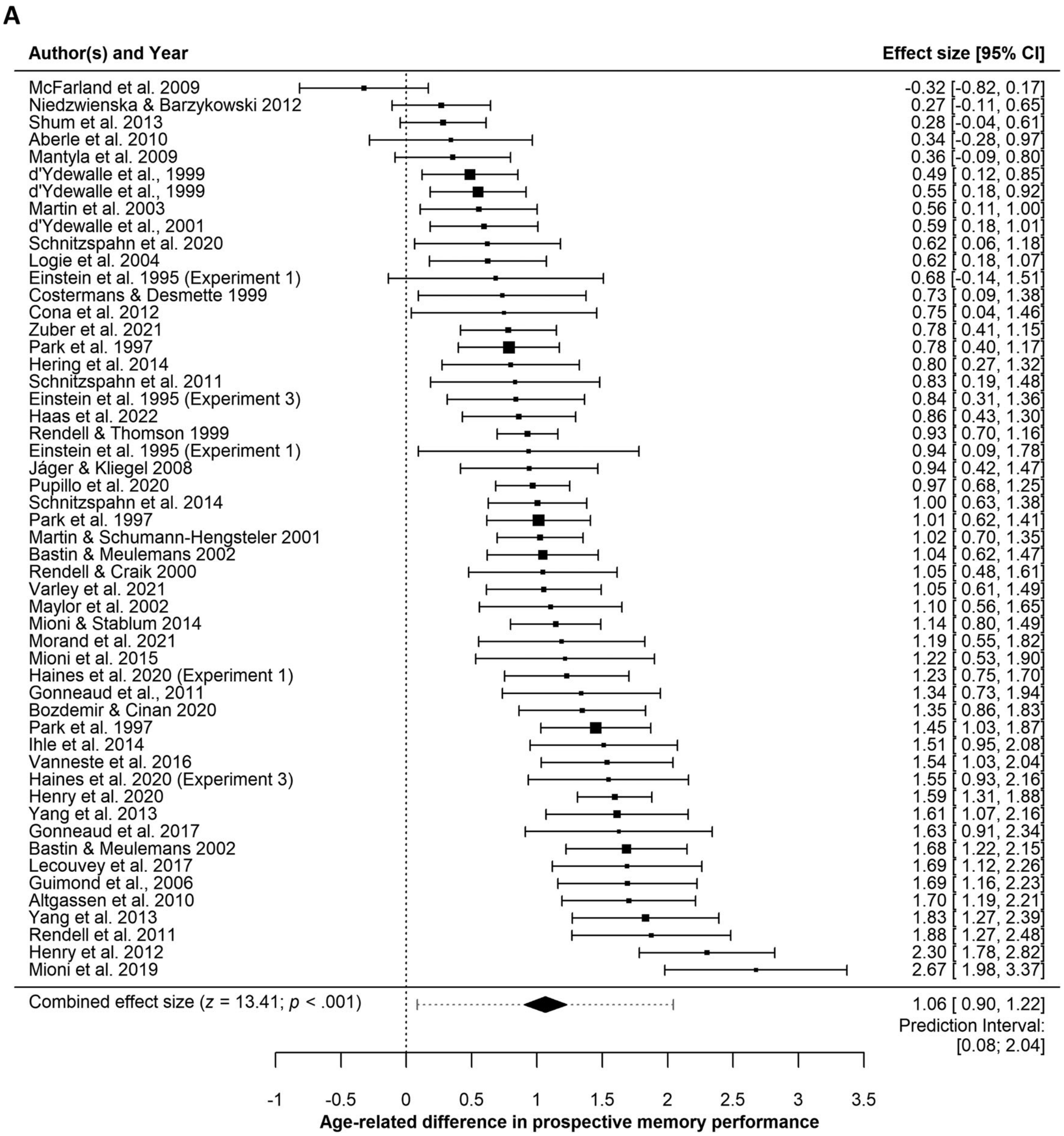

Figure 2. Age effects in time-based prospective memory performance and time monitoring. *Note*. Quantitative summary illustrating the combined effect sizes (Hedges' *g*) of age differences in time-based prospective memory performance (2A), and time monitoring (2B). Concerning time-based prospective memory performance (2A), points to the right of zero indicated negative effect of age (i.e.,: younger adults performing better than the older adults); concerning time monitoring (2B), points to the right of zero indicated negative effect of age (i.e.,: younger adults checked the clock more frequently than the older adults). In both forest plots, the size of the circles indicates the relative weight assigned to that study in the analysis. Error bars represent the 95% confidence interval of the effect size of each study and, below them, the combined effect size is reported with its confidence interval (the black diamond) and its prediction interval (the dotted line). Studies are sorted in ascending order by effect size.

frequency was associated with a reduction of the age effect in TBPM performance ($\beta = -0.05$; $p = 0.031$; Figure 4(A), upper right panel). Finally, there was a significant negative interaction between duration of the PM target time and PM task frequency ($\beta = -0.02$; $p = 0.009$), meaning that the longer the frequency task duration, the stronger was the effect of the PM target time duration in reducing age differences in TBPM performance (Figure 4(A), lower panel). Similarly with the results on TBPM performance, predictors explained the most of the variance of the age effect in time monitoring too ($r^2 = 59.36\%$). Specifically, the model indicated that longer durations of target time were associated with a reduction of the age effect in time monitoring ($\beta = -0.11$; $p = 0.014$; Figure 4B, upper left panel); moreover, while the main effect of PM task frequency ($\beta = -0.07$; $p = 0.110$; Figure 4(B), upper right panel) was not significant, there was a significant interaction between duration of the PM target time and PM task frequency ($\beta = -0.032$; $p = 0.026$, Figure 4(B), lower panel). The model indicated that the interval criterion for PM accuracy did not exert any significant effect on both TBPM performance ($\beta = -0.009$; $p = 0.948$) and time monitoring ($\beta = 0.355$; $p = 0.177$). The correlation between age effects on TBPM performance and time

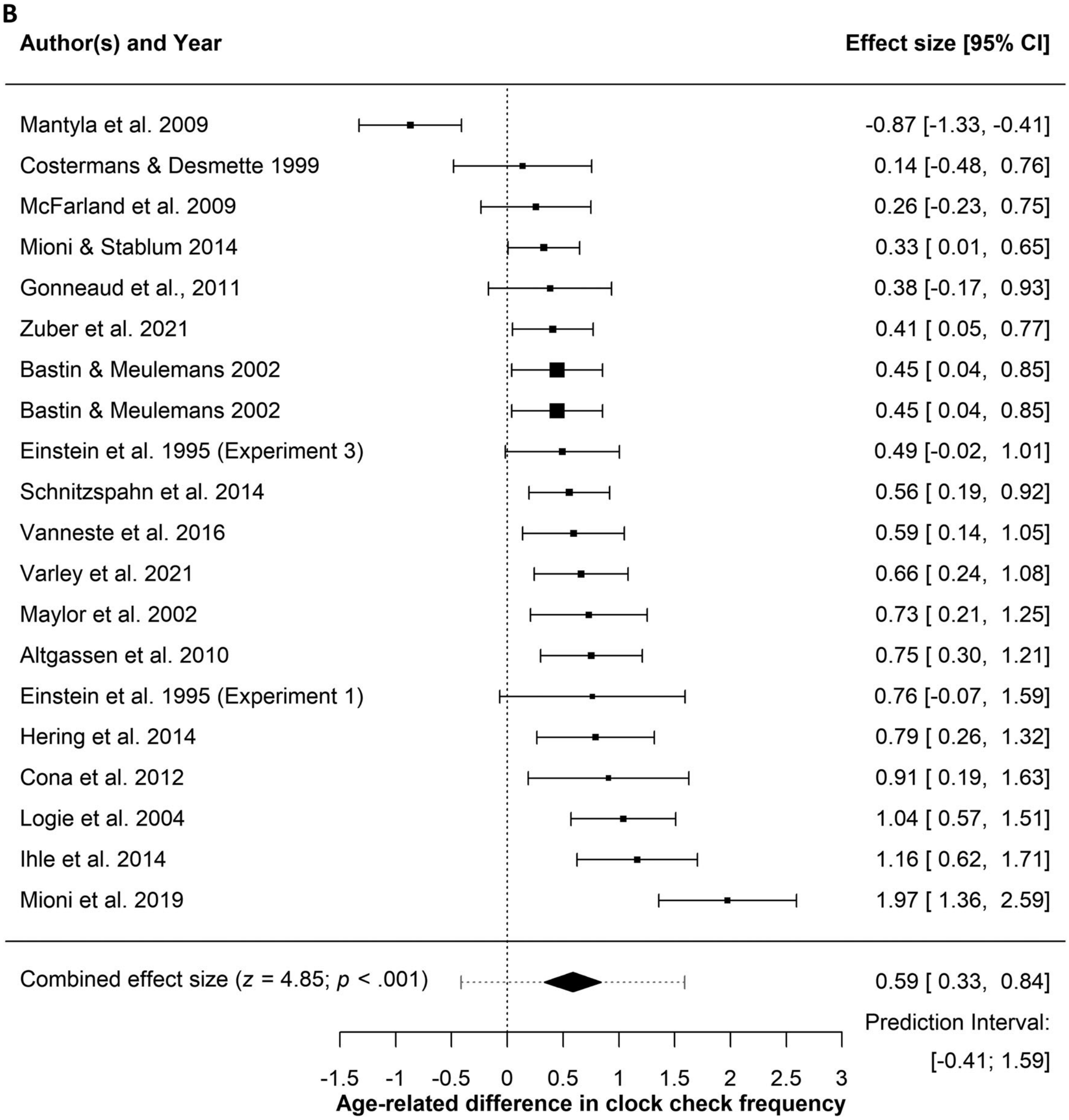

Figure 2 *Continued*

monitoring was $r = 0.61$, suggesting that there was a positive association between the age effect on time monitoring and its effect on TBPM performance.

Discussion

The present meta-analysis aimed (1) to quantify age-related differences in TBPM and time monitoring among studies that used laboratory tasks, (2) to estimate the relationship between age effects in both TBPM performance and time monitoring, in order to determine if studies with a high effect size on one outcome also have higher effect sizes on the other outcome, and (3) to measure whether and how specific task-related factors (i.e., the duration of the PM target time, the frequency of the PM task, and the interval criterion for correct PM responses) affect age-related differences in TBPM performance and time monitoring. The meta-analysis comprised 43 studies reporting compatible measures of PM performance (sum or proportional accuracy); 17 of which reported measures of total time monitoring.

Overview of the results

In summary, our meta-analysis found that younger adults performed better than older adults in TBPM tasks (*g*: 1.06; Figure 2(A)). This result is in line with the previous meta-analysis showing a negative effect of age in laboratory TBPM task, with an effect size of 0.39 (Henry et al., 2004). Our larger effect size might be due to the number of studies: while Henry and colleagues (2004) included 6 studies, we included 44 studies; this huge difference is due to the increasing number of studies in the last 15 years, which in turn affect the magnitude of the age

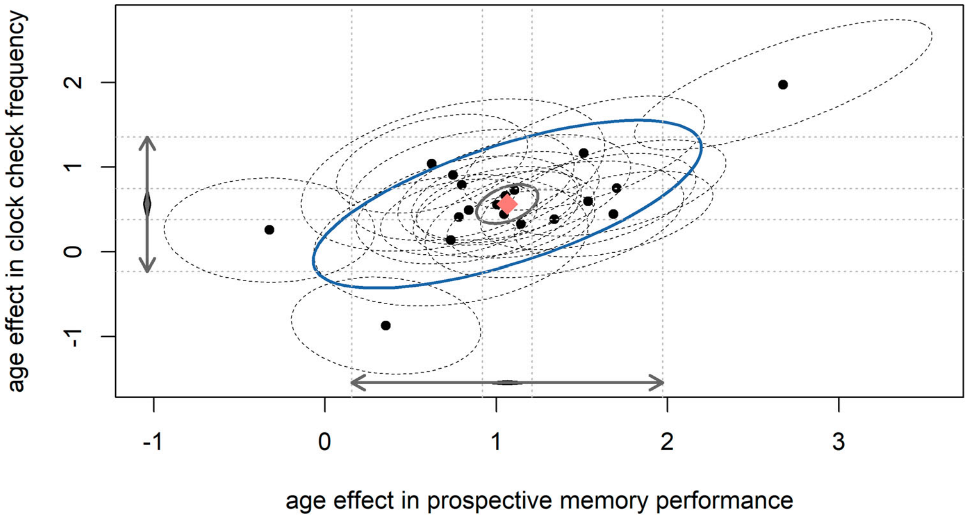

Figure 3. Association between age effects in time-based prospective memory performance and time monitoring. *Note*. Effect sizes and confidence ellipses for age differences in time-based prospective memory performance and total time monitoring. The x-axis displays the age effect in time-based prospective memory performance, while the y-axis displays the age effect in time monitoring. Along each respective axis, the pooled age effect and its 95% confidence interval is represented black diamonds; the bi-directional rows along each axis indicate the prediction intervals for each outcome. In the middle of the plot, the pooled effect of both variables is shown as a red diamond; the smaller dark-grey ellipse represents the 95% confidence interval of the pooled age effects, while the larger turquoise ellipse depicts the 95% prediction interval.

effect (Hak et al., 2016). We also found that younger adults checked the clock more often than older adults (*g*: 0.59; Figure 2(B)); however, there was substantial unexplained heterogeneity, which dropped considerably when outliers were removed (i.e., from from $I^2 = 66.39\%$ to 46.10%, and from $I^2 = 78.86\%$ to 12.81%, for TBPM performance and time monitoring, respectively; see Supplementary material for more information). The results from the multi-variate model showed that studies finding significant age effects in TBPM performance seemed to have 67% of probability to find significant age effects in time monitoring too, suggesting that there is a strong relationship between age effects in the two outcomes (Figure 3). The meta-regression analysis showed that the duration of the PM target time was negatively related to age effects in both TBPM performance and time monitoring (Figure 4(A), upper left panel, and 4B, upper left panel), and PM task frequency was negatively related to the age effect in TBPM performance (Figure 4(A), upper right panel) but not in time monitoring (Figure 4(B), upper right panel). For both time monitoring and TPM performance, a significant negative interaction between PM task frequency and duration of the PM target time was found, indicating that the negative relationship between age effects and duration of the PM target time was more pronounced for task blocks with multiple PM cues compared to 1- (single-item) PM task blocks (Figure 4(A,B), lower panels). Interestingly, this model seemed to explain the between-studies heterogeneity introduced by the outliers, as shown by the τ^2 which was considerably reduced in this model compared to the one without predictors (i.e., from $\tau^2 = 0.214$ to $\tau^2 = 0.075$ for TBPM performance, and from $\tau^2 = 0.164$ to $\tau^2 = 0.067$ for time monitoring).

Conceptual and methodological implications

As mentioned above, time monitoring is essential for TBPM accuracy (Ceci & Bronfenbrenner, 1985; Harris & Wilkins, 1982; Mäntylä et al., 2006; Mioni et al., 2019; Mioni & Stablum, 2014; Vanneste et al., 2016); yet, the cognitive processes underlying age differences in time monitoring and TBPM are still an open debate. Some authors argued that age-related impairments are related to time estimation ability, especially involved in time monitoring (Labelle et al., 2009; Mioni & Stablum, 2014; Vanneste et al., 2016), whereas others argued that attentional processes, such as task-switching, are responsible for age differences in TBPM (Lecouvey et al., 2017; Varley et al., 2021; Zuber & Kliegel, 2020). Other authors argued that shorter PM target times (≤ 2 min) involved more attentional control processes (e.g., task switching) than longer PM target times (> 2 min; Bastin & Meulemans, 2002; Conte & McBride, 2018). This argument was confirmed by a recent study showing that age differences in a 1-minute TBPM task were due to impairments in attentional processes (Varley et al., 2021). Nonetheless, it is not clear yet whether and how the duration of the PM target time affect attention and/or time estimation processes in aging (Block & Zakay, 2006; Mioni & Stablum, 2014). The average duration of the PM target time in the studies included in the meta-analysis was 4 min, ranging from 30 s (Gonneaud et al., 2017; Morand et al., 2021) to 10 min (Einstein et al., 1995; Niedźwieńska & Barzykowski, 2012). The results showed that younger adults were more accurate and checked the clock more frequently in the TBPM task, especially for shorter intervals (less than 4 min). It's possible that the age differences for shorter intervals are

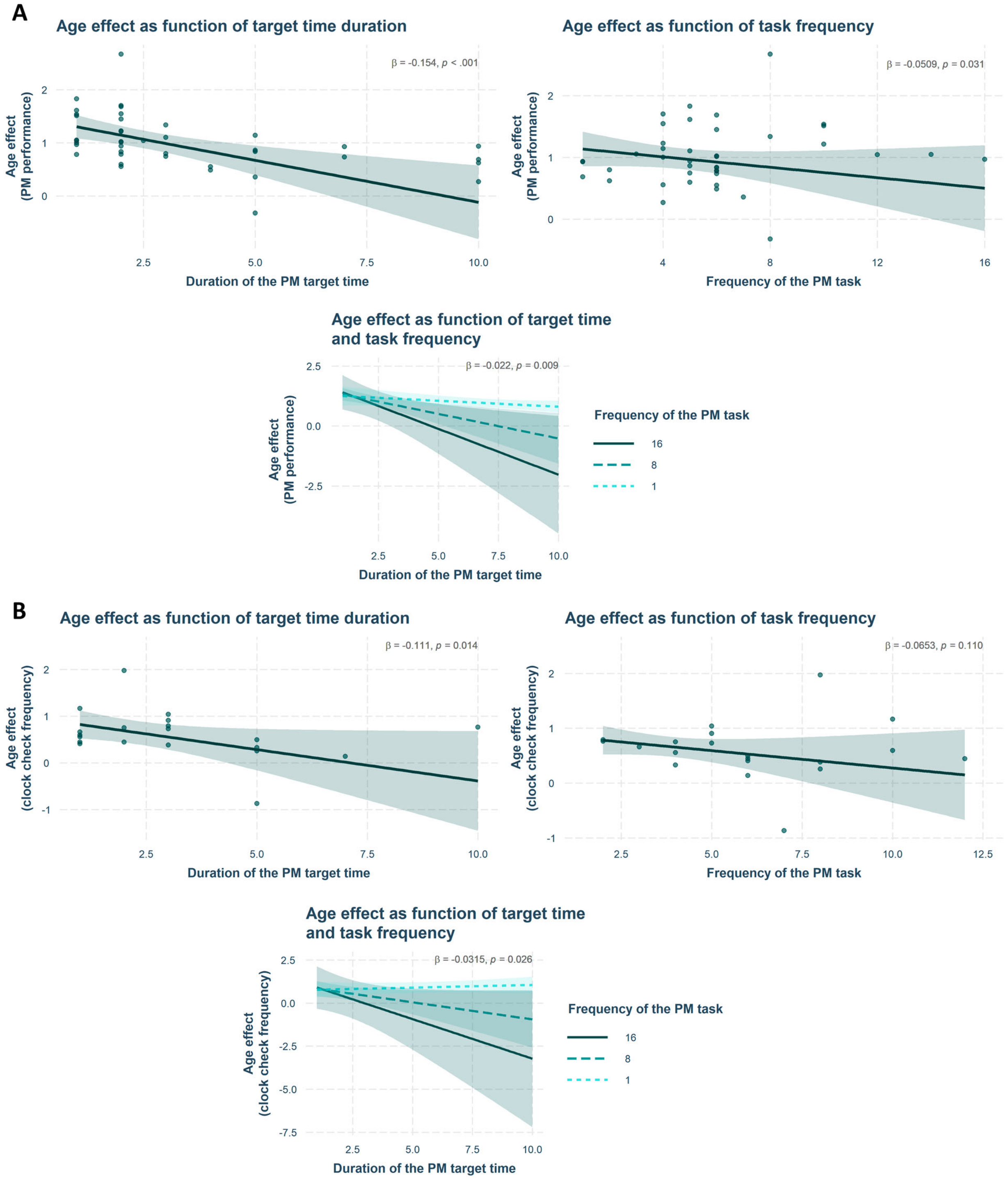

Figure 4. Age effect as function of PM target time's duration and PM task frequency. *Note*. Graphical representation of age differences in time-based prospective memory performance (4A) and time monitoring (4B) as function of target time's duration and frequency of the prospective memory task; PM: prospective memory.

due to the involvement of attentional control processes (Conte & McBride, 2018) that are particularly impaired with aging (Craik, 1986; Varley et al., 2021). Diversely, longer PM target times (i.e., ≥ 4 min) may either allow more time to better distribute the attentional resources between OT and PM task, as well as engage more time estimation abilities compared to short PM target times, reducing the age differences in time monitoring and TBPM accuracy (Mioni et al., 2020, 2021; Varley et al., 2021).

Among the studies included in the present meta-analysis, the frequency of PM target time ranged from 1 (Rendell & Thomson, 1999) to 16 (Pupillo et al., 2021; Shum et al., 2013). The meta-analysis showed that PM task frequency had a selective effect on TBPM performance, but not on time monitoring. Age differences in TBPM performance were influenced by PM task frequency, especially when the PM target time was longer, suggesting that learning from task repetition could counteract or reduce age effects. PM task frequency had no effect on age differences in time monitoring, indicating that these differences are not influenced by task repetition but by PM target time duration; however, this last finding should be taken carefully as there were fewer studies that measures time monitoring (i.e., 18) and thus fewer observations of these effects as for TBPM accuracy, for which there were more studies (i.e., 44); therefore, it is not possible to fully exclude the presence of learning effects in time monitoring too. Indeed, age differences in time monitoring were reduced by PM task frequency, but only when the PM target time was longer (Figure 4(B), lower panel), suggesting that learning from task repetition could counteract or attenuate age effects in time monitoring which, in turn, affect age differences in TBPM performance. Future studies are needed to investigate this specific effect experimentally.

The current meta-analysis has several important methodological implications for the design and interpretation of future studies in the field. Firstly, with regards to the design of future tasks, the results of the meta-analysis showed that age effects in TBPM performance and time monitoring were significant when the TBPM task consisted of 6 PM target times, lasting 4 min each. As such, future studies that aim to detect age differences in TBPM should replicate these parameters. It is important to note that these parameters can be changed based on the specific research needs; yet, researchers should be aware that changing the number of cues and target time duration can impact the magnitude of age effects, and this may be reflective of different cognitive mechanisms, such as task-switching and time estimation, which may interact with learning processes. Another methodological implication concerns the criterion used to determine TBPM accuracy and its impact on age differences in TBPM performance and time monitoring (Yang et al., 2013). The studies included in the meta-analysis used a wide range of interval criteria, ranging from 2% (Costermans & Desmette, 1999) to 100% (Haas et al., 2020; Martin & Schumann-Hengsteler, 2001; Varley et al., 2021; Yang et al., 2013) of the total PM target time. Despite this variability, the results of the meta-analysis showed that age differences were detected regardless of the interval criterion chosen by the researchers, indicating that the TBPM paradigm is robust for studying age effects. As such, the choice of criterion for PM accuracy should not be a concern for researchers, as significant age differences are likely to emerge regardless of the criterion used. Finally, the meta-analysis highlights the importance of including measures of time monitoring in any TBPM experiment, as it is highly correlated with TBPM and essential for understanding the cognitive processes underlying TBPM performance, as supported by previous research that has demonstrated the close relationship between time monitoring and TBPM (Ceci & Bronfenbrenner, 1985; Harris & Wilkins, 1982; Mäntylä et al., 2006; Mioni et al., 2019; Mioni & Stablum, 2014; Vanneste et al., 2016).

Finally, age effects in TBPM performance and time monitoring can be interpreted into the broader context of aging in human memory (Bopp & Verhaeghen, 2005). Specifically, while recollection is disrupted by aging, recognition is usually spared (Yonelinas, 2002). These differences in memory recollection are not solely due to the size of the hippocampus (Van Petten, 2004), but may also be determined by compensatory frontally-mediated executive functions that are engaged in self-initiated processes involved in memory recall (Cabeza et al., 2018; Craik, 1986; West, 1996). In line with such explanation, TBPM has been shown to be particularly difficult for older adults as it requires similar self-initiated processes (Lewis-Peacock et al., 2016; Martin-Ordas et al., 2010; McDaniel et al., 2015), involving executive functions and cognitive control processes (Cruz et al., 2017; Zuber & Kliegel, 2020) that are particularly impaired in aging (Cabeza et al., 2018; Craik, 1986; West, 1996). However, recent meta-analytic evidence challenged the explanation of self-initiated processes being the only source of age-related differences in memory performance (Craik, 1986; West, 1996), as age effects can also be observed in recognition (Fraundorf et al., 2019), which presumably should not involve frontally-mediated executive functions (McDaniel et al., 2015; Yonelinas, 2002; Yonelinas et al., 2010). Therefore, it is still unclear whether the age-related deficit reflects a general deficit related to self-initiated processes that affects globally all memory tasks, and further studies are needed to understand this aspect related to aging in human memory.

Limitations & future directions

The present meta-analysis has some limitations. The first limitation concerns the lack of analysis on the strategic aspects of time monitoring, which can be investigated measuring monitoring over time. Several studies have shown that people usually check the clock few times as the task starts, and then increase the number of clock checks as the PM target time approaches, forming a "J-shaped" curve (Labelle et al., 2009; Mäntylä et al., 2006; Mioni et al., 2019; Vanneste et al., 2016); such strategic behavior is associated with better PM accuracy (Ceci & Bronfenbrenner, 1985; Harris & Wilkins, 1982; Mäntylä et al., 2006; Mioni et al., 2012, 2019; Vanneste et al., 2016; Waldum & McDaniel, 2016). A recent study disentangled the respective contribution of total versus strategic time monitoring to the age differences in TBPM performance. The authors proposed a more fine-grained indicator of strategic behaviour (i.e., relative monitoring), which accounts for interindividual differences in the total frequency of clock checks (i.e., absolute monitoring). The results showed that both relative and absolute monitoring fully mediated the negative age effect on TBPM; yet, relative monitoring was a stronger predictor of TBPM performance than absolute monitoring (Joly-Burra et al., 2022). In the present meta-analysis, we could not code monitoring over time given that each study used different PM target times and analysed monitoring using different intervals, according to the specific research needs. Thus, any inference on strategic time monitoring should be taken carefully considering these current meta-analytic results, and future studies are needed to investigate the strategic aspect of time monitoring. Another limitation is the lack of comparison between laboratory and naturalistic setting (Cauvin et al., 2019; Haas et al., 2020; Kvavilashvili & Fisher, 2007; Maylor, 1990; McBride et al., 2013; Rendell & Thomson, 1993). As mentioned in the introduction, we decided to focus only on laboratory tasks because, as far as known by the authors, there are no studies in the literature that have developed a method for measuring time monitoring in naturalistic settings; hence, considering the relevant role of time monitoring in TBPM, we decided to focus only on laboratory studies. However, with the development of new technologies, such as electronic pads and smartwatches, future studies could develop an experimental protocol to measure time monitoring in naturalistic contexts, allowing future meta-analytic comparisons between laboratory and naturalistic assessments.

Finally, it is also possible that other factors, such as cognitive demands and stimulus material of the ongoing task, may have influenced the results. Indeed, the multivariate analysis with the predictors showed that there was still a significant portion of unexplained variance in the age effect in TBPM performance ($\tau^2 = 0.075$, $p = 0.038$); it cannot be excluded that such unexplained variance could be due to different OTs. Most of studies in the meta-analysis used traditional cognitive tasks such as working memory (Pupillo et al., 2021; Zuber et al., 2021) or arithmetic tasks (D'Ydewalle et al., 2001; Gonneaud et al., 2011); others used more passive OT such as watching a movie (Logie et al., 2004; Mioni & Stablum, 2014), whereas few studies used alternative OTs such as trivia or jigsaw puzzle (Einstein et al., 1995; Waldum & McDaniel, 2016); some studies used even different OTs across TBPM task blocks (D'Ydewalle et al., 1999; Niedźwieńska & Barzykowski, 2012). Our analysis did not examine the specific nature of the OT, but it cannot be excluded that these factors could also influence age effects in TBPM (D'Ydewalle et al., 2001; Khan et al., 2008; Meier & Zimmermann, 2015). Moreover, it is also possible that PM task frequency and/or the duration of the PM target time were related to the nature of the OT. Future research should consider conducting a more in-depth examination of the role of OT typology in time monitoring and TBPM performance, to provide a clearer understanding of the relationships between this further factor and age-related changes in TBPM.

Conclusions

Overall, this meta-analysis provided an update on age differences in TBPM accuracy and their potential effect size (Henry et al., 2004), as well as a first meta-analytic quantification of the age difference in time monitoring, investigating the contribution of task-related features, namely the duration of the PM target time, the frequency of the PM task, and the criterion of PM accuracy. Our meta-analytical results have both conceptual and methodological implications. Conceptually, the results of the meta-analysis suggested that the age effect emerged consistently for shorter (e.g., ≤ 4 min) rather than longer intervals, probably because of age-related impairment in the attentional processes. Moreover, the effect of the PM target time's duration interacted with the frequency of the PM task, suggesting that there might be some learning effects that can attenuate the magnitude of age effects, especially for longer durations. Concerning the possible methodological implications, it is reasonable to argue that, regardless how researchers code accuracy, TBPM paradigm can detect age-related differences consistently; yet researchers should be aware that changing task-related parameters such as the frequency of the PM task and the duration of the PM target time can affect the magnitude of the age effect in both time monitoring and TBPM performance. In summary, the present meta-analysis can help the conceptual understanding of the cognitive processes underlying age effect in time monitoring and TBPM performance, also providing a methodological framework that can guide future aging research.

Note

1. Age groups were not defined a-priori because there is no agreement on how to define age groups; indeed, all studies used different age ranges for younger and older adults: hence, we extracted the age groups as they were reported within each study, regardless of the differences in age ranges across studies.

Acknowledgements

The authors are grateful to the authors of the reviewed articles for their collaboration in providing the necessary information for this publication; moreover, the authors are grateful to Christina Moses-Pasini for the methodological and analytic advice. Matthias Kliegel acknowledges support from the Swiss National Science Foundation (SNSF).

Disclosure statement

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

Funding

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

ORCID

F. Borghese http://orcid.org/0000-0002-8473-3652
A. Hering http://orcid.org/0000-0002-7952-9032
M. Kliegel http://orcid.org/0000-0002-2001-2522
G. Mioni http://orcid.org/0000-0002-1212-4591

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., Kliegel, M., Brandimonte, M., & Filippello, P. (2010). Are older adults more social than younger adults? Social importance increases older adults' prospective memory performance. *Aging, Neuropsychology, and Cognition*, *17*(3), 312–328. https://doi.org/10.1080/13825580903281308

Assink, M., & Wibbelink, C. J. M. (2016). Fitting three-level meta-analytic models in R: A step-by-step tutorial. *The Quantitative Methods for Psychology*, *12*(3), 154–174. https://doi.org/10.20982/tqmp.12.3.p154

Balduzzi, S., Rücker, G., & Schwarzer, G. (2019). How to perform a meta-analysis with R: A practical tutorial. *Evidence Based Mental Health*, *22*(4), 153–160. https://doi.org/10.1136/ebmental-2019-300117

Bastin, C., & Meulemans, T. (2002). Are time-based and event-based prospective memory affected by normal aging in the same way? *Current Psychology Letters*, *7*, 105–121. https://doi.org/10.4000/cpl.154

Behrendt, S., Kliegel, M., Kräplin, A., & Bühringer, G. (2015). Performance of smokers with DSM-5 tobacco Use disorder in time-based complex prospective memory. *Journal of Psychoactive Drugs*, *47*(3), 203–212. https://doi.org/10.1080/02791072.2015.1054008

Bezdicek, O., Nikolai, T., Nepožitek, J., Peřinová, P., Kemlink, D., Dušek, P., Příhodová, I., Dostálová, S., Ibarburu, V., Trnka, J., Kupka, K., Mecková, Z., Keller, J., Vymazal, J., Růžička, E., Šonka, K., & Dušek, P. (2018). Prospective memory impairment in idiopathic REM sleep behavior disorder. *Clinical Neuropsychologist*, *32*(5), 1019–1037. https://doi.org/10.1080/13854046.2017.1394493

Block, R. A., & Zakay, D. (2006). Prospective remembering involves time estimation and memory processes. In J. Glicksohn & M. S. Myslobodsky (Eds.), *Timing the future* (pp. 25–49). World Scientific Publishing Co. https://doi.org/10.1142/9789812707123_0002

Bopp, K. L., & Verhaeghen, P. (2005). Aging and verbal memory span: A meta-analysis. *The Journals of Gerontology: Series B*, *60*(5), 223–P233. https://doi.org/10.1093/geronb/60.5.P223

Borenstein, M., Hedges, L., & Rothstein, H. (2009). *Introduction to Meta-Analysis*.

Bozdemir, M., & Cinan, S. (2021). Age-related differences in intentional forgetting of prospective memory. *The International Journal of Aging and Human Development*, *92*(3), 350–363. https://doi.org/10.1177/0091415019900165

Cabeza, R., Albert, M., Belleville, S., Craik, F. I. M., Duarte, A., Grady, C. L., Lindenberger, U., Nyberg, L., Park, D. C., Reuter-Lorenz, P. A., Rugg, M. D., Steffener, J., & Rajah, M. N. (2018). Maintenance, reserve and compensation: The cognitive neuroscience of healthy ageing. *Nature Reviews Neuroscience*, *19*(11), 701–710. https://doi.org/10.1038/s41583-018-0068-2

Cauvin, S., Moulin, C., Souchay, C., Schnitzspahn, K. M., & Kliegel, M. (2019). Laboratory vs. naturalistic prospective memory task predictions: Young adults are overconfident outside of the laboratory. *Memory (Hove, England)*, *27*(5), 592–602. https://doi.org/10.1080/09658211.2018.1540703

Ceci, S. J., & Bronfenbrenner, U. (1985). 'Don't forget to take the cupcakes out of the oven': Prospective memory, strategic time-monitoring, and context. *Child Development*, *56*(1), 152–164. https://doi.org/10.2307/1130182

Cheung, M. W.-L. (2014). Modeling dependent effect sizes with three-level meta-analyses: A structural equation modeling approach. *Psychological Methods*, *19*(2), 211–229. https://doi.org/10.1037/a0032968

Cheung, M. W.-L. (2015). metaSEM: An R package for meta-analysis using structural equation modeling. *Frontiers in Psychology*, *5*. https://www.frontiersin.org/articles/10.3389fpsyg.2014.01521

Cochran, W. G. (1954). The combination of estimates from different experiments. *Biometrics*, *10*(1), 101–129. https://doi.org/10.2307/3001666

Cona, G., Arcara, G., Tarantino, V., & Bisiacchi, P. S. (2012). Age-related differences in the neural correlates of remembering time-based intentions. *Neuropsychologia*, *50*(11), 2692–2704. https://doi.org/10.1016/j.neuropsychologia.2012.07.033

Conte, A. M., & McBride, D. M. (2018). Comparing time-based and event-based prospective memory over short delays. *Memory*, *26*(7), 936–945. https://doi.org/10.1080/09658211.2018.1432662

Costa, A., Peppe, A., Brusa, L., Caltagirone, C., Gatto, I., & Carlesimo, G. A. (2008). Levodopa improves time-based prospective memory in Parkinson's disease. *Journal of the International Neuropsychological Society*, *14*(4), 601–610. https://doi.org/10.1017/S135561770808082X

Costa, A., Zabberoni, S., Peppe, A., Serafini, F., Scalici, F., Caltagirone, C., & Carlesimo, G. A. (2015). Time-based prospective memory functioning in mild cognitive impairment associated with Parkinson's disease: Relationship with autonomous management of daily living commitments. *Frontiers in Human Neuroscience*, *9*(June), 1–10. https://doi.org/10.3389/fnhum.2015.00333

Costermans, J., & Desmette, D. (1999). A method for describing time-monitoring strategies in a prospective memory setting. *Current Psychology of Cognition*, *18*(3), 289–306.

Craik, F. I. M. (1986). A functional account of age differences in memory. In F. Klix & H. Hagendorf (Eds.), *Human memory and cognitive capabilities: Mechanisms and performances* (pp. 409–422). Amsterdam: Elsevier.

Cruz, G., Burgos, P., Kilborn, K., & Evans, J. J. (2017). Involvement of the anterior cingulate cortex in time-based prospective memory task monitoring: An EEG analysis of brain sources using independent component and measure projection analysis. *PLoS ONE*, *12*(9), 1–28. https://doi.org/10.1371/journal.pone.0184037

D'Ydewalle, G., Bouckaert, D., & Brunfaut, E. (2001). Age-related differences and complexity of ongoing activities in time- and event-based prospective memory. *The American Journal of Psychology*, *114*(3), 411–423. https://doi.org/10.2307/1423688

D'Ydewalle, G., Luwel, K., & Brunfaut, E. (1999). The importance of on-going concurrent activities as a function of age in time- and event-based prospective memory. *European Journal of Cognitive Psychology*, *11*(2), 219–237. https://doi.org/10.1080/713752309

Einstein, G. O., Holland, L. J., McDaniel, M. A., & Guynn, M. J. (1992). Age-related deficits in prospective memory: The influence of task complexity. *Psychology and Aging*, *7*(3), 471–478. https://doi.org/10.1037/0882-7974.7.3.471

Einstein, G. O., McDaniel, M. A., Richardson, S. L., Guynn, M. J., & Cunfer, A. R. (1995). Aging and prospective memory: Examining the influences of self-initiated retrieval processes. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(4), 996–1007. https://doi.org/10.1037/0278-7393.21.4.996

Esposito, M. J., Occhionero, M., & Cicogna, P. (2015). Sleep deprivation and time-based prospective memory. *Sleep*, *38*(11), 1823–1826. https://doi.org/10.5665/sleep.5172

Fraundorf, S. H., Hourihan, K. L., Peters, R. A., & Benjamin, A. S. (2019). Aging and recognition memory: A meta-analysis. *Psychological Bulletin*, *145*(4), 339–371. https://doi.org/10.1037/bul0000185

Gan, J., & Guo, Y. (2019). The cognitive mechanism of the practice effect of time-based prospective memory: The role of time estimation. *Frontiers in Psychology*, *10*(December), 1–7. https://doi.org/10.3389/fpsyg.2019.02780

Gonneaud, J., Kalpouzos, G. G., Bon, L., Viader, F., Eustache, F., & Desgranges, B. A. (2011). Distinct and shared cognitive functions mediate event and time-based prospective memory impairment in normal ageing. *Memory*, *19*(4), 360–377. https://doi.org/10.1080/09658211.2011.570765

Gonneaud, J., Lecouvey, G., Groussard, M., Gaubert, M., Landeau, B., Mézenge, F., de La Sayette, V., Eustache, F., Desgranges, B., & Rauchs, G. (2017). Functional dedifferentiation and reduced task-related deactivations underlie the age-related decline of prospective memory. *Brain Imaging and Behavior*, *11*(6), 1873–1884. https://doi.org/10.1007/s11682-016-9661-z

Gonneaud, J., Rauchs, G., Groussard, M., Landeau, B., Mézenge, F., de La Sayette, V., Eustache, F., & Desgranges, B. (2014). How do we process event-based and time-based intentions in the brain? An fMRI study of prospective memory in healthy individuals. *Human Brain Mapping*, *35*(7), 3066–3082. https://doi.org/10.1002/hbm.22385

Guimond, A., Braun, C. M. J., Rouleau, I., Bélanger, F., & Godbout, L. (2006). Remembering the past and foreseeing the future while dealing with the present: A comparison of young adult and elderly cohorts on a multitask simulation of occupational activities. *Experimental Aging Research*, *32*(3), 363–380. https://doi.org/10.1080/03610730600699100

Guo, Y., & Huang, X. (2019). Time-based prospective memory has plasticity in behavior under different monitoring conditions. *Current Psychology*, *40*, 3386–3392. https://doi.org/10.1007/s12144-019-00270-5

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*, *77*(4), 695–703. https://doi.org/10.1093/geronb/gbac012

Haas, M., Zuber, S., Kliegel, M., & Ballhausen, N. (2020). Prospective memory errors in everyday life: Does instruction matter? *Memory*, *28*(2), 196–203. https://doi.org/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*. https://doi.org/10.1016/j.cognition.2020.104305

Hak, T., van Rhee, H., & Suurmond, R. (2016). How to interpret results of meta-analysis. *Social Science Research Network*. https://doi.org/10.2139/ssrn.3241367

Harrer, M., Cuijpers, P., F, A., & Ebert, T., & D, D. (2021). *Doing meta-analysis with R: A hands-On guide* (1st ed.). Chapman & Hall/CRC Press.

Harris, J., & Wilkins, A. (1982). Remembering to do things: A theoretical framework and an illustrative experiment. *Human Learning: Journal of Practical Research Applications*, *1*, 123–136.

Hedges, L. V. (1981). Distribution theory for glass's estimator of effect size and related estimators. *Journal of Educational Statistics*, *6*(2), 107–128. https://doi.org/10.2307/1164588

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., Rendell, P. G., Phillips, L. H., Dunlop, L., & Kliegel, M. (2012). Prospective memory reminders: A laboratory investigation of initiation source and age effects. *Quarterly Journal of Experimental Psychology*, *65*(7), 1274–1287. https://doi.org/10.1080/17470218.2011.651091

Henry, J. D., Terrett, G., Grainger, S. A., Rose, N. S., Kliegel, M., Bugge, M., Ryrie, C., & Rendell, P. G. (2020). Implementation intentions and prospective memory function in late adulthood. *Psychology and Aging*, *35*(8), 1105–1114. https://doi.org/10.1037/pag0000563

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. https://doi.org/10.1017/S1355617718000152

Hering, A., Rendell, P. G., Rose, N. S., Schnitzspahn, K. M., & Kliegel, M. (2014). Prospective memory training in older adults and its relevance for successful aging. *Psychological Research*, *78*(6), 892–904. https://doi.org/10.1007/s00426-014-0566-4

Higgins, J. P. T., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine*, *21*(11), 1539–1558. https://doi.org/10.1002/sim.1186

Huang, T., Loft, S., & Humphreys, M. S. (2014). Internalizing versus externalizing control: Different ways to perform a time-based prospective memory task. *Journal of Experimental Psychology: Learning Memory and Cognition*, *40*(4), 1064–1071. https://doi.org/10.1037/a0035786

Ihle, A., Kliegel, M., Hering, A., Ballhausen, N., Lagner, P., Benusch, J., Cichon, A., Zergiebel, A., Oris, M., & Schnitzspahn, K. M. (2014). Adult age differences in prospective memory in the laboratory: Are they related to higher stress levels in the elderly? *Frontiers in Human Neuroscience*, *8*. https://doi.org/10.3389/fnhum.2014.01021

Jäger, T., & Kliegel, M. (2008). Time-based and event-based prospective memory across adulthood: Underlying mechanisms and differential costs on the ongoing task. *Journal of General Psychology*, *135*(1), 4–22. https://doi.org/10.3200/GENP.135.1.4-22

Joly-Burra, E., Haas, M., Laera, G., Ghisletta, P., Kliegel, M., & Zuber, S. (2022). Frequency and strategicness of clock-checking explain detrimental age effects in time-based prospective memory. *Psychology and Aging*, *37*(5), 637–648. https://doi.org/10.1037/pag0000693

Khan, A., Sharma, N. K., & Dixit, S. (2008). Cognitive load and task condition in event- and time-based prospective memory: An experimental investigation. *Journal of Psychology: Interdisciplinary and Applied*, *142*(5), 517–531. https://doi.org/10.3200/JRLP.142.5.517-532

Kvavilashvili, L., & Fisher, L. (2007). Is time-based prospective remembering mediated by self-initiated rehearsals? Role of incidental cues, ongoing activity, age, and motivation. *Journal of Experimental Psychology: General*, *136*(1), 112–132. https://doi.org/10.1037/0096-3445.136.1.112

Labelle, M. A., Graf, P., Grondin, S., & Gagné-Roy, L. (2009). Time-related processes in time-based prospective memory and in time-interval production. *European Journal of Cognitive Psychology*, *21*(4), 501–521. https://doi.org/10.1080/09541440802031000

Lecouvey, G., Gonneaud, J., Piolino, P., Madeleine, S., Orriols, E., Fleury, P., Eustache, F., & Desgranges, B. (2017). Is binding decline the main source of the ageing effect on prospective memory? A ride in a virtual town. *Socioaffective Neuroscience & Psychology*, *7*(1), 1304610. https://doi.org/10.1080/20009011.2017.1304610

Lewis-Peacock, J. A., Cohen, J. D., & Norman, K. A. (2016). Neural evidence of the strategic choice between working memory and episodic memory in prospective remembering. *Neuropsychologia*, *93* (October), 280–288. https://doi.org/10.1016/j.neuropsychologia.2016.11.006

Logie, R., Maylor, E., Della Sala, S., & Smith, G. (2004). Working memory in event- and time-based prospective memory tasks: Effects of secondary demand and age. *European Journal of Cognitive Psychology*, *16*(3), 441–456. https://doi.org/10.1080/09541440340000114

Mäntylä, T., Carelli, M. G., & Forman, H. (2006). Time monitoring and executive functioning in children and adults. *Journal of Experimental Child Psychology*, *96*(1), 1–19. https://doi.org/10.1016/j.jecp.2006.08.003

Mäntylä, T., Missier, F. D., & Nilsson, L.-G. (2009). Age differences in multiple outcome measures of time-based prospective memory. *Aging, Neuropsychology, and Cognition*, *16*(6), 708–720. https://doi.org/10.1080/13825580902912721

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

Martin, M., & Schumann-Hengsteler, R. (2001). How task demands influence time-based prospective memory performance in young and older adults. *International Journal of Behavioral Development*, *25*(4), 386–391. https://doi.org/10.1080/01650250042000302

Martin-Ordas, G., Haun, D., Colmenares, F., & Call, J. (2010). Keeping track of time: Evidence for episodic-like memory in great apes. *Animal Cognition*, *13*(2), 331–340. https://doi.org/10.1007/s10071-009-0282-4

Maylor, E. A. (1990). Age and prospective memory. *The Quarterly Journal of Experimental Psychology Section A*, *42*(3), 471–493. https://doi.org/10.1080/14640749008401233

Maylor, E. A., Smith, G., Sala, S. D., & Logie, R. H. (2002). Prospective and retrospective memory in normal aging and dementia: An experimental study. *Memory & Cognition*, *30*(6), 871–884. https://doi.org/10.3758/BF03195773

McBride, D. M., Beckner, J. K., & Abney, D. H. (2011). Effects of delay of prospective memory cues in an ongoing task on prospective memory task performance. *Memory and Cognition*, *39*(7), 1222–1231. https://doi.org/10.3758/s13421-011-0105-0

McBride, D. M., Coane, J. H., Drwal, J., & LaRose, S. A. M. (2013). Differential effects of delay on time-based prospective memory in younger and older adults. *Aging, Neuropsychology, and Cognition*, *20*(6), 700–721. https://doi.org/10.1080/13825585.2013.765937

McDaniel, M. A., Umanath, S., Einstein, G. O., & Waldum, E. R. (2015). Dual pathways to prospective remembering. *Frontiers in Human Neuroscience*, *9*, 392. https://doi.org/10.3389/fnhum.2015.00392

McFarland, C. P., & Glisky, E. L. (2009). Frontal lobe involvement in a task of time-based prospective memory. *Neuropsychologia*, *47*(7), 1660–1669. https://doi.org/10.1016/j.neuropsychologia.2009.02.023

Meier, B., & Zimmermann, T. D. (2015). Loads and loads and loads: The influence of prospective load, retrospective load, and ongoing task load in prospective memory. *Frontiers in Human Neuroscience*, *9* (June), 1–12. https://doi.org/10.3389/fnhum.2015.00322

Meier, B., Zimmermann, T. D., & Perrig, W. J. (2006). Retrieval experience in prospective memory: Strategic monitoring and spontaneous retrieval. *Memory*, *14*(7), 872–889. https://doi.org/10.1080/09658210600783774

Mioni, G., Capizzi, M., & Stablum, F. (2020). Age-related changes in time production and reproduction tasks: Involvement of attention and working memory processes. *Aging, Neuropsychology, and Cognition*, *27*(3), 412–429. https://doi.org/10.1080/13825585.2019.1626799

Mioni, G., Cardullo, S., Ciavarelli, A., & Stablum, F. (2021). Age-related changes in time discrimination: The involvement of inhibition, working memory and speed of processing. *Current Psychology: A Journal for Diverse Perspectives on Diverse Psychological Issues*, *40* (5), 2462–2471. https://doi.org/10.1007/s12144-019-00170-8

Mioni, G., Grondin, S., McLennan, S. N., & Stablum, F. (2019). The role of time-monitoring behaviour in time-based prospective memory performance in younger and older adults. *Memory*, *28*(1), 34–48. https://doi.org/10.1080/09658211.2019.1675711

Mioni, G., Rendell, P. G., Stablum, F., Gamberini, L., & Bisiacchi, P. S. (2015). Test–retest consistency of virtual week: A task to investigate prospective memory. *Neuropsychological Rehabilitation*, *25* (3), 419–447. https://doi.org/10.1080/09602011.2014.941295

Mioni, G., Santon, S., Stablum, F., & Cornoldi, C. (2017). Time-based prospective memory difficulties in children with ADHD and the role of time perception and working memory. *Child Neuropsychology*, *23*(5), 588–608. https://doi.org/10.1080/09297049.2016.1172561

Mioni, G., & Stablum, F. (2014). Monitoring behaviour in a time-based prospective memory task: The involvement of executive functions and time perception. *Memory*, *22*(5), 536–552. https://doi.org/10.1080/09658211.2013.801987

Mioni, G., Stablum, F., McClintock, S. M., & Cantagallo, A. (2012). Time-based prospective memory in severe traumatic brain injury patients: The involvement of executive functions and time perception. *Journal of the International Neuropsychological Society*, *18*(4), 697–705. https://doi.org/10.1017/S1355617712000306

Moher, D., Shamseer, L., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., Shekelle, P., & Stewart, L. A. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Systematic Reviews*, *4*(1). https://doi.org/10.1186/2046-4053-4-1

Morand, A., Segobin, S., Lecouvey, G., Gonneaud, J., Eustache, F., Rauchs, G., & Desgranges, B. (2021). Brain substrates of time-based prospective memory decline in aging: A voxel-based morphometry and diffusion tensor imaging study. *Cerebral Cortex*, *31* (1), 396–409. https://doi.org/10.1093/cercor/bhaa232

Morand, A., Segobin, S., Lecouvey, G., Gonneaud, J., Eustache, F., Rauchs, G., & Desgranges, B. (2022). Alterations in resting-state functional connectivity associated to the age-related decline in time-based prospective memory. *Cerebral Cortex*. https://doi.org/10.1093/cercor/bhac349

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. https://doi.org/10.1080/13825585.2011.628374

Nigro, G., Senese, V. P., Natullo, O., & Sergi, I. (2002). Preliminary remarks on type of task and delay in children's prospective memory. *Perceptual and Motor Skills*, *95*(2), 515–519. https://doi.org/10.2466/pms.95.6.515-519

Oksanen, K. M., Waldum, E. R., McDaniel, M. A., & Braver, T. S. (2014). Neural mechanisms of time-based prospective memory: Evidence for transient monitoring. *PLoS ONE*, *9*(3). https://doi.org/10.1371/journal.pone.0092123

Okuda, J., Fujii, T., Ohtake, H., Tsukiura, T., Yamadori, A., Frith, C. D., & Burgess, P. W. (2007). Differential involvement of regions of rostral prefrontal cortex (Brodmann area 10) in time- and event-based prospective memory. *International Journal of Psychophysiology*, *64*(3), 233–246. https://doi.org/10.1016/j.ijpsycho.2006.09.009

Park, D. C., Hertzog, C., Kidder, D. P., Morrell, R. W., & Mayhorn, C. B. (1997). Effect of age on event-based and time-based prospective memory. *Psychology and Aging*, *12*(2), 314–327. https://doi.org/10.1037/0882-7974.12.2.314

Patton, G. W. R., & Meit, M. (1993). Effect of aging on prospective and incidental memory. *Experimental Aging Research*, *19*(2), 165–176. https://doi.org/10.1080/03610739308253929

Platt, B., Kamboj, S. K., Italiano, T., Rendell, P. G., & Curran, H. V. (2016). Prospective memory impairments in heavy social drinkers are partially overcome by future event simulation. *Psychopharmacology*, *233*(3), 499–506. https://doi.org/10.1007/s00213-015-4145-1

Pupillo, F., Phillips, L., & Schnitzspahn, K. (2021). The detrimental effects of mood on prospective memory are modulated by age. *Emotion*, *21*(3), 569–583. https://doi.org/10.1037/emo0000723

Rakap, S., Rakap, S., Evran, D., & Cig, O. (2016). Comparative evaluation of the reliability and validity of three data extraction programs: UnGraph, GraphClick, and DigitizeIt. *Computers in Human Behavior*, *55*, 159–166. https://doi.org/10.1016/j.chb.2015.09.008

R Core Team. (2022). R: A language and environment for statistical computing [Manual]. https://www.R-project.org/

Rendell, P. G., & Craik, F. I. M. (2000). Virtual week and actual week: Age-related differences in prospective memory. *Applied Cognitive Psychology*, *14*(7), S43–S62. https://doi.org/10.1002/acp.770

Rendell, P. G., Phillips, L. H., Henry, J. D., Brumby-Rendell, T., de la Piedad Garcia, X., Altgassen, M., & Kliegel, M. (2011). Prospective memory, emotional valence and ageing. *Cognition and Emotion*, *25*(5), 916–925. https://doi.org/10.1080/02699931.2010.508610

Rendell, P. G., & Thomson, D. M. (1993). The effect of ageing on remembering to remember: An investigation of simulated medication regimens. *Australasian Journal on Ageing*, *12*(1), 11–18. https://doi.org/10.1111/j.1741-6612.1993.tb00578.x

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. https://doi.org/10.1093/geronb/54b.4.p256

Riley, R. D., Higgins, J. P. T., & Deeks, J. J. (2011). Interpretation of random effects meta-analyses. *BMJ*, *342*. https://doi.org/10.1136/bmj.d549

Schild, A. H. E., & Voracek, M. (2013). Less is less: A systematic review of graph use in meta-analyses. *Research Synthesis Methods*, *4*(3), 209–219. https://doi.org/10.1002/jrsm.1076

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. https://doi.org/10.1017/S1041610210001651

Schnitzspahn, K. M., & Kliegel, M. (2009). Age effects in prospective memory performance within older adults: The paradoxical impact of implementation intentions. *European Journal of Ageing*, *6*(2), 147–155. https://doi.org/10.1007/s10433-009-0116-x

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. https://doi.org/10.1007/s00426-018-1140-2

Schnitzspahn, K. M., Thorley, C., Phillips, L., Voigt, B., Threadgold, E., Hammond, E. R., Mustafa, B., & Kliegel, M. (2014). Mood impairs time-based prospective memory in young but not older adults: The mediating role of attentional control. *Psychology and Aging*, *29*(2), 264–270. https://doi.org/10.1037/a0036389

Shum, D. H. K., Cahill, A., Hohaus, L. C., O'Gorman, J. G., & Chan, R. C. K. (2013). Effects of aging, planning, and interruption on complex prospective memory. *Neuropsychological Rehabilitation*, *23*(1), 45–63. https://doi.org/10.1080/09602011.2012.716761

Smith-Spark, J. H., Zięcik, A. P., & Sterling, C. (2017). Adults with developmental dyslexia show selective impairments in time-based and self-initiated prospective memory: Self-report and clinical evidence. *Research in Developmental Disabilities*, *62*(2016), 247–258. https://doi.org/10.1016/j.ridd.2016.12.011

Sullivan, K. L., Woods, S. P., Bucks, R. S., Loft, S., & Weinborn, M. (2018). Intraindividual variability in neurocognitive performance is associated with time-based prospective memory in older adults. *Journal of Clinical and Experimental Neuropsychology*, 1–11. https://doi.org/10.1080/13803395.2018.1432571

Tracy, J. I., Faro, S. H., Mohamed, F. B., Pinsk, M., & Pinus, A. (2000). Functional localization of a "time keeper" function separate from attentional resources and task strategy. *NeuroImage*, *11*(3), 228–242. https://doi.org/10.1006/nimg.2000.0535

Van den Noortgate, W., López-López, J. A., Marín-Martínez, F., & & Sánchez-Meca, J. (2015). Meta-analysis of multiple outcomes: A multilevel approach. *Behavior Research Methods*, *47*(4), 1274–1294. https://doi.org/10.3758/s13428-014-0527-2

Vanneste, S., Baudouin, A., Bouazzaoui, B., & Taconnat, L. (2016). Age-related differences in time-based prospective memory: The role of time estimation in the clock monitoring strategy. *Memory*, *24*(6), 812–825. https://doi.org/10.1080/09658211.2015.1054837

Van Petten, C. (2004). Relationship between hippocampal volume and memory ability in healthy individuals across the lifespan: Review and meta-analysis. *Neuropsychologia*, *42*(10), 1394–1413. https://doi.org/10.1016/j.neuropsychologia.2004.04.006

van Rhee, H., Suurmond, R., & Hak, T. (2015). User manual for meta-essentials: Workbooks for meta-analysis. *Social Science Research Network*. https://doi.org/10.2139/ssrn.3241355

Varley, D., Henry, J. D., Gibson, E., Suddendorf, T., Rendell, P. G., & Redshaw, J. (2021). An old problem revisited: How sensitive is time-based prospective memory to age-related differences? *Psychology and Aging*, *36*(5), 616–625. https://doi.org/10.1037/pag0000625

Viechtbauer, W. (2005). Bias and efficiency of meta-analytic variance estimators in the random-effects model. *Journal of Educational and Behavioral Statistics*, *30*(3), 261–293. https://doi.org/10.3102/10769986030003261

Viechtbauer, W. (2010). Conducting meta-analyses in *R* with the **metafor** package. *Journal of Statistical Software*, *36*(3), 1–48. https://doi.org/10.18637/jss.v036.i03

Viechtbauer, W., & Cheung, M. W.-L. (2010). Outlier and influence diagnostics for meta-analysis. *Research Synthesis Methods*, *1*(2), 112–125. https://doi.org/10.1002/jrsm.11

Waldum, E. R., & McDaniel, M. A. (2016). Why are you late? Investigating the role of time management in time-based prospective memory. *Journal of Experimental Psychology: General*, *145*(8), 1049–1061. https://doi.org/10.1037/xge0000183

West, R. L. (1996). An application of prefrontal cortex function theory to cognitive aging. *Psychological Bulletin*, *120*(2), 272–292. https://doi.org/10.1037/0033-2909.120.2.272

Wojtyniak, J., Britz, H., Selzer, D., Schwab, M., & Lehr, T. (2020). Data digitizing: Accurate and precise data extraction for quantitative systems pharmacology and physiologically-based pharmacokinetic modeling. *CPT: Pharmacometrics & Systems Pharmacology*, *9* (6), 322–331. https://doi.org/10.1002/psp4.12511

Woods, S. P., Weinborn, M., Li, Y. R., Hodgson, E., Ng, A. R. J., & Bucks, R. S. (2015). Does prospective memory influence quality of life in community-dwelling older adults? *Aging, Neuropsychology, and Cognition*, *22*(6), 679–692. https://doi.org/10.1080/13825585.2015.1027651

Yang, T., Wang, Y., Lin, H., Zheng, L., & Chan, R. C. K. (2013). Impact of the aging process on event-, time-, and activity-based prospective memory. *PsyCh Journal*, *2*(1), 63–73. https://doi.org/10.1002/pchj.19

Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language*, *46*(3), 441–517. https://doi.org/10.1006/jmla.2002.2864

Yonelinas, A. P., Aly, M., Wang, W. C., & Koen, J. D. (2010). Recollection and familiarity: Examining controversial assumptions and new directions. *Hippocampus*, *20*(11), 1178–1194. https://doi.org/10.1002/hipo.20864

Zöllig, J., Martin, M., & Kliegel, M. (2010). Forming intentions successfully: Differential compensational mechanisms of adolescents and old adults. *Cortex*, *46*(4), 575–589. https://doi.org/10.1016/j.cortex.2009.09.010

Zuber, S., Haas, M., Framorando, D., Ballhausen, N., Gillioz, E., Künzi, M., & Kliegel, M. (2021). The Geneva space cruiser: A fully self-administered online tool to assess prospective memory across the adult lifespan. *Memory*, 1–16. https://doi.org/10.1080/09658211.2021.1995435

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