Canadian Journal of Experimental Psychology
2011, Vol. 65, No. 1, 57–68

1196-1961/11/$12.00 DOI: 10.1037/a0022843

Self-Report Measures of Prospective Memory Are Reliable but Not Valid

Bob Uttl
Mount Royal University

Mekale Kibreab
University of Calgary

Are self-report measures of prospective memory (ProM) reliable and valid? To examine this question, 240 undergraduate student volunteers completed several widely used self-report measures of ProM including the Prospective Memory Questionnaire (PMQ), the Prospective and Retrospective Memory Questionnaire (PRMQ), the Comprehensive Assessment of Prospective Memory (CAPM) questionnaire, self-reports of retrospective memory (RetM), objective measures of ProM and RetM, and measures of involvement in activities and events, memory strategies and aids use, personality and verbal intelligence. The results showed that both convergent and divergent validity of ProM self-reports are poor, even though we assessed ProM using a newly developed, reliable continuous measure. Further analyses showed that a substantial proportion of variability in ProM self-report scores was due to verbal intelligence, personality (conscientiousness, neuroticism), activities and event involvement (busyness), and use of memory strategies and aids. ProM self-reports have adequate reliability, but poor validity and should not be interpreted as reflecting ProM ability.

Keywords: prospective memory, self-reports, measurement, reliability, validity

Prospective memory (ProM) refers to our ability to make plans, to retain them, and to bring them back into consciousness at the right time and place (Graf & Uttl, 2001; Uttl, 2008). ProM is divided into several subdomains: ProM proper, vigilance/monitoring, and habitual ProM (Graf & Uttl, 2001). ProM proper is engaged in episodic, one-time tasks such as buying groceries en route home, whereas vigilance/monitoring is required for tasks such as watching a kettle so it does not boil over. Whereas the plan dominates consciousness in vigilance/monitoring, in ProM proper the plan leaves consciousness and must be brought back to consciousness at the right time and place in response to a ProM cue (Brandimonte, Ferrante, Feresin, & Delbello, 2001; Graf & Uttl, 2001; Meacham & Leiman, 1982; Uttl, 2008). Habitual ProM is similar to ProM proper, but the plan must be executed repeatedly (Graf & Uttl, 2001; Harris, 1984; Uttl, 2008); habitual ProM is used for completing tasks such as taking medication repeatedly at prescribed times over a long period of time. When discussing ProM in this article, we refer to ProM proper as opposed to vigilance/monitoring.

Although most of the research on ProM to date has employed objective task measures of ProM, recently, self-reports of ProM have become increasingly more popular and are often used to assess and make inferences about ProM ability (Chan, Qing, Wu, & Shum, 2010; Chau, Lee, Fleming, Roche, & Shum, 2007; Crawford, Smith, Maylor, Della Sala, & Logie, 2003; Heffernan & Bartholomew, 2006; Roche, Moody, Szabo, Fleming, & Shum, 2007). Surprisingly, a review of the published literature on the development and psychometric properties of ProM self-reports reveals that while these measures show sufficient to excellent reliability, evidence for their validity is lacking with the few available studies showing no evidence of convergent or divergent validity (Hannon, Adams, Harrington, Fries-Dias, & Gipson, 1995; Kliegel & Jager, 2006; Mantyla, 2003).

Table 1 summarizes the available evidence concerning the reliability and validity of three commonly used self-report measures of ProM: the Prospective Memory Questionnaire (PMQ; Hannon et al., 1995), the Prospective and Retrospective Memory Questionnaire (PRMQ; Smith, Della Sala, Logie, & Maylor, 2000), and the Comprehensive Assessment of Prospective Memory (CAPM; Roche et al., 2007). The evidence in Table 1 suggests that these instruments have acceptable to very good reliability but evidence for their validity is lacking; a comprehensive search for validity studies revealed only a few studies (listed in Table 1) and none of them found any appreciable validity coefficients. Correlations with objective measures of ProM are extremely low, explaining at best about 4% of the variability in objective ProM task measures (Hannon et al., 1995; Kliegel & Jager, 2006; Mantyla, 2003), indicating a lack of convergent validity. Correlations with objective measures of RetM (reported for PRMQ only) are similarly low (Mantyla, 2003), indicating a lack of divergent validity. In addition, the validity of self-reports has typically been examined using laboratory ProM tasks and their validity against naturalistic tasks is largely unknown.

Moreover, previous attempts to validate ProM self-reports suffer from several methodological problems. First, they have used inefficient and unreliable binary measures of ProM as the criterion objective measures of ProM (Foster, McDaniel, Repovs, & Hershey, 2009; Hannon et al., 1995; Kliegel & Jager, 2006; Mantyla, 2003). Given the low reliability of these binary measures,

Bob Uttl, Psychology Department, Mount Royal University; and Mekale Kibreab, Psychology Department, University of Calgary.

Mekale Kibreab is now at the Addiction Centre, Foothills Medical Centre, Calgary, AB.

We thank Amy L. Siegenthaler for careful reading and comments on the manuscript.

Correspondence concerning this article should be addressed to Bob Uttl, Department of Psychology, Mount Royal University, 4825 Mount Royal Gate SW, Calgary, AB T3E 6K6, Canada. E-mail: buttl@mtroyal.ca or uttlbob@gmail.com

Table 1
ProM Self-Reports: Review of Reliability and Validity Evidence

Study	Subscales	Cronbach's alpha	Test-retest reliability	Correlations			Notes
				ProM lab tasks	ProM naturalistic tasks	RetM lab tasks	
PMQ							
Hannon et al., 1995	Total, LTE, STH, IC	0.78 to 0.92	0.64 to 0.88	−0.17 to −0.25	−0.06 to + 0.01		
PRMQ							
Crawford et al., 2003		0.80 to 0.89					
Mantyla, 2003	ProM			−0.20		0.10	
	RetM			−0.22		0.05	
Kliegel & Jager, 2006	ProM			−0.19 to −0.22			
	RetM			ns/nr			
Thadeus Meeks, Hicks, & Marsh, 2007	ProM			−0.13 to −0.17			
	RetM			−0.09 to −0.19			
Foster et al., 2009	ProM			−0.17 to + 0.13			[1]
CAPM/A							
Fleming et al., 2009		0.74 to 0.94		−0.24 to + 0.22			[2]

Note. ns/nr = not statistically significant but not reported.
[1] Control participants only. [2] Criterion measures were CAM-PROMPT and MIST.

the lack of convergent validity may be an artifact of poor criterion measurement rather than poor validity of self-reports (Uttl, 2005a, 2008). Second, the low validity coefficients reported in some studies may be an artifact of ceiling effects in the criterion measures (Foster et al., 2009; Hannon et al., 1995; Kliegel & Jager, 2006). Third, the low validity coefficient may be the result of criterion measures that have confounded prospective and retrospective components (Fleming et al., 2009). Fourth, in at least one validation study (Fleming et al., 2009), examinees were allowed to use external memory aids to support their performance on criterion measure of ProM.

The present study has several objectives. The first objective was to develop a new continuous measure of laboratory ProM and to examine the validity of ProM self-reports against this new continuous measure. One of the limitations of the previous validation studies is the use of binary measures of ProM that have limited reliability relative to continuous measures of ProM, restricting the magnitude of any validity coefficients (Uttl, 2005a, 2005b, 2008). We have developed a new continuous measure of ProM patterned after the continuous measures developed by Uttl and his colleagues (Graf, Uttl, & Dixon, 2002; Uttl, 2006). For the present study, participants were instructed to circle any and all occurrences of a ProM cue while working through various questionnaires (Ellis, Kvavilashvili, & Milne, 1999; Kvavilashvili, 1998). The ProM cue appeared four times in an increasingly intrusive visual form and spatial location. The first ProM cue circled was used as an index of ProM ability; participants who circled the first occurrence of the ProM cue received 4 points, whereas participants who did not circle any cues received 0 points. To examine the reliability of this new measure, each participant completed two assessments.

The second objective was to examine the convergent and divergent validity of self-reports of ProM. Convergent validity refers to the degree to which two measures of the same or similar construct correlate with one another. Divergent validity is the extent to which two measures of different constructs do not correlate with each other. If various ProM self-reports are valid measures of ProM ability, correlations among different self-reports of ProM should be relatively high, indicating that they measure the same construct and that the reported findings would not vary with the specific self-report used. Moreover, correlations between self-reports of ProM and other constructs, for example, self-reports of RetM should be relatively low. For this purpose, we included the three widely used self-reports of ProM mentioned above (PMQ, PRMQ Pro, CAPM/A), a newer self-report measure of ProM – TCPMQ (Cuttler & Graf, 2009), the MemQ of ProM (Uttl, 2002a), and several self-reports of RetM – including PRMQ RetM (Smith et al., 2000), the EMQ-R (Royle & Lincoln, 2008a), and the MemQ RetM (Uttl, 2002a).

The third objective of the present study was to examine the convergent and divergent validity of self-reports of ProM against objective measures of both laboratory and naturalistic ProM. If ProM self-reports are valid, correlations between ProM self-reports and objective measures of lab and naturalistic ProM should be relatively high whereas correlations between ProM self-reports and objective measures of other constructs should be relatively low. Accordingly, the present study evaluated the validity of ProM self-reports against two objective binary laboratory measures of ProM, two objective continuous laboratory measures of ProM, and two binary naturalistic measures of ProM. In addition, we examined correlations between ProM self-reports and measures of verbal intelligence (Uttl, 2002b) and personality (Costa & McCrae, 1992).

The fourth and final objective of the present study was to determine whether the lack of correlations between ProM self-reports and objective measures of ProM could be due to two other factors. First, the low validity coefficients could be due to the use of compensatory strategies and aids in naturalistic settings (Groot, Wilson, Evans, & Watson, 2002; Thone-Otto & Walther, 2008). If people use external compensatory strategies and aids effectively, they are less likely to experience ProM failures, and therefore, less likely to report ProM failures on the self-report measures of ProM. Second, the low validity coefficients could occur because of vari-

ability among people in their involvement in various activities and events requiring ProM, sometimes called "busyness" (Marsh, Hicks, & Landau, 1998; Martin & Park, 2003; Ronnlund, Mantyla, & Nilsson, 2008). Busy people have more opportunities for ProM failures and thus are likely to experience more ProM failures even though their ProM ability may be comparable to less busy people. Accordingly, we also measured participants' use of compensatory strategies and aids and participants' involvement in activities and events normally requiring ProM using the Activities and Events Involvement Scale (AEI; Uttl, 2002c).

Method

Participants

Participants were 240 undergraduate student volunteers, 182 females and 58 males (age $M = 22.4$ years, $SD = 6.45$). The first language of the majority of participants was English (90.8%). The study was conducted with the approval of the Red Deer College Research Ethics Board. The study took approximately 2 hours and each participant received two course participation credits.

Instruments

As part of a larger study (Uttl & Kisinger, 2010), participants were administered several objective tests of ProM and RetM, several self-reports/questionnaires of ProM and RetM failures, several questionnaires enquiring about the use of memory aids and strategies, a questionnaire about involvement in activities and events, a personality test, and a verbal knowledge test. These instruments and tasks are described below.

Continuous Lab Measure of Event Cued ProM (Lab EC ProM/C). The continuous measure of EC ProM was developed for this study, patterned after a continuous measures previously used by Uttl and his colleagues (Graf et al., 2002; Uttl, 2006). To allow assessment of test–retest reliability, participants' ProM ability was assessed on two separate occasions. For each assessment, participants were instructed to circle any and all occurrences of the word "collection" (first assessment) and "demanding" (second assessment) as they worked through various paper-and-pencil tests and questionnaires. The specific instructions were:

> *We want to examine your ability to do something in the future. Thus, if you encounter the word "collection" at any point during this experiment, please circle it. I will not remind you again but it is important that you circle any and all occurrences of the word "collection."* Do you have any question about this task? [pause for any questions] *It is important that you circle any and all occurrences of the word "collection."*

Following the instructions, participants worked through the tasks and the ProM cue appeared four times in an increasingly intrusive visual form and spatial location. For the first assessment, the cue (word "collection") appeared in the following visual form and spatial locations: (1) in lower case, normal font, nonprominent location, as part of question #8 on CAPM/A; (2) in lower case, normal font, more prominent location, embedded in the last question (#17) on the PRMQ page (this question was added to the PRMQ for this purpose but was not included in the scoring of the PRMQ); (3) in lower case, bold, more prominent location, appearing as the last statement on the EMQ-R page (i.e., "Please check that you answered all questions in this **collection**. Thank you."); and (4) in capitals, bold, more prominent location, appearing as part of the instructions for Memory Aiding Strategies (i.e., "For each statement in this **COLLECTION**, choose the response alternative that best describes how often you use each of the strategies.") For the second assessment, the cue ("demanding") appeared in the following visual form and spatial locations: (1) in lower case, normal font, nonprominent location, as part of question #14 on Self-Efficacy for Studying; (2) in lower case, normal font, more prominent location, embedded in the last question (#61) on the NEO-FFI (this question was added to NEO-FFI for this purpose and was not included in the scoring of the NEO-FFI); (3) in lower case, bold, more prominent location, appearing as the last statement on the TCPMQ (i.e., "Please check you have answered every question, both easy and **demanding** ones. Thank you."); and (4) in capitals, bold, more prominent location, appearing as part of the instructions for Memory Failures task (i.e., "Don't forget to e-mail us even if you have a **DEMANDING** day."). For each assessment, participants received a score of 4 if the first circled cue was the first cue, 3 if the first circled cue was the 2nd cue, 2 if it was the 3rd cue, 1 if it was the 4th cue, and 0 if they did not circle any cue. Thus, each participant's test score was determined by the first cue to which the participant responded and the score depended on whether the first responded cue was the first cue, the second cue, and so forth At least 15 minutes elapsed between the ProM instructions and the first appearance of the ProM cue.

Binary Lab Measure of Event Cued ProM (Lab EC ProM/B). The performance on the first cue appearance of the continuous lab measure of EC ProM reflects performance on a binary measure of ProM, commonly used in studies of ProM, with 1 and 0 indicating success and failure in circling the first cue, respectively.

Memory Failures Task and Naturalistic Time Cued ProM Assessment (Naturalistic TC ProM 1 and 2). We asked participants to e-mail us back on the 3rd and 7th day following their participation in the study with a list of things they had failed to remember during the previous 24 hours and reasons why they thought they had forgotten them. Participants indicated on a response form on which days and at what times they would e-mail us with the list and the reasons for forgetting (participants could choose any time for their e-mails). If participants e-mailed us back within 24 hours of their chosen time, they would receive a score of 1. If they did not e-mail us or e-mailed us outside of the ± 24 hour window, they would receive a score of 0. E-mails to be sent on the 3rd and 7th day served as the first and second assessment of naturalistic time cued ProM (Naturalistic TC ProM 1 and 2), respectively.

Verbal Learning Test Unrelated 20 (VLT/U20). The VLT/U20 (Uttl, 2006) is a measure of explicit RetM patterned after the Rey Auditory Verbal Learning Test (Strauss, Sherman, & Spreen, 2006). The test involved a series of three study-test trials on one set of 20 semantically unrelated words. On each trial, examinees listened to a list of 20 words read by an experimenter and were required to write down as many words as they could remember, in any order, after the examiner had completed reading the list. For each trial, participants were given 90 seconds for recall. Thus, the scores on each trial as well as the average of the three trials score could range from 0 to 20.

Verbal knowledge (Words/A40). The Words/A40 is a 40-item multiple-choice tests designed to assess examinees' vocabulary knowledge (Uttl, 2002b). Each item consists of a target word and four other words out of which one word is similar in meaning. Each item is scored as correct (1 point), incorrect (0 points), or not answered (0.25 points to correct for a failure to guess), and the average across all items is the proportion correct.

Prospective and Retrospective Memory Questionnaire (PRMQ). The PRMQ (Smith et al., 2000) is a 16-item questionnaire assessing the frequency of memory failures on two main subscales: prospective memory subscale (PRMQ ProM; e.g., "Do you fail to mention or give something to a visitor that you were asked to pass on") and retrospective memory subscale (PRMQ RetM; e.g., "Do you forget something that you were told a few minutes before"). The PRMQ ProM is further subdivided into four 2-item scales: *Prospective short-term self-cued*; *Prospective short-term environmentally cued*; *Prospective long-term self-cued*; and *Prospective long-term environmentally cued*. Similarly, the PRMQ RetM is subdivided into four 2-item scales: *Retrospective short-term self-cued*; *Retrospective short-term environmentally cued*; *Retrospective long-term self-cued*; and *Retrospective long-term environmentally cued*. Examinees rate the frequency of their ProM and RetM failures on a 5-point scale: *Never* = 1, *Rarely* = 2, *Sometimes* = 3, *Quite often* = 4, and *Very often* = 5. To obtain scale scores, item responses were averaged for all items within each scale. Thus, higher PRMQ scores reflect more self reported memory failures.

Prospective Memory Questionnaire (PMQ). The PMQ (Hannon et al., 1995) is a 52-item self report assessing ProM failures as well as the frequency of use of memory aids. The questionnaire consists of four subscales. The *long-term episodic* scale (PMQ LTE) includes 14 items that refer to tasks that occur infrequently and must be performed following an extended number of hours or days after a cue (e.g., "I forgot to return my book to the library by the due date"). The *short-term habitual* scale (PMQ STM) consists of 14 items. These items refer to regularly occurring tasks that must be completed relatively soon after the cue (e.g., "I forgot to lock the door when leaving my apartment"). The *internally cued* scale (PMQ IC) includes 10 items referring to tasks that have no external cues to facilitate remembering (e.g., "I forgot what I wanted to say in the middle of a sentence"). The *techniques to assist memory* scale (PMQ TECH) includes 14 items that assess the use of memory aiding strategies [for example, "I make Post-It (sticky note) reminders and place them in obvious places"]. Examinees rate the frequency of each memory failure on a visual analog scale ranging from *0 times* to either *4 or more times* or *6 or more times* (depending on the specific item) per week, month, or year (depending on the specific item). A not applicable response can also be made if the item does not apply. However, examinees often find the specific visual analogue scale confusing (Cuttler & Graf, 2009). Thus, to simplify the response scoring and allow the use of machine readable questionnaire forms, examinees in our study rated the frequency of each memory failure on a 7-point scale ranging from *0 times* to *6 or more times* (i.e., 0, 1, 2, 3, 4, 5, 6 or more times) per specified time period used in the original PMQ and were also allowed to select a *N/A* response. To obtain scale scores, item response were averaged for all items within each scale. Thus, higher scores indicate more ProM failures, with the exception of the PMQ TECH scale where higher scores indicate the use of more memory strategies.

Comprehensive Assessment of Prospective Memory (CAPM). The CAPM (Roche et al., 2007) is a 54-item questionnaire based on Waugh's CAPM published earlier (Waugh, 1999). It is divided into three sections. Section A (CAPM/A) assesses the frequency of ProM lapses using 39 items with 10 items focusing on basic activities of daily living (e.g.,"Forgetting to eat a meal"; CAPM/A BADL), 23 items focusing on instrumental activities of daily living (e.g., "Not remembering to pay bills"; CAPM/A IADL), and 6 unclassified items contributing only to CAPM/A overall score. Section B (CAPM/B) includes the same 39 items as Section A but examinees are asked to indicate the degree of concern about each ProM failure. Section A and B items are rated on a 5-point scale: 1 = *Never*; 2 = *Rarely* (*once/month*); 3 = *Occasionally (2–3 times/month*); 4 = *Often (once/week*); and 5 = *Very often (daily*). A not applicable response can also be made if the item does not apply. Section C includes 15 items that examine the reasons for ProM failures [for example, "The more things (say two or three) I have to do, the more likely I will forget to do them."]. Examinees are asked to indicate their agreement with each statement using a 4-point Likert scale (i.e., 1 = *Strongly disagree*; 2 = *Disagree*; 3 = *Agree*; and 4 = *Strongly agree*). For sections A and B, item responses were averaged for all items within each scale. Accordingly, higher scores indicate more frequent ProM failures.

Time Cued Prospective Memory Questionnaire (TCPMQ). The TCPMQ (Cuttler & Graf, 2009) has three sections. The first section includes 39 items and measures the frequency of time-cued ProM failures (TCPMQ Freq). Examinees rate how often they forget to do various activities that must be executed at a specific time (e.g., "I forgot to go to my dentist appointment") on a 5-point scale: 1 = *Never*; 2 = *Seldom*; 3 = *Sometimes*; 4 = *Often*; and 5 = *Very often,* or n/a = *Not applicable*. Item responses are averaged across all items and higher scores on the TCPMQ Freq indicate more time-cued ProM failures. The second section of the questionnaire assesses punctuality (TCPMQ Punc). Examinees are presented with the same 39 items and asked to rate how punctual they are when performing various tasks. The third section, called Memory Aiding Strategies (TCPMQ MAS), is an 11-item self-report measuring the frequency of memory aids and strategy use including the use of Post-It notes, alarms, cell phones, and reminder services (e.g., "I use the alarm function on my PDA to help me remember to do things on time"). All items are answered on a 5-point scale: 1 = *Never*; 2 = *Seldom*; 3 = *Sometimes*; 4 = *Often*; and 5 = *Very often;* or n/a = *Not applicable*. The responses are averaged across all items, with higher scores indicating higher use of these external memory aids and strategies.

Memory Questionnaire (MemQ). The MemQ (Uttl, 2002a) is a 51-item self-report assessing the frequency of ProM and RetM memory failures in everyday life. The ProM subscale consists of 37 items asking about ProM failures (e.g., "I forget to stop at a store and get groceries as I planned."; "I forget to attend lectures, classes, and workshops."). The RetM subscale consists of 14 items asking about RetM failures (e.g., "I forget where I parked my car."; "I forget directions to places when I heard them only once."). Examinees rate the frequency of their memory failures on a 7-point scale: 1 = *Never*; 2 = *Very rarely*; 3 = *Rarely*; 4 = *Sometimes*;

5 = *Often*; 6 = *Very often*; and 7 = *Always*). A not applicable response can also be made if the item does not apply. The responses are averaged across all items for each scale and higher scores indicate a higher frequency of self-reported memory failures.

Everyday Memory Questionnaire–Revised (EMQ-R). The EMQ-R (Royle & Lincoln, 2008b) is a 13-item measure assessing subjective beliefs about everyday memory failures. Almost all items concern RetM failures (e.g., "Getting the details of what someone has told you mixed up and confused."; "Forgetting when it was that something happened;, e.g., whether it was yesterday or last week.") with only two items concerning ProM failures (i.e., "Completely forgetting to do things you said you would do, and things you planned to do" and "Forgetting to tell somebody something important, perhaps forgetting to pass on a message or remind someone of something."). Examinees rate how often things happen to them on a 5-point scale (i.e., 1 = *Once or less in the last month*; 2 = *More than once a month but less than once a week*; 3 = *About once a week*; 4 = *More than once a week or less than once a day*; and 5 = *Once or more in a day*). Item responses are averaged across all items and higher scores reflect primarily RetM failures.

Memory Aids Questionnaire (MAidQ). The MAidQ (Uttl, 2002d) is a 27-item questionnaire asking about the frequency of use of various memory aids and strategies, both internal (e.g., "I associate to-be-remembered information with other things") and external (e.g., "I use built-in alarms in cars, ovens, etc. to remind me to do something"). Fifteen items focus on aids and strategies supporting ProM, eight items focus on aids and strategies supporting RetM, and four items focus on aids and strategies that could be used to support either ProM or RetM. Examinees rate the frequency with which they use aids and strategies on a 7-point scale: 1 = *Never*; 2 = *Very rarely*; 3 = *Rarely*; 4 = *Sometimes*; 5 = *Often*; 6 = *Very often*; and 7 = *Always*. The responses are averaged across all items for each scale, with higher scores indicating more frequent use of aids and strategies.

Activity and Events Involvement (AEI). The AEI (Uttl, 2002c) is a 15-item questionnaire which examines the frequency of participation and involvement in various activities and events (e.g., "I attend a meeting or an activity of a volunteer organization I belong to"; "I think about things I have to do."). Examinees rate the frequency of their participation on an 8-point scale (i.e., 1 = *Never or almost never*; 2 = *Monthly*; 3 = *2–3 times a month*; 4 = *Weekly*; 5 = *2–6 times a week*; 6 = *Daily*; 7 = *2–3 times a day*; and 8 = *4 or more times a day*). The responses are averaged across all items and higher scores indicate participation and involvement with more activities and events.

NEO FFI Personality Inventory (NEO FFI). The NEO FFI (Costa & McCrae, 1992) is a standardized self-report personality inventory measuring five personality domains: Neuroticism, Extraversion, Openness to Experience, Agreeableness, and Conscientiousness (the Big Five). Each domain is measured by a scale consisting of 12 statements. Participants rate their agreement with each statement on a 5-point scale: *Strongly Disagree*, *Disagree*, *Neutral*, *Agree*, and *Strongly Agree*. We calculated each scale's score as the average across the scale's 12 items (after reverse scoring the required items), with higher scores indicating higher levels of the personality traits.

Procedure

Participants were tested in small groups, in a single session lasting about 2 hours, in a small classroom, seated widely separated. After providing written, informed consent and basic demographic information (age, gender, whether their first language was English), they completed the set of tasks listed in Table 2. The table indicates the order of the tasks, the approximate time required for completing each task, and the tasks that included ProM cues.

In the first phase, the participants were given EC ProM Task 1 instructions as described above. After ensuring that all questions were answered the experimenter distributed the first set of questionnaires—AEI, MemQ, and MAidQ—and participants completed these questionnaires at their own pace, placing each completed page on the desk directly behind them (this prevented participants from going back and circling cues they may have not responded to at the right time). After all participants completed all questionnaires, the second phase began. In the second phase, participants watched a short movie clip and completed several tasks unrelated to this study (Uttl & Kisinger, 2010). In the third phase, the experimenter distributed another set of questionnaires—CAPM/A + B, PRMQ, EMQ-R, and TCPMQ MAS. Again, participants completed these questionnaires at their own pace, placing each completed page on the desk directly behind them. In the fourth phase, the experimenter administered the VLT/U20 and

Table 2
Sequence of All Instruments and Tasks Completed by Participants, Including ProM Cue Appearances, With the Approximate Time Required to Complete Each Instrument/Task

Task	Approximate time required (min)
EC ProM Task 1 instructions	2
Activities & Events Inventory (AEI)	5
Memory Questionnaire (MemQ)	5
Memory Aids Questionnaire (MAidQ)	5
Accident movie	2
Accident recall 1 & 2	5
Accident questionnaire	5
Accident causes questionnaire	3
Comprehensive Assesssment of Prospective Memory (CAPM/A + B) [1st cue]	6
Prospective Retrospective Memory Questionnaire (PRMQ) [2nd cue]	4
Everyday Memory Questionnaire–Revised (EMQ-R) [3rd cue]	4
Memory aid strategies [4th cue]	4
Verbal Learning Test Trials A1–3 (VLT/U20)	10
—5-min break—	5
EC ProM Task 2 instructions	2
Verbal Knowledge Test (Words/A40)	10
Prospective Memory Questionnaire (PMQ)	8
Self Efficacy for Learning (SEL) [1st cue]	5
NEO-FFI Personality Inventory (NEO-FFI) [2nd cue]	8
Time Cued Prospective Memory Questionnaire (TCPMQ) [3rd cue]	5
Memory Failures Instructions (Naturalistic TC ProM) [4th cue]	5
Movie check + comments	2

after this test participants were allowed a 5-min break with refreshments (e.g., soda, chocolates).

Following the break, in the fifth phase, participants were given EC ProM Task 2 instructions as described above and ensured that all questions were answered. Next, the experimenter distributed Words/A40 questionnaire and participants were allowed 10 minutes to complete this task. When all participants completed Words/A40, the experimenter distributed another set of questionnaires—PMQ, SEL, NEO-FFI, TCPMQ, and Memory Failures Instructions—and participants completed these questionnaires at their own pace, placing each completed page directly behind them. When finished, the experimenter checked with each participant that they understood the Memory Failures Instructions and selected the times for sending the two e-mails. Next, the participant was given the last form that asked them to indicate whether they had seen the shown movie previously and to provide any comments on the study.

Results

Data Screening

The data were screened for univariate outliers defined as scores falling 1.5 interquartile ranges below the 25th percentile or above the 75th percentile. Less than 1.7% of data values were univariate outliers. To reduce their influence on any analyses, all univariate outliers were replaced with corresponding outlier caps (i.e., a value 1.5 interquartile ranges either below the 25th percentile or above the 75th percentile, as appropriate).

Binary Versus Continuous Objective Measures of ProM

Table 3 shows the means, standard deviations, assessment to assessment (test-retest) reliabilities for individual measures of ProM—Lab EC ProM/B 1 and 2, Lab EC ProM/C 1 and 2, and Naturalistic TC ProM/B 1 and 2—as well as for composite measures of ProM (i.e., averages of two assessments). Consistent with prior research findings, performance increased from the first to the second assessment on both binary and continuous laboratory measures of EC ProM, and performance declined on the naturalistic measure of TC ProM with delay, that is, from the first assessment at 3 days to the second assessment at 7 days. As expected, for the lab measures of EC ProM, the test–retest reliability was substantially higher for individual continuous versus binary measures (Graf & Uttl, 2001; Uttl, 2005a, 2006, 2008). Similarly, Cronbach's alpha was also substantially higher for composite continuous versus binary measures (Kaplan & Saccuzzo, 2008). For the naturalistic measures of TC ProM, the correlation between the success at 3 days and 7 days was 0.59 and Cronbach's alpha for the composite (average of the two trials) was 0.74. Thus, the reliabilities of the lab EC ProM/C composite measure as well as the naturalistic TC ProM/B measures were respectable but the reliability of the lab EC ProM/B measure was below what is generally considered acceptable (Kaplan & Saccuzzo, 2008).

Table 3
Objective Measures of ProM and RetM: Descriptive Statistics, Individual Measures Test-Retest Reliability and Composites Reliability

	M	*SD*	r_{xx}	α
ProM 1/B	0.24	0.43	0.32	
ProM 2/B	0.43	0.50		
ProM/B	0.34	0.38		0.48
ProM 1/C	2.16	1.41	0.56	
ProM 2/C	2.86	1.35		
ProM/C	2.51	1.22		0.71
TC ProM 1/B	0.35	0.48	0.59	
TC ProM 2/B	0.25	0.43		
TC ProM	0.30	0.41		0.74
VLT/U20 A1	8.89	1.90		
VLT/U20 A2	12.54	2.34		
VLT/U20 A3	14.80	2.47		
VLT/U20	12.08	1.95		0.83

Note. $N = 240$, except for EC ProM 1/C, ProM 2/C, and ProM/C where $N = 176$; r_{xx} = test-retest reliability using parallel forms; α = Cronbach's α.

Table 3 also shows the means, standard deviations, and reliabilities for individual measures of RetM—VLT/U20 A1, VLT/U20 A2, VLT/U20 A3—as well as for a composite measure of RetM (i.e., average of the three study-test trial; VLT/U20). Cronbach's alpha of the composite was 0.83.

Correlations Among Objective Measures of ProM, RetM, Verbal Intelligence, and Personality

The laboratory measure of EC ProM/C did not correlate significantly with the naturalistic measure of TC ProM, $r = .05$, suggesting that the performance on these two tasks depends on different processes (see Table 4). The laboratory measure of EC ProM/C was weakly correlated with the measure of RetM (VLT/U20), $r = .25$, and verbal intelligence, $r = .36$, confirming the previous research findings (Uttl, 2006, 2008). The EC ProM/C was also correlated with personality, specifically, with agreeableness, $r = .19$, and neuroticism, $r = -0.20$. In contrast, the naturalistic measure of TC ProM/B was weakly correlated with verbal intelligence only, $r = .14$. Consistent with prior research, the measure of RetM (VLT/U20) correlated with verbal intelligence, $r = .35$, and agreeableness, $r = .18$.

Reliabilities of Self-Reports

Table 4 includes the means, standard deviations, and Cronbach's alphas for all measures, including the self-reports of ProM. The ProM self-report reliabilities ranged from a low of 0.75 (CAPM/B BADL) to a high of 0.92 (TCPMQ Freq), with most of the reliabilities higher than 0.84. The reliabilities of self-reported ProM aids use ranged from 0.71 (MAidQ ProM) to 0.89 (PMQ Tech). The reliabilities of RetM self-reports ranged from 0.73 (PRMQ RetM) to 0.87 (EMQ-R). The reliability of the single RetM aids use index was only 0.46 (MAidQ RetM).

Validity of Self-Reports

Table 5 shows the correlation matrix of all measures. If self-reports of ProM are valid measures of ProM, then correlations among the self-reports of ProM should be relatively high and correlations between self-reports of ProM and measures of other

Table 4
Descriptive Statistics, Reliability, and Correlations With Objective Measures of ProM, RetM, and Verbal Intelligence

				Correlations				
	M	*SD*	α	EC ProM/B	EC ProM/C	TC ProM	VLT/U20	Words/A40
EC ProM/B	0.34	0.38	0.48	1.00	**.82**	.08	**.14**	**.31**
EC ProM/C	2.51	1.22	0.71	**.82**	**1.00**	.05	**.25**	**.36**
TC ProM	0.30	0.41	0.74	.08	.05	1.00	.01	**.14**
VLT/U20	12.08	1.95	0.83	**.14**	**.25**	.01	1.00	**.35**
Words/A40	0.66	0.15	0.80	**.31**	**.36**	**.14**	**.35**	1.00
PMQ LTE	0.84	0.57	0.90	**−.17**	**−.33**	**−.15**	−.01	**−.26**
PMQ STE	0.28	0.27	0.84	**−.16**	**−.35**	**−.22**	−.11	**−.32**
PMQ ICH	1.05	0.75	0.89	**−.15**	**−.32**	**−.16**	−.03	**−.22**
PRMQ ProM	2.61	0.50	0.80	.01	−.04	−.09	.05	**.16**
MemQ ProM	2.80	0.66	0.93	**−.17**	**−.23**	**−.15**	.06	.03
CAPM/A IADL	2.27	0.49	0.90	**−.13**	**−.21**	**−.13**	.10	−.02
CAPM/B BADL	1.79	0.46	0.75	−.12	**−.25**	−.06	.07	**−.13**
TCPMQ Freq	1.61	0.34	0.92	**−.15**	**−.28**	−.07	.01	**−.13**
PRMQ RetM	2.34	0.48	0.73	−.08	−.11	−.09	.01	.06
MemQ RetM	3.52	0.77	0.85	−.12	**−.15**	−.12	.01	.04
EMQ−R	1.09	0.63	0.87	−.09	**−.23**	**−.18**	−.02	**−.20**
PMQ TECH	2.28	1.23	0.89	−.12	**−.21**	.10	−.02	−.01
MAidQ ProM	2.94	0.69	0.71	−.10	−.13	.09	.04	−.05
MAidQ RetM	3.43	0.86	0.46	−.06	−.13	.08	.02	.08
TCPMQ MAS	2.41	0.69	0.82	−.04	−.13	−.03	.09	.04
AEI	3.49	0.61	0.65	**−.20**	**−.18**	.00	.06	**−.21**
Openness	3.31	0.52	0.72	.07	.14	.05	.07	**.20**
Conscientiousness	3.61	0.55	0.83	.08	.13	.12	−.02	−.04
Extraversion	3.55	0.53	0.79	−.11	−.07	−.09	.05	**−.17**
Agreeableness	3.62	0.51	0.76	**.22**	**.19**	.10	**.18**	.12
Neuroticism	2.85	0.71	0.87	**−.18**	**−.20**	−.05	−.10	−.04
Gender (F = 1, M = 0)	0.76	0.43		.06	.11	.10	.12	.01

Note. $N = 240$, except for EC ProM 1/C, ProM 2/C, and ProM/C where $N = 176$. Bold print = $p < .05$.

constructs (e.g., RetM) should be relatively low. However, the pattern of correlations provides no evidence of divergent validity; correlations among the self-reports of ProM (range: 0.31 to 0.75; average 0.58) are generally comparable with correlations between the self-reports of ProM and self-reports of RetM (range: 0.28 to 0.73; average 0.52). The lack of divergent validity is further highlighted by high correlations between ProM and RetM subscales of PRMQ and MemQ, 0.72 and 0.73, respectively, indicating that the ProM and RetM subscales in these measures tap substantially the same construct.

More important, if at least some self-reports are valid measures of ProM, the correlations between these self-reports and objective measures of ProM should be high, and relatively higher than between the self-reports and measures of other constructs, for example, RetM and verbal intelligence. An examination of the relevant correlations suggests only weak evidence of convergent validity with objective measures of both laboratory and naturalistic ProM. First, ProM self-reports show weak correlations with composite continuous measure of laboratory EC ProM, with correlations ranging from −0.04 (PRMQ ProM) to −0.35 (PMQ STE). As expected from the reliability data, ProM self-reports show only inconsistent and very weak correlations with composite binary measures of laboratory EC ProM, with correlations ranging from 0.01 (PRMQ ProM) to −0.17 (PMQ LTE). Second, ProM self-reports showed weak but generally significant correlations with composite binary measure of naturalistic TC ProM, ranging from −0.06 (CAPM/A BADL) to −0.22 (PMQ STE).

The examination of correlations between ProM self-reports and other constructs provide only very limited evidence of divergent validity. ProM self-reports showed no significant correlations with objective measures of RetM (correlation range: −0.11 to 0.10). However, ProM self-reports displayed weak but significant correlations with ProM memory aids use, *r*s ranging from 0.04 to 0.36; weak but significant correlations with activity and events involvement scores, $r = .13$ to 0.28; weak but significant correlations with an objective measure of verbal intelligence (*r*s range: 0.03 to −0.32), weak to moderate correlations with neuroticism (*r*s range: 0.08 to 0.29); and weak to moderately strong negative correlations with conscientiousness (*r*s range: −0.25 to −0.50).

One of the most commonly used self-reports of ProM, the PRMQ ProM subscale, did not correlate significantly with any of the objective ProM, RetM, or verbal intelligence measures. In contrast, PMQ subscales showed the strongest correlations with all of the reliable objective ProM measures, ranging from −0.32 to −0.35 with the laboratory measure of EC ProM/C and from −0.15 to −0.22 with the naturalistic measure of TC ProM/B.

Activities and Events Involvement, Memory Aids Use, and Gender

Activities and events involvement scores were correlated with extraversion ($r = .27$) and memory aids use (*r*s range: 0.20 to 0.48, see Table 5). Female participants were generally more likely to use memory aids, were more conscientious, and more agreeable. Con-

Table 5
Correlation Matrix

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27
1. EC ProM/B		**.82**	.08	**.14**	**.31**	**−.17**	**−.16**	**−.15**	.01	**−.17**	**−.13**	−.12	**−.15**	−.08	−.12	−.09	−.12	−.10	−.06	−.04	**−.20**	.07	.08	−.11	**.22**	**−.18**	.06
2. EC ProM/C			.05	**.25**	**.36**	**−.33**	**−.35**	**−.32**	−.04	**−.23**	**−.21**	**−.25**	**−.28**	−.11	**−.15**	**−.23**	**−.21**	**−.13**	**−.13**	**−.13**	**−.18**	**.14**	**.13**	−.07	**.19**	**−.20**	.11
3. TC ProM				.01	**.14**	**−.15**	**−.22**	**−.16**	−.09	**−.15**	**−.13**	−.06	−.07	−.09	−.12	**−.18**	.10	.09	.08	−.03	.00	.05	.12	−.09	.10	−.05	.10
4. VLT/U20					**.35**	−.01	−.11	−.03	.05	.06	.10	.07	.01	.01	.01	−.02	−.02	.04	.02	.09	.06	.07	−.02	.05	**.18**	−.10	.12
5. Words/A40						**−.26**	**−.32**	**−.22**	**.16**	.03	−.02	**−.13**	**−.13**	.06	.04	**−.20**	−.01	−.05	.08	.04	**−.21**	**.20**	−.04	**−.17**	.12	−.04	.01
6. PMQ LTE							**.58**	**.71**	**.49**	**.58**	**.63**	**.56**	**.67**	**.45**	**.44**	**.60**	**.23**	**.18**	**.15**	**.29**	**.28**	−.04	**−.33**	.04	−.01	**.16**	.04
7. PMQ STE								**.56**	**.31**	**.41**	**.47**	**.53**	**.51**	**.28**	**.28**	**.47**	.12	.04	.06	**.18**	**.20**	−.03	**−.25**	.12	−.02	.08	−.10
8. PMQ IC									**.49**	**.52**	**.60**	**.55**	**.59**	**.50**	**.51**	**.71**	**.22**	.11	.14	**.24**	**.19**	−.10	**−.34**	.02	−.05	**.23**	.07
9. PRMQ ProM										**.66**	**.64**	**.44**	**.54**	**.72**	**.64**	**.56**	**.28**	**.26**	**.22**	**.28**	**.15**	.10	**−.30**	**−.13**	−.02	**.26**	.10
10. MemQ ProM											**.79**	**.62**	**.71**	**.58**	**.73**	**.50**	**.15**	**.17**	**.18**	**.31**	**.13**	.12	**−.50**	−.05	.03	**.26**	**.14**
11. CAPM/A IADL												**.73**	**.75**	**.59**	**.63**	**.61**	**.22**	**.23**	**.23**	**.36**	**.20**	.11	**−.37**	−.09	.04	**.18**	**.13**
12. CAPM/A BADL													**.61**	**.42**	**.44**	**.52**	**.18**	**.15**	**.20**	**.30**	**.22**	−.02	**−.32**	−.03	.04	**.13**	.07
13. TCPMQ Freq														**.41**	**.50**	**.48**	**.16**	**.19**	**.18**	**.33**	**.13**	.03	**−.42**	.00	.02	**.29**	**.19**
14. PRMQ RetM															**.69**	**.64**	**.22**	**.19**	**.19**	**.21**	**.19**	.07	**−.32**	−.07	−.09	**.19**	.02
15. MemQ RetM																**.59**	**.20**	**.18**	**.16**	**.28**	**.14**	.10	**−.37**	−.11	.01	**.31**	**.14**
16. EMQ−R																	**.20**	.07	.02	.19	**.22**	−.07	**−.30**	−.06	−.05	**.18**	−.05
17. PMQ Tech																		**.71**	**.55**	**.39**	**.48**	.06	**.18**	.03	.10	.07	**.19**
18. MAidQ ProM																			**.66**	**.49**	**.48**	−.01	**.17**	.00	.07	.12	**.26**
19. MAidQ RetM																				**.32**	**.26**	.08	.06	.06	.06	.05	**.23**
20. TCPMQ MAS																					**.20**	−.01	**−.20**	−.02	.05	**.13**	.10
21. AEI																						.00	.09	**.27**	.08	−.08	.05
22. Openness																							−.05	.12	**.21**	**−.14**	.00
23. Conscientiousness																								**.17**	**.24**	**−.35**	**.14**
24. Extraversion																									**.33**	**−.36**	.00
25. Agreeableness																										**−.30**	**.20**
26. Neuroticism																											.12
27. Gender																											

Note. $N = 240$, except for EC ProM 1/C, ProM 2/C, and ProM/C, where $N = 176$. Bold print = $p < .05$.

scientiousness was correlated with two ProM memory aids use measures (PMQ Tech $r = .18$, MAidQ ProM $r = .17$) but negatively correlated with the third ProM memory aids use measure (TCPMQ MAS $r = -0.20$).

Hierarchical Regression Analyses

To gain a better understanding of variables that predict performance on the EC ProM/C an TC ProM/B, we also conducted two hierarchical regression analyses to find out whether personality explains additional variability in ProM above that explained by verbal intelligence, and whether PMQ (the self-report measure of ProM with the highest validity coefficients, see below) explains additional variability in ProM above and beyond that explained by verbal intelligence and personality. The main findings of these hierarchical regression analyses are reported in Table 6. Verbal intelligence, personality (agreeableness and neuroticism), and PMQ each explained additional variability in the EC ProM/C above and beyond the variability explained by the variables entered on the preceding steps, explaining 26.8% of variability in total. In contrast, only verbal intelligence and PMQ explained additional variability in the TC ProM/B, explaining only 3.6% of variability in total.

In the next hierarchical analysis, summarized in Table 7, we examined the amount of variability explained in PMQ by verbal intelligence, personality (conscientiousness, neuroticism), activities and events involvement ("busyness"), and memory strategies and aids use. Each of these variables, with the exception of neuroticism, explained additional variability above and beyond the variability explained by the variables entered previously. In total, these variables explained 34.3% of variability in the PMQ (Multiple $R = 0.59$).

Discussion

Our study resulted in several important new findings. Our newly developed, easily administered continuous measure of ProM has favorable psychometric properties when compared to binary success/failure measures. The higher reliability of continuous versus binary ProM measures resulted in generally higher correlations with RetM measures, verbal intelligence, as well as ProM self-reports. Accordingly, poor reliabilities of the binary criterion measures of ProM used in the previous validation studies may have contributed to the extremely low validity coefficients reported previously (see Table 1).

Table 6
Hierarchical Regression Analyses of Performance on EC ProM/C and TC ProM/B

	Increase R^2	R^2	Multiple R
EC ProM/C			
1. Words/A40	0.127*	0.127	0.30
2. Agreeableness Neuroticism	0.036*	0.163	0.40
3. PMQ	0.107*	0.268	0.52
TC ProM/B			
1. Words/A40	0.019*	0.019	0.14
2. PMQ	0.017*	0.036	0.19

Note. $N = 176$ for EC ProM/C. $N = 240$ for TC ProM/B.
* $p < .05$.

Table 7
Hierarchical Regression Analysis of Performance on PMQ

	Increase R^2	R^2	Multiple R
1. Words/A40	0.090*	0.090	0.30
2a. Conscientiousness	0.102*	0.192	0.44
2b. Neuroticism	0.002	0.194	0.44
3. AEI	0.083*	0.277	0.53
4. PMQ Tech	0.066*	0.343	0.59

Note. $N = 240$.
* $p < .05$.

Correlations among the self-reports of ProM were generally moderate and comparable to correlations between the self-reports of ProM and RetM. These results indicate that self-reports of ProM lack both convergent and divergent validity with other self-reports of ProM and RetM, respectively. To illustrate, PRMQ ProM correlated the highest with PRMQ RetM ($r = .72$) rather than with other self-reports of ProM. These high intercorrelations between PRMQ ProM and RetM are consistent both with the previous findings suggesting that PRMQ is unidimensional rather than having tripartite factor structure (Crawford et al., 2003). First, the single factor model of PRMQ had acceptable fit and the tripartite model (general memory factor plus two orthogonal ProM and RetM factors) had an only marginally better fit. Moreover, the tripartite model's general factor versus specific PRMQ ProM and RetM factors explained about 10 times as much variance in each of the 16 PRMQ items (Crawford et al., 2003). Second, the previous research has consistently found that PRMQ ProM and RetM scores are highly intercorrelated, with the reported correlations generally above 0.70 (Macan, Gibson, & Cunningham, 2010; Mantyla, 2003; Ronnlund et al., 2008). Accordingly, the available evidence to date strongly suggests that the three PRMQ scores measure primarily the same thing, rather than distinct components of memory.

Correlations between the self-reports of ProM and objective measures of both laboratory and naturalistic ProM were generally very weak, indicating a lack of convergent validity with criterion measures. The laboratory ProM performance was best predicted by PMQ subscales (rs $= -0.32$ to -0.35) and the least predicted by PRMQ ProM ($r = -0.04$). Thus, ProM self-reports predict at best only about 12% of variability in the criterion measure. Naturalistic ProM performance was also best predicted by PMQ subscales (rs $= -0.15$ to -0.22). In contrast, PRMQ ProM, TCPMQ, and CAPM/B BADL showed no significant correlations with the naturalistic ProM performance. Thus, in general, the relationship between the self-reports of ProM and performance on the naturalistic versus laboratory tasks was even weaker, explaining at best 5% of variance in the criterion measures. Moreover, some ProM self-reports display no divergent validity with objective measures of RetM. To illustrate, the correlation between PRMQ ProM and objective measures of ProM was comparable to the correlation between PRMQ ProM and the objective measure of RetM.

The present study identified several additional factors (beyond poor reliability of criterion measures in previous studies) that may be responsible for low correlations between ProM self-reports and performance on objective measures of ProM: individual differences in demands on ProM, use of compensatory strategies, and

personality. First, self reports of ProM were correlated with high demands on ProM, as measured by the AEI scale. High demands on ProM lead to more opportunities for ProM failure, increased self-reported ProM failure, and in turn, reduced correlations between self-reports and objective ProM measures. Previous research has found an inconsistent relationship between busyness and ProM. Using the Environmental Demands Inventory (Martin & Park, 2003) to assess busyness, Martin and Park (2003) reported that busy people experienced more self-reported ProM failures, but Cuttler and Graf (2007) found no relationship between busyness and objective ProM task measures. In contrast to both of these studies, we used the AEI scale that more directly measures examinees' involvement in events and activities requiring ProM (i.e., not forgetting about events, keeping appointments) rather than general busyness measured by the Environmental Demands Inventory. As expected, the AEI scale also correlated with participants' use of memory aids and strategies, with *r*s ranging from 0.20 (TCPMQ MAS) to 0.48 (PMQ Tech, MAidQ ProM) as well as extroversion ($r = .27$).

Second, self-reports were also correlated with the use of compensatory strategies. Individuals who experience more frequent ProM failures (e.g., because of their poor ProM and/or because of their involvement with activities and events requiring ProM) are more likely to compensate for their ProM problems by using ProM strategies and aids. In turn, the use of ProM strategies and aids is likely to reduce correlations between self-reports and objective ProM measures (Groot et al., 2002).

Third, self-reports of ProM were correlated with both conscientiousness (*r*s ranging from -0.25 to -0.50) and neuroticism (*r*s ranging from 0.08 to 0.29). Consistent with previous research findings (Cuttler & Graf, 2007; Pearman & Storandt, 2005), conscientious people reported fewer ProM failures and they also performed better on the objective continuous measure of ProM. In contrast, neuroticism was associated with more reported ProM failures and with worse performance on objective measures of ProM (Cuttler & Graf, 2007; Pearman & Storandt, 2005). Although agreeableness was not associated with ProM self-reports, it was positively related to better performance on both objective measures of ProM as well as RetM. Hierarchical regression analysis showed that verbal intelligence, personality, activities, and event involvement, and use of ProM strategies and aids, together explained 34% of variability in PMQ scores.

Another possible explanation for the lack of validity not addressed by the present study is that most of the current ProM questionnaires do not distinguish between ProM subdomains (vigilance/monitoring, ProM proper, habitual ProM), and thus, measure a mixture of partially overlapping but distinct constructs. Results, however, may vary depending on ProM subdomain. To illustrate, Uttl (2008) demonstrated that the magnitude of age declines in ProM proper is much larger than the magnitude of age declines on vigilance/monitoring tasks.

Rabbitt and colleagues (Rabbitt, Maylor, McInnes, Bent, & Moore, 1995) have argued some time ago that self-reports of cognitive functioning depend on many variables at the same time, including the to-be-measured cognitive functions but also personality variables, individual differences in demands on cognition, use of memory aids, depression, and so forth, and warned against interpreting self-report scores as measures of ability. We extend that warning to the domain of ProM self-reports. Self-report measures of ProM do not strongly correlate with each other, are influenced by many of the noncognitive factors discussed by Rabbit et al., and they explain only minute proportions of variance in performance on both laboratory and naturalistic tasks.

The low validity of ProM self-reports aside, the present study suggests that one of the most promising self-report measures of ProM may be the PMQ whereas one of the least valid measures of ProM may be the PRMQ. The PMQ shows generally the highest (although still small) correlations with ProM tasks and low correlations with RetM tasks. In contrast, the PRMQ ProM and RetM subscales are strongly intercorrelated and neither was significantly correlated with ProM nor RetM tasks in the present study. We speculate that one reason for PMQ's better performance may be its objective frequency scale where examinees are required to state how often per week, month, or year they experience specific memory failures. In contrast, the other ProM self-reports ask examinees to rate the frequency of their ProM failure using nonobjective scales; for example, the PRMQ asks participants to rate the frequency of memory failures using a 5-point scale ranging from *never* to *very often*. Thus, examinees themselves have to interpret the meaning of these labels. If they interpret them relative to their peers, they may experience lots of failures but since people around them experience them too, they may answer "rarely" rather than "often" (Rabbitt et al., 1995).

In conclusion, self-report measures of ProM are reliable but not sufficiently valid. Although some self-reports predict performance on both the laboratory and naturalistic ProM tasks, they explain too small a percentage of variability in the criterion measures to be considered valid and useful in predicting criterion performance. Accordingly, ProM self-report scores should not be interpreted as reflecting ProM ability.

Résumé

Les mesures auto-rapportées de la mémoire prospective (MPro) sont-elles fidèles et valides? Afin de répondre à cette question, 240 étudiants volontaires de premier cycle universitaire ont complété plusieurs mesures auto-rapportées de MPro largement répandues, incluant le *Prospective Memory Questionnaire* (PMQ), le *Prospective and Retrospective Memory Questionnaire* (PRMQ), le questionnaire *Comprehensive Assessment of Prospective Memory* (CAPM), des auto-rapports de mémoire rétrospective (MRet), ainsi que des mesures objectives de la MPro et de la MRet, de la participation à des événements et activités, des stratégies mnémoniques et des aide-mémoire utilisés, de la personnalité et de l'intelligence verbale. Les résultats ont montré que la validité convergente et la validité divergente des auto-rapports de la MPro sont toutes deux faibles, m]]

Mots-clés : mémoire prospective, auto-rapports, mesure, fidélité, validité

References

Brandimonte, M., Ferrante, D., Feresin, C., & Delbello, R. (2001). Dissociating prospective memory from vigilance processes. *Psicologica, 22,* 97–113.

Chan, R. C. K., Qing, Y., Wu, Q., & Shum, D. (2010). Prospective memory

in healthy Chinese people: The latent structure of the Comprehensive Assessment of Prospective Memory Questionnaire. *Neuropsychological Rehabilitation, 20,* 459–470. doi:10.1080/09602010903425710

Chau, L., Lee, J., Fleming, J., Roche, N., & Shum, D. (2007). Reliability and normative data for the Comprehensive Assessment of Prospective Memory (CAPM). *Neuropsychological Rehabilitation, 17,* 707–722. doi:http://dx.doi.org/10.1080/09602010600923926

Costa, P., & McCrae, R. (1992). *Revised NEO Personality inventory (NEO PI-R) and NEO five-factor inventory (NEO-FFI): Professional manual.* Odessa, FL: Psychological Assessment Resources.

Crawford, J. R., Smith, G., Maylor, E. A., Della Sala, S., & Logie, R. H. (2003). The Prospective and Retrospective Memory Questionnaire (PRMQ): Normative data and latent structure in a large non-clinical sample. *Memory (Hove, England), 11,* 261–275.

Cuttler, C., & Graf, P. (2007). Personality predicts prospective memory task performance: An adult lifespan study. *Journal of Psychology, 48,* 215–231. doi:http://dx.doi.org/10.1111/j.1467–9450.2007.00570.x

Cuttler, C., & Graf, P. (2009). Sub-clinical compulsive checkers show impaired performance on habitual, event- and time-cued episodic prospective memory tasks. *Journal of Anxiety Disorders, 23,* 813–823. doi:http://dx.doi.org/10.1016/j.janxdis.2009.03.006

Ellis, J., Kvavilashvili, L., & Milne, A. (1999). Experimental tests of prospective remembering: The influence of cue-event frequency on performance. *Journal of Psychology, 90,* 9–23. doi:http://dx.doi.org/10.1348/000712699161233

Fleming, J., Kennedy, S., Fisher, R., Gill, H., Gullo, M., & Shum, D. (2009). Validity of the Comprehensive Assessment of Prospective Memory (CAPM) for use with adults with traumatic brain injury. *Brain Impairment, 10,* 34–44. doi:http://dx.doi.org/10.1375/brim.10.1.34

Foster, E., McDaniel, M., Repovs, G., & Hershey, T. (2009). Prospective memory in Parkinson disease across laboratory and self-reported everyday performance. *Neuropsychology, 23,* 347–358. doi:http://dx.doi.org/10.1037/a0014692

Graf, P., & Uttl, B. (2001). Prospective memory: A new focus for research. *Consciousness Cognition, 10,* 437–450.

Graf, P., Uttl, B., & Dixon, R. A. (2002). Prospective and retrospective memory in adulthood. In P. Graf, & N. Ohta, (Eds.), *Lifespan development of human memory* (pp. 255–282). London: The MIT Press.

Groot, Y. C., Wilson, B., Evans, J., & Watson, P. (2002). Prospective memory functioning in people with and without brain injury. *Journal of the International Neuropsychological Society, 8,* 645–654. doi:http://dx.doi.org/10.1017/S1355617702801321

Hannon, R., Adams, P., Harrington, S., Fries-Dias, C., & Gipson, M. (1995). Effects of brain injury and age on prospective memory self-rating and performance. *Rehabilitation Psychology, 40,* 289–298. doi:http://dx.doi.org/10.1037/0090–5550.40.4.289

Harris, J. E. (1984). Remembering to do things: A forgotten topic. In In J. E. Haris, & P. E. Morris, (Eds.), *Everyday memory: Actions and absentmindedness* (pp. 71–92). London: Academic Press.

Heffernan, T., & Bartholomew, J. (2006). Does excessive alcohol use in teenagers affect their everyday prospective memory? *Journal of Adolescent Health, 39,* 138–140. doi:http://dx.doi.org/10.1016/j.jadohealth.2005.10.010

Kaplan, R. M., & Saccuzzo, D. P. (2008). *Psychological testing: Principles, applications, and issues* (7th ed.). Beverly, MA: Wadsworth Publishing.

Kliegel, M., & Jager, T. (2006). Can the Prospective and Retrospective Memory Questionnaire (PRMQ) predict actual prospective memory performance? *Current Psychology, 25,* 182–191.

Kvavilashvili, L. (1998). Remembering intentions: Testing a new method of investigation. *Applied Cognitive Psychology, 12,* 533–554. doi:http://dx.doi.org/10.1002/%28SICI%291099–0720%281998120%2912:6%3C533::AID-ACP538%3E3.0.CO;2–1

Macan, T., Gibson, J., & Cunningham, J. (2010). Will you remember to read this article later when you have time? The relationship between prospective memory and time management. *Personality and Individual Differences, 48,* 725–730. doi:http://dx.doi.org/10.1016/j.paid.2010.01.015

Mantyla, T. (2003). Assessing absentmindedness: Prospective memory complaint and impairment in middle-aged adults. *Memory & Cognition, 31,* 15–25.

Marsh, R., Hicks, J., & Landau, J. (1998). An investigation of everyday prospective memory. *Memory & Cognition, 26,* 633–643.

Martin, M., & Park, D. C. (2003). The Martin and Park Environmental Demands (MPED) Questionnaire: Psychometric properties of a brief instrument to measure self-reported environmental demands. *Aging Clinical and Experimental Research, 15,* 77–82.

Meacham, J., & Leiman, B. (1982). Remembering to perform future actions. In *Memory observed: Remembering in natural contexts* (pp. 327–336). San Francisco: Freeman.

Pearman, A., & Storandt, M. (2005). Self-discipline and self-consciousness predict subjective memory in older adults. *The Journals of Gerontology Series B, Psychological Sciences and Social Sciences, 60,* P153–P157.

Rabbitt, P., Maylor, E., McInnes, L., Bent, N., & Moore, B. (1995). What goods can self-assessment questionnaires deliver for cognitive gerontology? *Applied Cognitive Psychology, 9,* S127–S152. doi:10.1002/acp.2350090709

Roche, N., Moody, A., Szabo, K., Fleming, J., & Shum, D. H. (2007). Prospective memory in adults with traumatic brain injury: An analysis of perceived reasons for remembering and forgetting. *Neuropsychological Rehabilitation, 17,* 314–334. doi:http://dx.doi.org/10.1080/09602010600831004

Ronnlund, M., Mantyla, T., & Nilsson, L. (2008). The Prospective and Retrospective Memory Questionnaire (PRMQ): Factorial structure, relations to global subjective memory ratings, and Swedish norms. *Journal of Psychology, 49,* 11–18.

Royle, J., & Lincoln, N. B. (2008a). The Everyday Memory Questionnaire-revised: Development of a 13-item scale. *Disability and Rehabilitation, 30,* 114–121. doi:10.1080/09638280701223876

Royle, J., & Lincoln, N. B. (2008b). The Everyday Memory Questionnaire – revised: Development of a 13-item scale. *Disability & Rehabilitation, 30,* 114. doi:10.1080/09638280701223876

Smith, G., Della Sala, S., Logie, R., & Maylor, E. (2000). Prospective and retrospective memory in normal ageing and dementia: A questionnaire study. *Memory, 8,* 311–321. doi:http://dx.doi.org/10.1080/096582100406801

Strauss, E., Sherman, E. M. S., & Spreen, O. (2006). *A compendium of neuropsychological tests: Administration, norms, and commentary* (3rd ed.). Oxford, U. K.: Oxford University Press.

Thadeus Meeks, J., Hicks, J. L., & Marsh, R. L. (2007). Metacognitive awareness of event-based prospective memory. *C onsciousness and Cognition, 16,* 997–1004. doi:10.1016/j.concog. 2006.09.005

Thone-Otto, A. I., & Walther, 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. London: Psychology Press.

Uttl, B. (2002a). *Memory questionnaire*. Unpublished manuscript.

Uttl, B. (2002b). *Words/A40: Multiple choice index of verbal intelligence*. Unpublished manuscript.

Uttl, B. (2002c). *Activities and event involvement questionnaire*. Unpublished manuscript.

Uttl, B. (2002d). *Memory aids questionnaire*. Unpublished manuscript.

Uttl, B. (2005a). Age-related changes in event-cued prospective memory proper. In In N. Ohta, C. M. MacLeod, & B. Uttl, (Eds.), *Dynamic cognitive processes* (pp. 273–303). Tokyo: Springer.

Uttl, B. (2005b). Measurement of individual differences: Lessons from memory assessment in research and clinical practice. *Psychological Science, 16,* 460–467.

Uttl, B. (2006). Age-related changes in event-cued visual and auditory prospective memory proper. *Aging, Neuropsychology, and Cognition, 13,* 141–172. doi:http://dx.doi.org/10.1080/138255890969618

Uttl, B. (2008). Transparent meta-analysis of prospective memory and aging. *PLoS ONE, 3,* e1568.

Uttl, B., & Kisinger, K. (2010). Meaning of missing values in eyewitness recall and accident records. *PLoS ONE, 5,* e12539.

Waugh, N. (1999). *Self-report of the young, middle-aged, young-old, and old-old individuals on prospective memory functioning* (Unpublished undergraduate honors thesis). Griffith University, Brisbane, Australia.

Received September 1, 2010
Accepted December 29, 2010 ■