**Supplementary Material**

## Data reduction strategy

Participants did not always open the SAM at the exact time that it was sent. As we were interested in the affective state that participants were in when they had to remember their intentions, we used the following algorithm in order to include/exclude measurements. First, we excluded the (later) SAM if it was completed within 5 minutes of a previous SAM. SAMs were then automatically assigned to one of the three time windows based on when they were opened, rather than when they were sent. The same process was applied to the SAMs with the EBPM cues.

We kept a maximum of three SAMs per time window. If participants filled in more than three SAMs within each time window, the later ones were excluded. If participants filled in more than one SAM containing an EBPM task cue within one time window, only the first was considered. If participants did not answer any SAMs within a time window, those data points were counted as missing (for the affect and the EBPM measure). Finally, the affect ratings were averaged across the time windows, for each day and participant. The average affect ratings within each time-window were used to predict EBPM and TBPM performance within the respective time window.

## Variability in affect and PM performance in the naturalistic tasks.

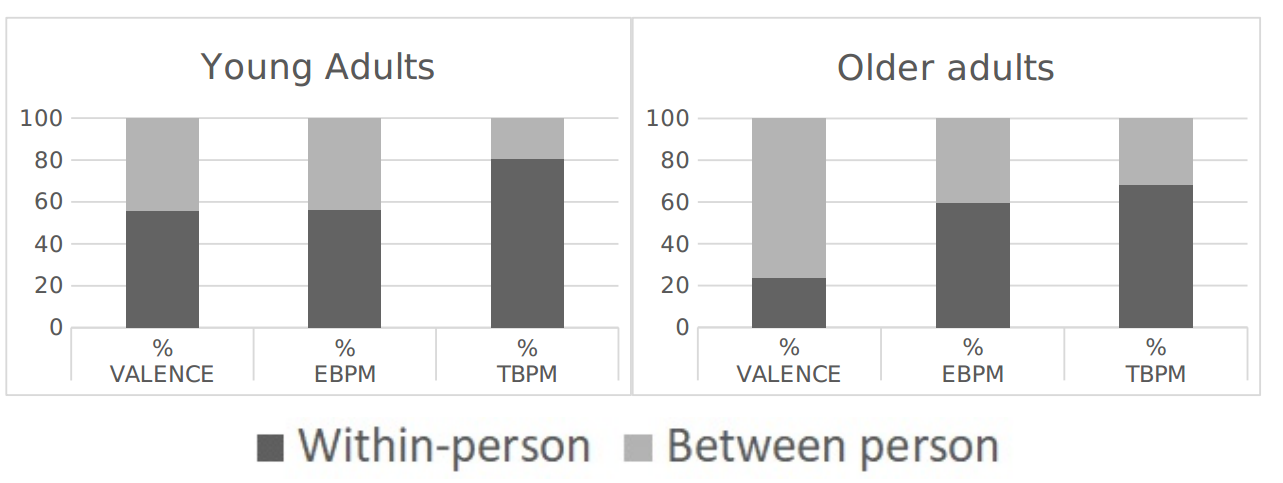
## Before performing the main analyses testing predictors of naturalistic PM performance, we explored the amount of variance in affect and PM performance that was within-participants and between-participants to ensure that there was sufficient variance in both independent and dependent variables within both age groups suggesting that our samples are representative. We calculated intraclass correlations using a null model, i.e. with no predictors (Singer & Willett, 2003). To explore possible age differences in the percentage of variance that was within- versus between-participants, observed percentages of variance within- and between-participants were compared with expected percentages using Pearson’s chi square test (Fisher, 1922). Odds ratios (OR) were calculated as measures of effect sizes, as the percentage of the variance within-participants divided by the percentage of the variance between participants.

Results are shown in Figure 1. There was a significant age difference between the amount of variation in affect that was within-participants, χ2(1)= 21.33, *p* < .001. Based on the ORs, the amount of variance in affect levels within-participants was 4.10 times higher in young than in older adults.

There was an age-related difference in the amount of variance within-participants for performance on the TBPM task, with young adults having 3.15 times more variation within participants compared to their older counterparts, χ2(1)=10.35, *p* = .001. The amount of variance in EBPM performance did not change as a function of age group, χ2(1)=0.262, *p* = .669.

*Figure 1.*

Variability in affect and PM performance for young and older adults.

******

## Predictors of PM performance in the naturalistic tasks

Table 1 shows the results of the backward elimination regression that was used to identify the control variables that contributed to explaining PM performance in the naturalistic tasks. The -2 log-likelihood (-2LL) of the model including all six predictors (full model) was compared with models in which each variable was in turn excluded. Since smaller -2LL corresponds to a better fit, a significant increase in -2LL in models which excluded a specific variable reflected a loss in explanatory power and thus an increase in unexplained information. The exclusion of task importance and busyness resulted in a significant decrease in -2LL, while the exclusion of stress did not significantly affect the -2LL. Exclusions of motivation, metacognitive awareness and use of strategies all resulted in significantly increased -2LL. Thus, these variables were used as predictors of PM performance in the final model.

Table 1.

*Results of the Backward Elimination Regression.*

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Variable |  | *df* |  | *-2LL* |  | χ2 |  | *p* |
| Full model |  | - |  | 9593.98 |  | - |  | - |
| Stressa |  | 1 |  | 9597.28 |  | 3.31 |  | .069 |
| Importancea |  | 1 |  | 9587.10 |  | -6.88 |  | .009 |
| Busynessa |  | 1 |  | 9578.24 |  | -15.73 |  | <.001 |
| Motivationa |  | 1 |  | 9600.66 |  | 6.69 |  | .010 |
| Metacognitive Awarenessb |  | 1 |  | 9799.180 |  | 205.20 |  | <.001 |
| Strategyc |  | 2 |  | 10092.33 |  | 498.35 |  | <.001 |

*aImportance, Motivation, Stress, and Business were self-assessed. Participants were asked to rate the extent to which they were stressed, busy, motivated, and found the task important, using a 5-point Likert scale, ranging from 1 (Totally agree) to 5 (Strongly disagree).*

*bMetacognitive awareness represents the differences between participants’ predictions and their actual performances.*

*cStrategy represents the average strategy use for EBPM and TBPM, where “1” means that participants used strategies for both EBPM and TBPM; “0.5” means that participants used strategies either for TBPM or for EBPM, and “0” means that participants used no strategy for both EBPM and TBPM.*

## Building the baseline model

The baseline model (Table 2) aims to predict PM in the naturalistic tasks, collapsing across event- and time-based tasks. In addition, it considers the amount of variance in PM performance that was within/between participants. It was built in two steps: In the first step (Model 1), an unconditional model allowed to estimate the amount of variance in PM performance that was within/between participants; in the second step (Model 2), the identified predictors of PM performance (i.e. motivation, metacognitive awareness, strategy use) were added to the model.

In Table 2 results of the GLMMs used to create the baseline model are shown, with estimated coefficients of fixed effects and variances for random effects. The results show that the model including strategy, metacognitive awareness, and motivation significantly reduced log-likelihood compared to the model without those predictors, and thus significantly reduced the amount of unexplained variability, thereby increasing the power to detect associations between affect and PM (Tabachnick et al., 2007).

Table 2.

Predictors of Naturalistic PM: Estimate Coefficients for Fixed Effects and Variances for Random Effects. PM Performance is Collapsed Across Event- and Time- Based Tasks.

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | Model1  (unconditional) | Model 2  (predictors) |
| *Fixed effects* |  |  |  |
| Composite Model | Intercept | 0.349\* | -2.502\*\*\* |
|  | Motivation |  | 0.764\*\*\* |
|  | Metacognitive awareness |  | -0.021\*\* |
|  | Strategya = 1 |  | 0.149 |
|  | Strategy= 0.5 |  | 0.479 |
| *Variance components* |  |  |  |
| Level 1 | Within-person residuals | 0.915\*\*\* | 0.941\*\*\* |
| Level-2 between-person | Intercept | 1.932\*\*\* | 0.741\*\*\* |
| Goodness of fit | -2 Log Likelihood | 10992.252\*\*\* | 9563.348\*\*\* |

*p < .05, \*\*p < .01, \*\*\*p < .001*

*aThe variable “strategy” represents the average strategy use between EBPM and TBPM, where “1” means that participants used strategies for both EBPM and TBPM; “0.5” means that participants used strategies either for TBPM or for EBPM, and “0” means that participants used no strategy for both EBPM and TBPM. The reference category is 0.*

*Note. The Level 1 Within-person residuals refer to unexplained within-participants variance in PM performance. The level-2 between-person intercept refers to the unexplained between-participants variance in PM performance.*



## Descriptive statistics on reminder use and exploratory analyses

Table 3 shows the number of participants who reported to have used strategies for EBPM and TBPM tasks, as a function of age group.

Table 4 shows the kind of strategies participants used.

Table3.

*Number of Young and Older Participants who Reported to Have Used Strategies in the EBPM and TBPM tasks.*

|  |  |  |
| --- | --- | --- |
|  | Task | |
| Age Group | EBPM | TBPM |
| Young | N = 6 (16%) | N=13 (35%) |
| Older | N = 10 (38%) | N = 27 (79%) |

Table 4.

*Type of Strategy Participants Reported to Have Used, as a Function of Age Group.*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Young Adults | | Older Adults | |
| Type of strategy | EBPM | TBPM | EBPM | TBPM |
| Asking someone else to remind | - | 1 | - | 1 |
| Memorising the times (TBPM) | - | - | - | 3 |
| Reading first sentence carefully (EBPM) | 2 | - | 6 | - |
| Rehearsing the intention offline | 4 | - | 2 | - |
| Setting alarms | - | 5 | - | 19 |
| Using post-it (either physical or digital) | - | 9 | 2 | 2 |
| Others | - | - | - | 2 |

**Exploratory analyses: Effect of Everyday Affect, Age, and Task Type on Naturalistic PM in participants who used external reminders vs in participants who did not use external reminders**

To explore the possibility that results of our main analysis differ for participants who used external reminders vs participants who did not use external reminders, we ran the full model separately for participants who did not report using any external reminder (YA = 25, OA = 12), and on the remaining participants who reported to have used some form of external reminder (YA = 12, OA = 22). In the first analysis focusing on participants who used external reminders, effects of age (*p* = .026) and task type (*p* < .001) remained, while the effect of valence measured within participants did not reach significance (*p* = .282). The same pattern of results emerged for participants who did not use any external reminder: There were still significant effects of age (*p* < .001) and task type (*p* < .001), but the effect of valence was not significant (*p* = .136). However, the statistical power to detect an effect for the within-participants valence variable was rather low in both analyses (1 – β = 0.52 and 0.44, respectively).