Journal of Applied Research in Memory and Cognition 8 (2019) 244–254

Contents lists available at ScienceDirect

Journal of Applied Research in Memory and Cognition

journal homepage: www.elsevier.com/locate/jarmac

Prospective Memory: Comparing Self- and Proxy-Reports with Cognitive Modeling of Task Performance☆

Nina R. Arnold*
University of Mannheim, Germany

Heinrich-Heine-Universität Düsseldorf, Germany

Ute J. Bayen
Heinrich-Heine-Universität Düsseldorf, Germany

Prospective memory (PM) refers to remembering to perform an action in the future and is crucial in everyday life. Self-report questionnaires are sometimes used to assess PM problems. In two studies, we compared self-ratings on the Prospective and Retrospective Memory Questionnaire (PRMQ, G. Smith, Della Sala, Logie, & Maylor, 2000) with actual performance in laboratory PM tasks using Bayesian hierarchical multinomial modeling. In Study 2, we additionally collected parents' ratings of high-school students via the PRMQ. Results indicate a relationship between parents' ratings and self-ratings of prospective and retrospective memory. There was, however, no relationship of any of the PRMQ measures with PM performance or model-based estimates of retrospective and prospective components of PM. The findings suggest that questionnaires not be used in lieu of performance measures of PM.

General Audience Summary

In everyday life, it is often important to remember to do something in the future. For example, you may have to remember to give a colleague a message when you see her. This type of memory is called prospective memory. People often have problems with prospective memory. Therefore, researchers have developed questionnaires that ask people how often they experience problems with prospective memory in daily life. The questionnaire that is used most often is the Prospective and Retrospective Memory Questionnaire (PRMQ). We used this questionnaire to ask parents how good they thought their teenage children's prospective memory was. We found that the parents' answers were quite similar to what the teenagers themselves thought of their own prospective memory. Prospective memory can also be measured in the laboratory where research participants have to perform a task at a computer and in addition have to remember, for example, to press a certain key on the keyboard when a particular word appears on the computer screen. In the research reported in this article, we had the same people fill out the questionnaire and perform a laboratory prospective-memory task. The answers on the questionnaire were then compared with actual performance in the laboratory task. It turned out that there was no relationship between the questionnaire answers and performance in the task. We conclude that the questionnaire and the laboratory task do not measure the same thing and that other instruments must be developed to measure problems with prospective memory.

Keywords: Prospective memory, MPT model, PRMQ, Questionnaire

Author Note.

Nina Rebecca Arnold, School of Social Sciences, University of Mannheim, Germany; Institute of Experimental Psychology, Heinrich-Heine-Universität Düsseldorf, Germany.

Ute Johanna Bayen, Institute of Experimental Psychology, Heinrich-Heine-Universität Düsseldorf, Germany.

The data can be accessed via the Open Science Framework at https://osf.io/3yph4/?view_only=c543125d614b4e849c6abfe1a0fef3ec.

* Correspondence concerning this article should be addressed to Nina R. Arnold, Department of Psychology, School of Social Sciences, University of Mannheim, L13 17, D-68161 Mannheim, Germany. Contact: arnold@uni-mannheim.de.

Prospective memory (PM) refers to remembering to perform an action in the future and is very important in everyday life (e.g., Brandimonte, Einstein, & McDaniel, 1996). Examples of PM are remembering to take medicine at a certain point in time (i.e., time-based PM) or after a meal (i.e., event-based PM) as well as remembering to make a phone call, to turn off the oven, give someone a message, and so on. PM failures in daily life may have serious consequences; hence, the assessment of PM performance is of paramount importance. While efforts have been made to develop standardized tests of PM (Kamat et al., 2014; Wilson et al., 2005), questionnaire self-ratings are the most frequently used method for the assessment of memory problems, especially in clinical contexts (Thöne-Otto & Walter, 2008). However, while self-reports may adequately reflect perceptions of one's own PM performance, they may not yield valid assessments of actual memory performance. To increase the validity of PM assessments, the use of proxy ratings by relatives has been suggested (Thöne-Otto & Walter, 2008). In the current study, we compared self and proxy versions of the most frequently used PM questionnaire with PM performance measures derived from the most commonly used PM laboratory task. Specifically, we compared the Prospective and Retrospective Memory Questionnaire (PRMQ; G. Smith et al., 2000) with the standard laboratory PM task (as introduced by Einstein & McDaniel, 1990) combined with cognitive modeling (R.E. Smith & Bayen, 2004). If both types of measurement capture, at least in part, the same construct, there should be a correlation.

Self-Report of Prospective and Retrospective Memory via the PRMQ

There are several questionnaires available for the self-report of PM and PM failures (for an overview, see Thöne-Otto & Walter, 2008). Among these, we chose the PRMQ for our studies, because it is the most frequently used self-report measure of PM. The PRMQ addresses PM as well as retrospective memory (memory for information that was encountered in the past) as both are important in daily life. The PRMQ is easy to administer and consists of 16 items that ask about failure in prospective memory (8 items) and retrospective memory (8 items) in everyday life. Participants indicate how often they experience specific memory failures on a 5-point Likert scale from *very often* to *never*. Crawford, Smith, Maylor, Sala, and Logie (2003) tested construct validity of the PRMQ via confirmatory factor analysis and found a model with a common factor plus specific prospective and retrospective factors. The reliabilities (Cronbach's alpha) of the self-report total score, the prospective scale, and the retrospective scale were high (*rs* of .89, .84, and .80, respectively), as were the reliabilities of proxy PRMQ ratings (*rs* of .92, 87, and .83, respectively; Crawford, Henry, Ward, & Blake, 2006). The authors recommended the questionnaire for research and as a screening tool for health practitioners.

In recent years, the PRMQ has been used in a host of studies in order to assess individual differences in reports of prospective and retrospective memory and their association with health issues (e.g., Avgan et al., 2014; Bigdeli, Farzin, & Talepasand, 2014; Costa et al., 2015; Chan et al., 2012; Cuttler, Alcolado, & Taylor, 2013; Cuttler, Relkov, & Taylor, 2013; Fernie, Bennett, Currie, Perrin, & Reid, 2014; Ling, Campbell, Heffernan, & Greenough, 2007; Hsu, Huang, Tu, & Hua, 2014; Mohammadi, Keshavarzi, & Talepasand, 2014; Ross, Macdiarmid, Rostron, Watt, & Crawford, 2013; G. Smith et al., 2000; Thelen, Lynch, Bruce, Hancock, & Bruce, 2014; Wamsley, Donjacour, Scammell, Lammers, & Stickgold, 2014; Willians, Jarrold, Grainger, & Lind, 2014), aging (e.g., Debarnot et al., 2015), or substance use (e.g., Downey et al., 2015; Heffernan & O'Neill, 2014; Weinborn et al., 2013). Some of these studies assume that the measures derived from the PRMQ are valid measures of prospective memory and retrospective memory. However, this assumption may not hold as research on metamemory has shown that people have difficulties judging their own retrospective-memory performance (e.g., Hultsch, Hertzog, & Dixon, 1990).

For PM, the results are less clear. Dobbs and Rule (1987) compared self-reported PM ability and actual PM performance in real life and found that self-ratings and performance were often negatively correlated. There are only three published studies thus far on the relationship between PM problems self-reported on the PRMQ and performance on PM laboratory tasks (for a summary, see Table 1). The first of these was published by Mäntylä (2003) who compared middle-aged adults who complained about PM problems in response to a newspaper advertisement with noncomplainers. Complainers performed significantly worse on all PM tasks, but there were no differences in retrospective memory tasks. Differences between the two groups in the PRMQ were much larger on the prospective scale than on the retrospective scale. Kliegel and Jäger (2006) investigated whether the PRMQ predicted PM performance and found that time-based PM performance was predicted by the PM subscale but not by the retrospective-memory subscale. The correlation between event-based PM performance in the laboratory task and the prospective subscale of the PRMQ did not reach significance, but showed a small effect (PM subscale and time-based: $r = -.22$, $p < .050$; PM subscale and event-based: $r = -.19$, $p = .081$; no correlations for PRMQ total score and the retrospective subscale). Finally, Zeintl, Kliegel, Rast, and Zimprich (2006) screened older participants according to their memory complaints assessed with the PRMQ. For high complainers, prospective memory complaints were significantly associated with depressive symptoms and self-reported memory capacity but not with PM performance in a laboratory task. For low complainers, PM complaints were significantly correlated solely with PM performance. Due to these partially contradictory findings, we designed our studies to contribute a more fine-grained analysis of the relationship between PRMQ-based measures and different components of PM performance. Performance in PM tasks generally depends on two components (Einstein & McDaniel, 1990): (a) the prospective component—remembering *that* one must do something, and (b) the retrospective component—remembering *what* action to perform and *when* to perform it. For example, an employee may intend to give a message to a colleague when she sees him. She may later remember that she wanted to give someone a message (prospective component), but not the full contents of the message or which colleague to give it to (retrospective component).

Table 1
Prior Studies on the Relationship between PM Performance and PRMQ Self-ratings

Study	Population	PM task	Correlation of PRMQ prospective scale and PM performance
Mäntylä (2003), Experiment 2	200 adults (35–55 years)	Remind the experimenter to sign a piece of paper when the session is completed (1 target event)	$r = -.20$, $p < .01$
Kliegel and Jäger (2006)	87 adults ($M_{age} = 44.11$ years)	Remember to press a target key whenever a picture in an n-back task shows an animal	$r = -.19$, n.s.
Zeintl et al. (2006)	364 older adults ($M_{age} = 73.0$ years)	Repeat the words “red pencil” whenever the experimenter mentions these words during the session.	• Total sample: $r = .09$, n.s. • High complainers (according to median split in prospective PRMQ scores; $n = 197$): $r = .08$, n.s. • Low complainers ($n = 167$): $r = .21$, $p < .001$

Note. PRMQ: Prospective and Retrospective Memory Questionnaire (PRMQ, G. Smith et al., 2000). We only considered studies that report associations of PRMQ scores and event-based PM performance measures. Higher PRMQ scores indicate more self-reported PM failures (i.e., poorer self-reported memory). Higher PM performance scores indicate more PM hits (i.e., better memory).

As the PRMQ seeks to separate prospective from retrospective memory failures, we wanted to compare PRMQ measures with laboratory-based measures of the prospective and retrospective components of PM. Note that the items of the prospective scale of the PRMQ refer to PM tasks that involve both prospective and retrospective components of PM (as is usually the case in everyday PM tasks) with items such as “Do you fail to mention or give something to a visitor that you were asked to pass on?”, a situation that requires remembering that something needs to be done (prospective component), but also what to pass on and to which visitor (retrospective component). The retrospective scale of the PRMQ, by contrast, is a self-report of pure retrospective memory: Its items ask about recognition memory (e.g., “Do you fail to recognize a place you have visited before?”), recall from memory and (in one case) source memory. Therefore, the retrospective scale of the PRMQ should correlate with the retrospective component of a laboratory task only, whereas the prospective scale of the PRMQ should correlate with both the prospective and the retrospective component as well as with an overall PM performance measure.

Assessing Components of PM with Laboratory Tasks

To test these predictions, we need separate measures of the prospective and the retrospective components of a laboratory PM task. We will first describe the standard laboratory paradigm that we used, then describe how separate measures of both components can be obtained from the data with the help of a mathematical model.

The Standard Laboratory PM Task

Since, in daily life, a PM action often interrupts an ongoing activity (e.g., stopping to read a book to take medicine), standard PM laboratory tasks are also embedded in an ongoing activity. The ongoing activity is usually a computer-based task, such as a lexical-decision task. In event-based PM tasks, the appropriate PM action must be carried out in response to specific PM target events that appear during the ongoing task (Einstein & McDaniel, 1990). For example, the PM response must be given in response to certain initial letters of the words and non-words that appear during the lexical-decision task. The PM target events appear only rarely (e.g., in 10% of the ongoing-task trials). Overall PM performance is usually measured as PM hit rate, that is the proportion of PM target trials correctly responded to. This measure is a conglomerate of prospective and retrospective task components. That is, to achieve a PM hit participants must remember that they need to perform a PM task in addition to the ongoing task (prospective component). They must also discriminate PM target events from distractor events (retrospective recognition component), and which key to press upon occurrence of a target event. Therefore, we need separate measures to assess the prospective and the retrospective task components. To obtain these, we used a MPT model of PM.

The Multinomial Model of PM

Multinomial processing tree (MPT) models are a class of measurement models for the estimation of probabilities of latent cognitive processes that underlie performance in specific cognitive tasks (for an introduction to MPT models, see Batchelder & Riefer, 1999). A model specifically developed to provide separate and unconfounded measurements of the prospective and retrospective components of PM is the MPT model of event-based PM by R.E. Smith and Bayen (2004). This model was designed for laboratory event-based PM tasks that are embedded in an ongoing binary classification task. The ongoing task may, for example, be a lexical decision task with the stimulus classes *word* and *non-word* (as in our Study 2). In ongoing lexical-decision tasks, participants have two response options, namely class 1 (“word”) and class 2 (“non-word”). PM targets occur on trials with either stimulus class. Upon occurrence of a PM target, participants are supposed to press a designated key as the PM response instead of responding to the ongoing task.

Refer to Figure 1 for an illustration of the MPT model. The four trees of the model represent the four possible trial types of the laboratory task: Class 1 of the ongoing task occurs with a PM target, Class 2 of the ongoing task occurs with a PM target, Class 1 of the ongoing task occurs without a PM target, and Class 2 of the ongoing task occurs without a PM target. The

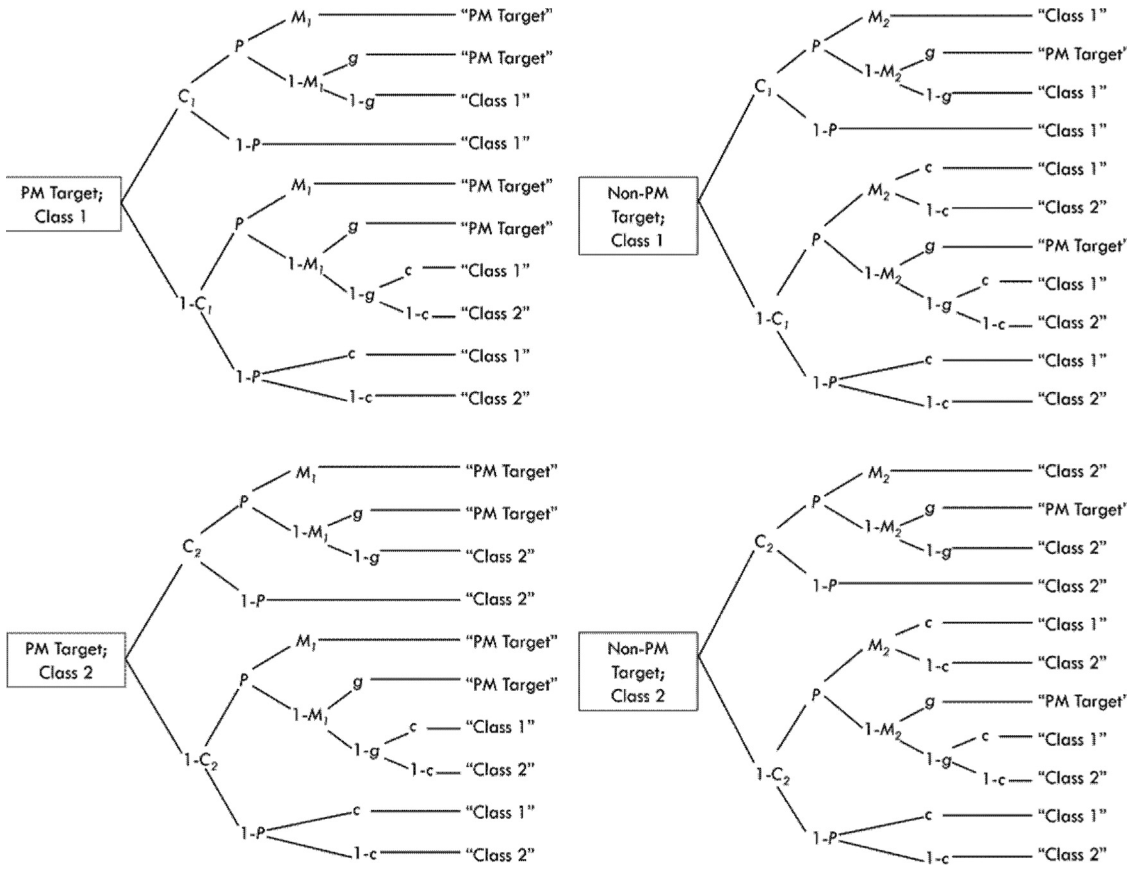

Figure 1. The multinomial model of event-based PM. PM, prospective memory; C_1, probability of detecting Class 1 in the ongoing task; C_2, probability of detecting Class 2 in the ongoing task; P, prospective component; M_1, probability of recognizing PM targets; M_2, probability of noticing that a stimulus is not a PM target; g, probability of guessing that a stimulus is a PM target; c, probability of guessing that the answer to the ongoing task is Class 1.
Adapted from "A multinomial model of event-based prospective memory" by R.E. Smith and U.J. Bayen, 2004, *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*, p. 758.

ability to perform the ongoing task is captured by Parameters C_1 and C_2. The prospective component is measured by parameter P, that is, the probability that the participant remembers that there is an additional PM task. On PM target trials, M_1 indicates the probability that a participant successfully recognizes a PM target, which results in a PM response. M_2 indicates the probability that a participant recognizes that an event is not a PM target resulting in an ongoing-task response. M_1 and M_2 thus represent the retrospective component in the model capturing discrimination of PM targets and non-targets (i.e., *when* to perform the action). Note that M parameters of the MPT model capture the *when* of retrospective memory, not the *what*. If recognition whether an item is a PM target or not fails (with probabilities $1 - M_1$ or $1 - M_2$, respectively), participants must either guess that the trial includes a PM target (with probability g), or that the trial does not include a PM target ($1 - g$). If participants do not remember the PM task ($1 - P$), they respond to the ongoing task. Parameter c indicates the probability of guessing Class 1, and $1 - c$ is the probability of guessing Class 2. To obtain mathematical identifiability of the model, some parameters are restricted based on theoretical assumptions (cf. R.E. Smith & Bayen, 2004): Parameter M ($=M_1 = M_2$) measures discrimination of PM targets and non-targets. The guessing parameters (c and g) are set equal to the ratios of ongoing-task items and PM targets, respectively. The resulting model with four free parameters (P, M, C_1, C_2) has been validated (Horn, Bayen, Smith, & Boywitt, 2011; R.E. Smith & Bayen, 2004). It has been successfully used to separately measure the prospective component (Parameter P) and the retrospective component (Parameter M) in a number of studies with various populations and task conditions (Pavawalla, Schmitter-Edgecombe, & Smith, 2012; Schnitzspahn, Horn, Bayen, & Kliegel, 2012; R.E. Smith & Bayen, 2005, 2006; R.E. Smith, Bayen, & Martin, 2010; Walter & Bayen, 2016).

The Current Research

In each of two studies, we administered both the PRMQ and a standard laboratory event-based PM task to the same participants. We correlated the prospective and retrospective scales of the PRMQ with the MPT model-based measures of the prospective and retrospective components of PM in the laboratory task as well as with overall PM performance (i.e., PM hit rate). The prospective and the retrospective component of PM can counteract each other in their effects on overall PM performance (e.g., R.E. Smith & Bayen, 2004). Therefore, if there are no correlations of the PRMQ with PM performance, this does not exclude correlations with the prospective and the retrospective component (see also Walter & Bayen, 2016). Thus, we correlated the PRMQ with PM hit rate as well as the two components. Note that parameter *M* of the MPT model captures the *when* component of retrospective memory, not the *what* component, whereas the retrospective scale of the PRMQ includes both *when* and *what*.

Study 1 is a reanalysis of the PM data from Arnold, Bayen, and Böhm (2015) to explore a possible correlation of laboratory-based measures with questionnaire data. In Study 2, we sought to replicate the results with a different sample and also added a proxy-report of PM to assess its relation with the self-report and the laboratory-based measures.

Study 1

Part of the data were included in the study by Arnold et al. (2015). They presented the relationship between MPT-model based estimates of PM with self-reports of depression and anxiety. In the current study, we present the relationship of MPT-model based estimates with data obtained with the PRMQ questionnaire. For complete methodological details, refer to Arnold et al.

Participants

One-hundred and thirty students from the Heinrich-Heine-Universität Düsseldorf (all native German speakers) participated. We excluded the data of five participants from analyses, because one never gave a "PM" response in the PM task suggesting that she had not understood the task, and four had missing values on the PRMQ. The final sample consisted of 76 females and 49 males ($M_{\text{age}} = 22.37$ years, range 18–52).

Procedure and Materials

Participants first performed the computerized PM task. The ongoing task was a color matching task with the *M* and *V* keys assigned to the answers *match* and *nonmatch*. One trial consisted of four colored rectangles presented in the middle of a black screen for 500 ms. After the fourth rectangle, a colored word appeared, and participants judged whether or not the color of the word matched one of the colors of the previously presented rectangles. The PM task was to press the space bar (instead of the match or nonmatch key) whenever one of the PM target words appeared.

Word items for the PM task were 168 German words of which fifteen served as PM targets and the others as distractors. Each item was presented twice. Thus, there were 336 trials in total, of which 30 were PM trials. After the PM task, participants filled out several paper-pencil questionnaires. Among these was the PRMQ (German version in paper-pencil format; Kaschel, 2002).[1]

Results and Discussion

The data of both studies can be accessed at https://osf.io/3yph4/?view_only=c543125d614b4e849c6abfe1a0fef3ec.

PRMQ. The total sum scores on the PRMQ can range from 16 to 80 (with higher scores indicating better self-reported PM performance), and the scores for each subscale from 8 to 40. In this study, total scores ranged from 29 to 72 ($M = 56.95$, $SD = 7.49$) with scores on the prospective PRMQ subscale ranging from 10 to 37 ($M = 27.01$, $SD = 4.44$), and on the retrospective subscale from 19 to 38 ($M = 29.94$, $SD = 4.02$). The correlation between the two subscales was .56 (a medium correlation according to Cohen, 1969).

Performance in the laboratory task. We measured PM performance in the laboratory task as PM hit rate, that is, the percentage of PM target items correctly responded to. Participants correctly responded on 69.57% of the PM trials and 89.59% of the ongoing-task trials. If they corrected themselves, the corrected response was used. Participants gave false alarms on a total of 313 ongoing-task trials (0.82%).

Correlations between PRMQ measures and PM performance. PM hit rate did not correlate with PRMQ total score, $r = -.004$, Bayes factor (BF) = 14.11 in favor of the null hypothesis. Neither did PM hit rate correlate with the PRMQ prospective subscale, $r = .003$, $BF = 14.11$, nor with the retrospective subscale, $r = -.012$, $BF = 14.01$.

Cognitive modeling. The MPT model of PM (R.E. Smith & Bayen, 2004) as described in the introduction has been applied in several studies. In most of these studies, the data were aggregated over participants in order to yield sufficient numbers of observations per response category. In recent years, there has been rising awareness that the traditional method of applying MPT models to aggregated data is limited (e.g., Klauer, 2006, 2010; Matzke, Dolan, Batchelder, & Wagenmakers, 2015; J.B. Smith & Batchelder, 2008, 2010). Hierarchical Bayesian extensions have been developed to account for participant heterogeneity by assuming that the individual parameters follow a continuous hierarchical distribution (e.g., Heck, Arnold, & Arnold, 2018; Klauer, 2010; Smith & Batchelder, 2010). These approaches allow us to include predictors in the MPT model and, thus, relate individual questionnaire data to individual MPT model parameters (e.g., Arnold et al., 2015; Schaper, Kuhlmann, & Bayen, 2018).

We calculated MPT model parameters using the *R* package TreeBUGS (Heck et al., 2018). We used the prospective and retrospective subscales of the PRMQ as continuous predictors of the estimated latent-trait hierarchical MPT parameters of the prospective component *P* and the retrospective component *M*, respectively. Note that we used a different parameter estimation method than we used in Arnold et al. (2015). The latent-trait approach (Klauer, 2010) used in the current study, in contrast to the beta-MPT approach we used in the earlier study, has the

[1] The following questionnaires were administered prior to the PRMQ: a questionnaire on caffeine consumption, the short version of a questionnaire investigating achievement motivation (Leistungsmotivationsinventar-kurz; Schuler & Prochaska, 2001), and German versions of the following tests: Fagerström Test for Nicotine Dependence (Bleich, Havemann-Reinecke, & Kornhuber, 2002), the Epworth Sleepiness Scale (Bloch, Schoch, Zhang, & Russi, 1999), and the Karolinska Sleepiness Scale (Åkerstedt & Gillberg, 1990).

Table 2
Posterior Distributions of the Parameters of the Group-Level Parameter Estimates of Study 1

	Mean [95% BCI]	*SD* [95% BCI]
C_1	.75 [.71, .79]	0.72 [0.62, 0.83]
C_2	.91 [.90, .92]	0.41 [0.35, 0.49]
P	.82 [.77, .86]	0.75 [0.62, 0.90]
M	.93 [.91, .95]	0.69 [0.57, 0.83]

Note. C_1, probability of detecting a color match; C_2, probability of detecting a color nonmatch; *P*, prospective component of PM; *M*, retrospective component of PM; BCI, Bayesian confidence interval; *Mean*, mean of the hierarchical distribution; *SD*, standard deviation of latent-trait values (probit scale) across individuals.

advantage that predictors can be directly included as regression weights. We further restricted the guessing parameters *c* and *g* to .5 and .10, respectively (i.e., equal to the ratios of ongoing-task items and PM targets, respectively, as explained in the introduction). The algorithm was run with 60,000 iterations using the first 10,000 iterations removed as burn-in period. For all parameters, the potential scale reduction factor was <1.05, indicating good convergence. Group means of the parameter estimates are presented in Table 2.

Continuous predictors are implemented in TreeBUGS by a linear regression on the probit scale. A positive regression weight indicates a higher probability whereas a negative regression weight indicates a lower probability that a cognitive process occurs as the predictor increases. We used the TreeBUGS standard priors, that is, weakly informative, multivariate Cauchy priors, for estimation. For the relationship between the PRMQ prospective subscale and the prospective component *P*, the regression weight was 0.01, 95% BCI [-0.02, 0.04]. For the relationship between the PRMQ prospective subscale and the retrospective component *M*, the regression weight was 0.01, 95% BCI [-0.05, 0.02]. For the relationship between the PRMQ retrospective subscale and the retrospective component *M*, the regression weight was -0.01, 95% BCI [-0.04, 0.02] Thus, both predictors clearly do not differ from zero. We used Bayesian statistics to weigh the null and alternative hypotheses against each other. The Bayes factor (BF) indicates the likelihood of the data under one hypothesis as compared to the other: Here, the BF indicates the likelihood of the null hypothesis compared to the alternative hypothesis (e.g., Rouder, Morey, & Pratte, 2017). A *BF* between 1 and 3 can be interpreted as “weak evidence,” a *BF* between 3 and 20 as “positive evidence,” a *BF* between 20 and 150 as “strong evidence,” and a *BF* larger than 150 as “very strong evidence” according to Raftery (1995, p. 139). The *BF* for *P* was 14.68, supporting the null hypothesis that the PRMQ prospective subscale was not associated with the prospective component *P*; the *BF* for *M* was 9.99, supporting the null hypothesis that the PRMQ prospective subscale was not associated with the retrospective component *M*. The *BF* for *M* was 14.52, supporting the null hypothesis that the PRMQ retrospective subscale was not associated with the retrospective component *M*. Thus, in Study 1, we did not find a relationship between PRMQ subscales and PM in the laboratory as measured by MPT model parameters. We also calculated split-half reliabilities (odd vs. even trials) for the parameters *M* and *P*. Both reliabilities were very high, $Rel(P) > .99$, $Rel(M) > .99$. However, the parameter for the retrospective component *M* was close to ceiling in Study 1 (with a mean estimate of .93, see Table 2) which may have obscured correlations. We addressed this issue in Study 2.

Study 2

Study 2 had two objectives. First, we sought to replicate Study 1 with a different sample and a different task. To avoid a ceiling effect on parameter *M*, we chose a task that would yield lower estimates of *M* than the ones obtained in Study 1. Specifically, in Study 2, we used a lexical-decision task similar to the task that yielded estimates of *M* of about .8 in the study by R.E. Smith, Persyn, and Butler (2011).

Second, we included proxy-ratings. Crawford et al. (2006) found that proxy-ratings might be more useful for the assessments of PM problems because they often correlate more strongly with objective assessments than self-ratings do (Sunderland, Harris, & Baddeley, 1988). Therefore, in Study 2, we also investigated the relationship between proxy-ratings and laboratory assessment of PM, as well as the relationship between self-ratings and proxy-ratings.

High-school students filled out the self-report version of the PRMQ and participated in a laboratory PM task with an ongoing lexical-decision task. Their parents filled out the proxy version of the PRMQ.

Participants

A total of 72 students of two different high schools ($n = 41$ and 31) in Düsseldorf and their parents participated. The students received 10 € for participation. We excluded 6 students who never pressed the PM key, and 10 students with missing values in the PRMQ either in the self-ratings or the parents’ ratings, leaving 56 students in the sample.[2] These students were aged between 16 and 19 years ($M = 17.5$); 75.8% were female. Parents who filled out the parent questionnaire were 91.1% female.

Procedure and Materials

PRMQ. For the students, we again used the German version of the PRMQ by Kaschel (2002). For parents, we rephrased all items of the PRMQ so that they asked about memory failures in everyday life for their son or daughter.

Procedure. Students were required to bring the completed parents’ questionnaire to the university laboratory. They filled out informed consent. Students below the age of 18 years also had to bring consent signed by a parent. They were tested in groups of up to six in separate computer booths. First, they filled out the paper-pencil PRMQ. After that, all instructions were computer based.

The ongoing task in Study 2 was a lexical decision task with the *R* and *I* keys assigned to the response options *word* and *non-*

[2] Including all possible data points per analysis did not change the pattern of the results.

word (counterbalanced). Items for the task were words from the CELEX database (Baayen, Piepenbrock, & Gulikers, 1995) between 8 and 12 letters in length with a frequency smaller than one per million. Non-words were created by changing the vowels. Altogether, 484 words and 484 non-words were presented. Responses were self-paced. Both speed and accuracy were emphasized in the instructions. After 48 ongoing-task practice trials, the PM task was introduced which was to press the space bar instead of the *word* or *nonword* key whenever one of the PM targets appeared. PM targets were letter strings that started with one of the letters G, H, or M. Sixteen PM trials appeared as word strings and 16 as nonword strings. The task was divided in eight blocks with 125 trials containing four PM targets each. There were 60 s breaks between blocks.

Results and Discussion

PRMQ. For the self-ratings, PRMQ total scores ranged from 45 to 74 ($M = 60.66$, $SD = 7.82$) with the prospective subscale ranging from 20 to 37 ($M = 29.12$, $SD = 4.28$), and the retrospective subscale from 20 to 39 ($M = 31.53$, $SD = 4.35$). The correlation between the two subscales was .64 (a large correlation according to Cohen, 1969).

For the proxy-ratings, PRMQ total scores ranged from 34 to 78 ($M = 66.05$, $SD = 8.29$) with the prospective subscale ranging from 18 to 40 ($M = 31.63$, $SD = 4.86$), and the retrospective subscale from 16 to 40 ($M = 34.43$, $SD = 3.99$). The correlation between the two subscales was $r = .75$ (a large correlation according to Cohen, 1969).

Correlations between the self- and the proxy-ratings were $r = .29$ for the prospective subscale, $r = .57$ for the retrospective subscale, and $r = .42$ for the total score. These correlations reflect medium to large effects sizes.

Performance in the laboratory task. Participants correctly responded to 48.80% of the PM trials and 91.05% of the ongoing-task trials. Again, if they corrected themselves the corrected response was used. Participants gave false alarms on a total of 367 ongoing-task trials (0.81%).

Correlations between PRMQ measures and PM performance. PM hit rate did not correlate with PRMQ total score, $r = .163$, $BF = 4.82$ in favor of the null hypothesis. Further, PM hit rate neither correlated with the PRMQ prospective subscale, $r = .121$, $BF = 6.32$, nor with the retrospective subscale, $r = .174$, $BF = 4.43$. For the parents' ratings the pattern was similar: PRMQ total score $r = .027$, $BF = 9.36$, prospective subscale $r = -.038$, $BF = 9.18$, retrospective subscale $r = .107$, $BF = 7.24$.

Cognitive modeling. Again, we calculated MPT model parameters using the *R* package TreeBUGS (Heck et al., 2018). We calculated two models, one for the self-ratings and one for the parents' ratings. First, we used the prospective and retrospective subscales of the PRMQ self-ratings as continuous predictors of the estimated latent-trait hierarchical MPT parameters of the prospective component *P* and the retrospective component *M* of the laboratory task, respectively. Based on the results of Study 1, we expected small effects in Study 2. Therefore, we used the information gained from Study 1 to set more informative priors for the regression weights in Study 2; specifically, we set the hyperprior on the precision (i.e., the inverse of the variance) of the slope parameters for the *z*-standardized PRMQ subscales to IVprec = "dgamma(.5,.5)" (see Heck et al., 2018). Second, we used the prospective and retrospective subscales of the PRMQ parents' ratings as continuous predictors of the estimated latent-trait hierarchical MPT parameters of the prospective component *P* and the retrospective component *M* of their children's laboratory task, respectively. Here, we used the weakly informative, multivariate Cauchy priors as in Study 1. Guessing parameters *c* and *g* were restricted to .50 and .03, respectively (again corresponding to the proportions of color matches and PM target occurrences in the task). Further restrictions and specifications were the same as in Study 1. For all parameters, the potential scale reduction factor was $R < 1.05$, indicating good convergence. Group parameter estimates (based on the estimation with self-ratings)[3] are presented in Table 3. As shown, we succeeded in obtaining an estimate of *M* well below ceiling (.78).

Table 3
Posterior Distributions of the Parameters of the Group-Level Parameter Estimates of Study 2

	Mean [95% BCI]	*SD* [95% BCI]
C_1	.76 [.71, .83]	0.64 [0.51, 0.82]
C_2	.93 [.89, .96]	0.83 [0.66, 1.07]
P	.70 [.58, .81]	1.07 [0.79, 1.44]
M	.78 [.68, .87]	1.05 [0.78, 1.45]

Note. C_1, probability of detecting a color-match; C_2, probability of detecting a color-nonmatch; *P*, prospective component of PM; *M*, retrospective component of PM; BCI, Bayesian confidence intervals; *Mean*, mean of the hierarchical distribution; *SD*, standard deviation of latent-trait values (probit scale) across individuals.

For the self-ratings, the regression weight indicating the relationship between the PRMQ prospective subscale and the parameter estimate for the prospective component *P* was 0.02, 95% BCI [-0.05, 0.08]. For the relationship between the PRMQ prospective subscale and the retrospective component *M*, the regression weight was 0.03, 95% BCI [-0.05, 0.02]. The regression weight indicating the relationship between the PRMQ retrospective subscale and the parameter estimate for the retrospective component *M* was 0.03, 95% BCI [-0.03, 0.09]. Thus, again, both regression weights clearly did not differ from zero. For the calculation of the BFs, we had to adjust the priors and used weakly informative, multivariate Cauchy priors. The *BF* for *P* was 7.10, supporting the null hypothesis that the PRMQ prospective subscale was not associated with the prospective component *P*; the *BF* for *M* was 3.92, providing weak evidence for the null hypothesis that the PRMQ prospective subscale was not associated with the retrospective component *M*. The *BF* for *M* was 15.52, supporting the null hypothesis that the PRMQ retrospective subscale was not associated with the retrospective component *M*.

For the parents' ratings, the regression weight indicating the relationship between the PRMQ prospective subscale and the

[3] Parameters based on the estimation with parents' ratings differ only slightly, and the overall pattern and all conclusions are the same.

parameter for the prospective component *P* was -0.05, 95% BCI [-0.12, 0.01]. For the relationship between the PRMQ prospective subscale and the retrospective component *M*, the regression weight was -0.01, 95% BCI [-0.09, 0.09]. The regression weight indicating the relationship between the PRMQ retrospective subscale and the parameter for the retrospective component *M* was 0.01, 95% BCI [-0.05, 0.09]. Thus, for the parents' ratings, both regression weights also did not differ from zero. The *BF* for *P* was 2.28, providing weak evidence for the null hypothesis that the PRMQ prospective subscale was not associated with the prospective component *P*; the *BF* for *M* was 7.97, supporting the null hypothesis that the PRMQ prospective subscale was not associated with the retrospective component *M*. The *BF* for *M* was 7.31, supporting the null hypothesis that the PRMQ retrospective subscale was not associated with the retrospective component *M*. Again, split-half reliabilities for the two parameters *P* and *M* were very high, $Rel(P) = .91$, $Rel(M) = .94$.

General Discussion

To address the important issue of measuring PM problems in daily life, the use of questionnaires for self-assessment and proxy assessment has been suggested. The validity of self-reports as measures of actual performance is, however, questionable. We set out to determine the relationship of self-reports and proxy reports with the PRMQ, which is the most frequently used self-report measure of PM, with PM performance and mathematical model-based measures obtained with the standard laboratory PM paradigm. In two studies, we correlated the PRMQ total score as well as prospective and retrospective scales of the PRMQ with PM performance and with MPT model-based measures of the prospective and retrospective components of PM in laboratory tasks. In Study 2, we also added PRMQ proxy-reports. In neither study was there a relationship between the PRMQ scores and laboratory measures of PM. Bayesian analyses yielded positive evidence for these null findings. Self- and proxy-ratings were highly correlated indicating convergent validity of the questionnaire measures.

One possible explanation for the lack of correlations between questionnaire variables and performance-based variables is that self-reports are generally inadequate proxies for actual performance due to demand characteristics such as social desirability. In clinical assessments, this issue may be particularly pronounced because patients being evaluated for memory problems may be in denial and may fear for their independence in case of negative evaluations. In addition, patients with severe memory problems may not remember their own everyday memory failures well enough to indicate them on a questionnaire. Our research shows that even in young healthy adults who were assured anonymity and did not have to fear negative consequences from admission of memory problems, questionnaire data did not relate to actual performance. In addition to limitations of questionnaire data, another contributor to the null correlations may be limited ecological validity of the laboratory task (for a discussion of ecological validity of PM tasks, see Phillips, Henry, & Martin, 2008). Unsworth, Brewer, and Spillers (2012) found that generally speaking, the ecological validity of laboratory cognitive ability measures is very high. However, for PM, they concluded in line with Phillips et al. (2008) that laboratory-based PM tasks may not have the best ecological validity. More naturalistic assessments of PM performance such as those obtained with the virtual-week task (Rendell & Craik, 2000) might possibly yield higher correlations with PRMQ measures.

We chose the standard laboratory paradigm introduced by Einstein and McDaniel (1990), because it allows us, in combination with MPT modeling, to disentangle prospective and retrospective components of PM. This is important as the PRMQ is designed with a similar objective: to separately measure PM and retrospective memory. The prospective scale of the PRMQ, however, measures self-report of PM performance, which is a conglomerate of prospective and retrospective components. This confound of the prospective scale may decrease its correlation with the model-based parameter *P*, which is a pure measure of the prospective component. However, the PRMQ prospective scale did not correlate with *M*, the model-based measure of retrospective memory, either, nor with PM hits, the laboratory-based overall performance measure. Although the latter correlation was a bit stronger, all *BF*s favored the null hypotheses.

The retrospective scale of the PRMQ is composed of items that regard different types of retrospective memory tasks: recall, recognition memory, and source memory. The retrospective component *M* of the MPT model, by contrast, exclusively captures the recognition of PM targets (i.e., when to perform the action) and not the recall of the PM response (i.e., what action to perform). This may obviate a correlation between the questionnaire-based measure and the model-based measure of retrospective memory. It is possible that both laboratory tasks and survey measures account for variability in naturalistic PM tasks even though they do not covary. This study does not address this question as we did not include performance measures for naturalistic intentions. As said, the inclusion of naturalistic tasks is an important venue for future research. Additionally, it is possible that some PM tasks described in the PRMQ may require less strategic monitoring for PM target occurrence than our laboratory task which was non-focal to the ongoing task (cf. McDaniel & Einstein, 2000), and that this may also contribute to the lack of correlation.

Future research should also investigate the relationship of different types of PM measures in different populations. Self- and proxy reports of PM may not reflect the extent of actual problems with PM tasks. The further development of performance-based assessments with naturalistic yet standardized PM tasks for research and clinical use may be more promising.

Author Contributions

Nina R. Arnold had the idea for the studies. She planned the study design under Ute J. Bayen's supervision. Nina R. Arnold supervised the data collections and performed all data analyses. Both authors contributed to the write-up of the studies, addressed reviewer comments, and approved of the final version of the manuscript.

Conflicts of interest

We declare no conflict of interests.

Acknowledgements

We thank Liliane Arzi, Mateja Böhm, Pia Ewerdwalbesloh, Laura Kemper, Elena Steegmüller, and Christina Trepkowski for assisting with the data collections, Silja Wübbelmann for assistance with the literature search, and Daniel Heck for help with questions concerning hyperprior selection. Parts of this research were presented at the "Tagung experimentell arbeitender Psychologen" in Gießen, Germany (2014), and at the "Hengstenberger Symposium–Theoretical and Applied Advances in Prospective Memory Research", in Heidelberg, Germany (2018).

References

Åkerstedt, T., & Gillberg, M. (1990). Subjective and objective sleepiness in the active individual. *International Journal of Neuroscience*, *52*, 29–37. http://dx.doi.org/10.3109/00207459008994241

Arnold, N. R., Bayen, U. J., & Böhm, M. F. (2015). Is prospective memory related to depression and anxiety? A hierarchical MPT modeling approach. *Memory*, *23*, 1215–1228. http://dx.doi.org/10.1080/09658211.2014.969276

Avgan, N., Sutherland, H. G., Spriggens, L. K., Rodriguez-Acevedo1, A. J., Haupt, L. M., Shum, D. H. K., & Griffiths, L. R. (2014). Variation H452Y in HTR2A gene affects immediate visual memory. *Journal of Genetics and Genome Research*, *1*, 1–4. http://dx.doi.org/10.23937/2378-3648/1410007

Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). *The CELEX lexical database (CDROM)*. Philadelphia, PA: University of Pennsylvania. Linguistic Data Consortium.

Batchelder, W. H., & Riefer, D. M. (1999). Theoretical and empirical review of multinomial process tree modeling. *Psychonomic Bulletin & Review*, *6*, 57–86. http://dx.doi.org/10.3758/BF03210812

Bleich, S., Havemann-Reinecke, U., & Kornhuber, J. (2002). *(Fagerström Test for Nicotine Dependency) FTNA – Fagerström-Test für Nikotinabhängigkeit*. Göttingen: Hogrefe, Beltz Test GmbH.

Bloch, K. E., Schoch, O. D., Zhang, J. N., & Russi, E. W. (1999). German version of the Epworth Sleepiness Scale. *Respiration*, *66*, 440–447. http://dx.doi.org/10.1159/000029408

Bigdeli, I., Farzin, A., & Talepasand, S. (2014). Prospective memory impairments in schizophrenic patients. *Iranian Journal of Psychiatry and Behavioral Sciences*, *8*, 57–63.

Brandimonte, M., Einstein, G. O., & McDaniel, M. A. (1996). *Prospective memory: Theory and applications*. Hillsdale, NJ: Erlbaum. http://dx.doi.org/10.4324/9781315806488

Chan, R. C. K., Lui, S. S. Y., Wang, Y., Liu, A. C. Y., Chui, W. W. H., Shum, D. H. K., & Cheung, E. F. C. (2012). Patients with bipolar disorders share similar but attenuated prospective memory impairments with patients with schizophrenia. *Psychological Medicine*, *43*, 1639–1649. http://dx.doi.org/10.1017/S003329171200236X

Cohen, J. (1969). *Statistical power analysis for the behavioral sciences*. New York, NY: Academic Press.

Costa, A., Peppe, A., Zabberoni, S., Serafini, F., Barban, F., Scalici, F., . . . & Carlesimo, G. A. (2015). Prospective memory performance in individuals with Parkinson's disease who have mild cognitive impairment. *Neuropsychology*, *29*, 782–791. http://dx.doi.org/10.1037/neu0000184

Crawford, J. R., Henry, J. D., Ward, A. L., & Blake, J. (2006). The prospective and retrospective memory questionnaire (PRMQ): Latent structure, normative data, and discrepancy analysis for proxy-ratings. *British Journal of Clinical Psychology*, *45*, 83–104. http://dx.doi.org/10.1348/014466505X28748

Crawford, J. R., Smith, G., Maylor, E. A., Sala, S. D., & Logie, R. H. (2003). The Prospective and Retrospective Memory Questionnaire (PRMQ): Normative data and latent structure in a large non-clinical sample. *Memory*, *11*, 261–275. http://dx.doi.org/10.1080/09658210244000027

Cuttler, C., Alcolado, G. M., & Taylor, S. (2013). Mediation and interaction effects of doubt, dysfunctional beliefs and memory confidence on the compulsion to check. *Journal of Obsessive-Compulsive and Related Disorders*, *2*, 157–166. http://dx.doi.org/10.1016/j.jocrd.2013.02.002

Cuttler, C., Relkov, T., & Taylor, S. (2013). Quick to act, quick to forget: The link between impulsiveness and prospective memory. *European Journal of Personality*, *28*, 560–568. http://dx.doi.org/10.1002/per.1926

Debarnot, U., Crépon, B., Orriols, E., Abram, M., Charron, S., Lion, S., . . . & Ergis, A.-M. (2015). Intermittent theta burst stimulation over left BA10 enhances virtual reality-based prospective memory in healthy aged subjects. *Neurobiology of Aging*, *36*, 2360–2369. http://dx.doi.org/10.1016/j.neurobiolaging.2015.05.001

Dobbs, A. R., & Rule, B. G. (1987). Prospective memory and self-reports of memory abilities in older adults. *Canadian Journal of Psychology/Revue canadienne de psychologie*, *41*, 209–222. http://dx.doi.org/10.1037/h0084152

Downey, L. A., Sands, H., Jones, L., Clow, A., Evans, P., Stalder, T., & Parrott, A. C. (2015). Reduced memory skills and increased hair cortisol levels in recent ecstasy/MDMA users: Significant but independence neurocognitive and neurohormonal deficits. *Human Psychopharmacology*, *30*, 199–207. http://dx.doi.org/10.1002/hup.2474

Einstein, G. O., & McDaniel, M. A. (1990). Normal aging and prospective memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*, 717–726. http://dx.doi.org/10.1037/0278-7393.16.4.717

Fernie, G., Bennett, D. M., Currie, J., Perrin, J. S., & Reid, I. C. (2014). Detecting objective and subjective cognitive effects of electroconvulsive therapy: Intensity, duration and test utility in a large clinical sample. *Psychological Medicine*, *44*, 2985–2994. http://dx.doi.org/10.1017/S0033291714000658

Heck, D. W., Arnold, N. R., & Arnold, D. (2018). TreeBUGS: An R package for hierarchical multinomial-processing-tree modeling. *Behavior Research Methods*, *50*, 264–284. http://dx.doi.org/10.3758/s13428-017-0869-7

Heffernan, T. M., & O'Neill, T. S. (2014). Exposure to second-hand smoke damages everyday prospective memory. *Addiction*, *108*, 420–426. http://dx.doi.org/10.1111/j.1360-0443.2012.04056.x

Horn, S. S., Bayen, U. J., Smith, R. E., & Boywitt, C. D. (2011). The multinomial model of prospective memory: Validity of ongoing-task parameters. *Experimental Psychology*, *58*, 247–255. http://dx.doi.org/10.1027/1618-3169/a000091

Hsu, Y.-H., Huang, C.-F., Tu, M.-C., & Hua, M.-S. (2014). The clinical utility of informants' appraisals on prospective and retrospective memory in patients with early Alzheimer's disease. *PLoS One*, *9* http://dx.doi.org/10.1371/journal.pone.0112210

Hultsch, D. F., Hertzog, C., & Dixon, R. A. (1990). Ability correlates of memory performance in adulthood and aging. *Psychology and Aging*, *5*, 356–368. http://dx.doi.org/10.1037/0882-7974.5.3.356

Kamat, R., Weinborn, M., Kellog, E. J., Bucks, R. S., Velnoweth, A., & Woods, S. P. (2014). Construct validity of the Memory for Intentions Screening Test (MIST) in healthy older adults. *Assessment*, *21*, 742–753. http://dx.doi.org/10.1177/1073191114530774

Kaschel, R. (2002). (*Remembering errands*) *Erinnern von Erledigungen*.. Retrieved from http://www.psy.ed.ac.uk/psy_research/documents/prmq/grprmq-1.pdf

Klauer, K. C. (2006). Hierarchical multinomial processing tree models: A latent-class approach. *Psychometrika*, *71*, 7–31. http://dx.doi.org/10.1007/s11336-004-1188-3

Klauer, K. C. (2010). Hierarchical multinomial processing tree models: A latent-trait approach. *Psychometrika*, *75*, 70–98. http://dx.doi.org/10.1007/s11336-009-9141-0

Kliegel, M., & Jäger, T. (2006). Can the Prospective and Retrospective Memory Questionnaire (PRMQ) predict actual prospective memory performance? *Current Psychology: Developmental° Learning° Personality° Social*, *25*, 182–191. http://dx.doi.org/10.1007/s12144-006-1002-8

Ling, J., Campbell, C., Heffernan, T. M., & Greenough, C. G. (2007). Short-term prospective memory deficits in chronic back pain patients. *Psychosomatic Medicine*, *69*, 144–148. http://dx.doi.org/10.1097/PSY.0b013e31802e0f22

Mäntylä, T. (2003). Assessing absentmindedness: Prospective memory complaint and impairment in middle-aged adults. *Memory & Cognition*, *31*, 15–25. http://dx.doi.org/10.3758/BF03196078

Matzke, D., Dolan, C. V., Batchelder, W. H., & Wagenmakers, E.-J. (2015). Bayesian estimation of multinomial processing tree models with heterogeneity in participants and items. *Psychometrika*, *80*, 205–235. http://dx.doi.org/10.1007/s11336-013-9374-9

McDaniel, M. A., & Einstein, G. O. (2000). Strategic and automatic processes in prospective memory retrieval: A multiprocess framework. *Applied Cognitive Psychology*, *14*, S127–S144.

Mohammadi, M. R., Keshavarzi, Z., & Talepasand, S. (2014). The effectiveness of computerized cognitive rehabilitation training program in improving cognitive abilities of schizophrenia clients. *Iranian Journal of Psychiatry*, *9*, 209–215.

Pavawalla, S. P., Schmitter-Edgecombe, M., & Smith, R. E. (2012). Prospective memory after moderate-to-severe traumatic brain injury: A multinomial modeling approach. *Neuropsychology*, *26*, 91–101. http://dx.doi.org/10.1037/a0025866

Phillips, L. H., Henry, J. D., & Martin, M. (2008). Adult aging and prospective memory. The importance of ecological validity. In M. Kliegel, M. A. McDaniel, & G. O. Einstein (Eds.), *Prospective memory. Cognitive, neuroscience, developmental, and applied perspectives* (pp. 161–185). New York, London: Taylor & Francis Psychology Press.

Raftery, A. E. (1995). Bayesian model selection in social research. *Sociological Methodology*, *25*, 111–163. http://dx.doi.org/10.2307/271063

Rendell, P. G., & Craik, F. I. M. (2000). Virtual week and actual week: Age-related differences in prospective memory. *Applied Cognitive Psychology*, *14*, 43–62. http://dx.doi.org/10.1002/acp.770

Ross, J. A., Macdiarmid, J. I., Rostron, C. L., Watt, S. J., & Crawford, J. R. (2013). Psychological and physical correlates of musculoskeletal symptoms in male professional divers and offshore workers. *Extreme Physiology & Medicine*, *2*, 5. http://dx.doi.org/10.1186/2046-7648-2-5

Rouder, J. N., Morey, R. D., & Pratte, M. S. (2017). Bayesian hierarchical models of cognition. In W. H. Batchelder, H. Colonius, E. N. Dzhafarov, & J. Myung (Eds.), *New handbook of mathematical psychology* (pp. 504–551). Cambridge, UK: Cambridge University Press.

Schaper, M. L., Kuhlmann, B. G., & Bayen, U. J. (2018). Metamemory expectancy illusion and schema-consistent guessing in source monitoring. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, http://dx.doi.org/10.1037/xlm0000602. Advance online publication

Schuler, H., & Prochaska, M. (2001). (*AMI: Achievement Motivation Inventory*) *LMI: Leistungsmotivationsinventar*. Göttingen: Hogrefe.

Schnitzspahn, K. M., Horn, S. S., Bayen, U. J., & Kliegel, M. (2012). Age effects in emotional prospective memory: Cue valence differentially affects the prospective and retrospective component. *Psychology and Aging*, *27*, 498–509. http://dx.doi.org/10.1037/a0025021

Smith, G., Della Sala, S., Logie, R. H., & Maylor, E. A. (2000). Prospective and retrospective memory in normal ageing and dementia: A questionnaire study. *Memory*, *8*, 311–321. http://dx.doi.org/10.1080/09658210050117735

Smith, J. B., & Batchelder, W. H. (2008). Assessing individual differences in categorical data. *Psychonomic Bulletin & Review*, *15*, 713–731. http://dx.doi.org/10.3758/PBR.15.4.713

Smith, J. B., & Batchelder, W. H. (2010). Beta-MPT: Multinomial processing tree models for addressing individual differences. *Journal of Mathematical Psychology*, *54*, 167–183. http://dx.doi.org/10.1016/j.jmp.2009.06.007

Smith, R. E., & Bayen, U. J. (2004). A multinomial model of event-based prospective memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*, 756–777. http://dx.doi.org/10.1037/0278-7393.30.4.756

Smith, R. E., & Bayen, U. J. (2005). The effects of working memory resource availability on prospective memory: A formal modeling approach. *Experimental Psychology*, *52*, 243–256. http://dx.doi.org/10.1027/1618-3169.52.4.243

Smith, R. E., & Bayen, U. J. (2006). The source of adult age differences in event-based prospective memory: A multinomial modeling approach. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*, 623–635. http://dx.doi.org/10.1037/0278-7393.32.3.623

Smith, R. E., Bayen, U. J., & Martin, C. (2010). The cognitive processes underlying event-based prospective memory in school-age children and young adults: A formal model-based study. *Developmental Psychology*, *46*, 230–244. http://dx.doi.org/10.1037/a0017100

Smith, R. E., Persyn, D., & Butler, P. (2011). Prospective memory, personality, and working memory: A formal modeling approach. *Zeitschrift für Psychologie/Journal of Psychology*, *219*, 108–116. http://dx.doi.org/10.1027/2151-2604/a000055

Sunderland, A., Harris, J. E., & Baddeley, A. D. (1988). Do laboratory tests predict everyday memory? A neuropsychological study. *Journal of Verbal Learning and Verbal Behaviour*, *22*, 341–357. http://dx.doi.org/10.1016/S0022-5371(83)90229-3

Thelen, J. M., Lynch, S. G., Bruce, A. S., Hancock, L. M., & Bruce, J. M. (2014). Polypharmacy in multiple sclerosis: Relationship with fatigue, perceived cognition, and objective cognitive performance. *Journal of Psychosomatic Research*, *76*, 400–404. http://dx.doi.org/10.1016/j.jpsychores.2014.02.013

Thöne-Otto, A. I. T., & Walter, K. (2008). Assessment and treatment of prospective memory disorders in clinical practice. In M. Kliegel, M. A. McDaniel, & G. O. Einstein (Eds.), *Prospective memory. Cognitive, neuroscience, developmental, and applied perspectives* (pp. 321–345). New York, London: Taylor & Francis Psychology Press.

Unsworth, N., Brewer, G. A., & Spillers, G. J. (2012). Variation in cognitive failures: An individual differences investigation of everyday attention and memory failures. *Journal of Memory and Language*, *67*, 1–16.

Walter, N. T., & Bayen, U. J. (2016). Selective effects of acute alcohol intake on the prospective and retrospective components of a prospective-memory task with emotional targets. *Psychopharmacology*, *233*, 325–339. http://dx.doi.org/10.1007/s00213-015-4110-z

Wamsley, E., Donjacour, C. E., Scammell, T. E., Lammers, G. J., & Stickgold, R. (2014). Delusional confusion of dreaming and reality in narcolepsy. *Sleep*, *37*, 419–422. http://dx.doi.org/10.1007/s00213-015-4110-z

Weinborn, M., Moyle, J., Bucks, R. S., Stritzke, W., Leighton, A., & Woods, S. P. (2013). Time-based prospective memory predicts engagement in risk behaviors among substance users: Results from clinical and nonclinical samples. *Journal of the International Neuropsychological Society*, *19*, 284–294. http://dx.doi.org/10.1017/S1355617712001361

Willians, D. M., Jarrold, C., Grainger, C., & Lind, S. E. (2014). Diminished time-based, but undiminished event-based, prospective memory among intellectually high-functioning adults with autism spectrum disorder: Relation to working memory ability. *Neuropsychology*, *28*, 30–42. http://dx.doi.org/10.1037/neu0000008

Wilson, B. A., Emslie, H., Foley, J., Shiel, A., Watson, P., Hawkins, K., Groot, Y., & Evans, J. (2005). *The Cambridge Prospective Memory Test (CAMPROMT*. London: Harcourt Assessment.

Zeintl, M., Kliegel, M., Rast, P., & Zimprich, D. (2006). Prospective memory complaints can be predicted by prospective memory performance in older adults. *Dementia and Geriatric Cognitive Disorders*, *22*, 209–215. http://dx.doi.org/10.1159/000094915

Received 12 October 2018;
accepted 10 April 2019
Available online 27 May 2019