This is the published version


Available from Deakin Research Online

http://hdl.handle.net/10536/DRO/DU:30058471

Reproduced with the kind permission of the copyright owner

Copyright: 2011, Taylor and Francis
Event-based prospective remembering in a virtual world

Steven L. Trawley, Anna S. Law & Robert H. Logie

Accepted author version posted online: 27 May 2011. Published online: 08 Jul 2011.

To cite this article: Steven L. Trawley, Anna S. Law & Robert H. Logie (2011) Event-based prospective remembering in a virtual world, The Quarterly Journal of Experimental Psychology, 64:11, 2181-2193, DOI: 10.1080/17470218.2011.584976

To link to this article: http://dx.doi.org/10.1080/17470218.2011.584976
Event-based prospective remembering in a virtual world

Steven L. Trawley¹, Anna S. Law², and Robert H. Logie¹

¹Human Cognitive Neuroscience, University of Edinburgh, Edinburgh, UK
²Department of Psychology, Liverpool John Moores University, Liverpool, UK

Most laboratory-based prospective memory (PM) paradigms pose problems that are very different from those encountered in the real world. Several PM studies have reported conflicting results when comparing laboratory- with naturalistic-based studies (e.g., Bailey, Henry, Rendell, Phillips, & Kliegel, 2010). One key contrast is that for the former, how and when the PM cue is encountered typically is determined by the experimenter, whereas in the latter case, cue availability is determined by participant actions. However, participant-driven access to the cue has not been examined in laboratory studies focused on healthy young adults, and its relationship with planned intentions is poorly understood. Here we report a study of PM performance in a controlled, laboratory setting, but with participant-driven actions leading to the availability of the PM cue. This uses a novel PM methodology based upon analysis of participant movements as they attempted a series of errands in a large virtual building on the computer screen. A PM failure was identified as a situation in which a participant entered and exited the “cue” area outside an errand related room without performing the required errand whilst still successfully remembering that errand post test. Additional individual difference measures assessed retrospective and working memory capacity, planning ability and PM. Multiple regression analysis showed that the independent measures of verbal working memory span, planning ability, and PM were significant predictors of PM failure. Correlational analyses with measures of planning suggest that sticking with an original plan (good or bad) is related to better overall PM performance.

Keywords: Prospective memory; Planning; Virtual reality; Working memory.

Our ability to remember to perform an activity at a specific future time or place is known as prospective memory (PM). As such intentions can only be realized at a later time, and as subsequent tasks demand our attention, we typically encode these intentions in memory and then “forget” them until the appropriate situation arises. An individual must frequently recall an intention when there is no explicit reminder to prompt them. For example, imagine a man driving home who suddenly realizes, shortly before driving past a supermarket, that he had intended earlier that day to buy a comic for his
daughter; appearing relieved, he quickly decreases his speed and pulls over. Laboratory-based PM research, following the standard Einstein and McDaniel (1990; e.g., McDaniel & Scullin, 2010) paradigm, attempts to mimic such situations by requiring participants to perform both an ongoing task, such as lexical decision, and a concurrent “background” PM task that requires the participant to make a specified response to a particular target embedded in the ongoing task (e.g., during a lexical decision task press the space bar if you see an example of a fruit).

Although very successful, this approach has never been used to explore performance in situations where either the timing or appearance of the PM cue can be influenced by the participant. Furthermore, with abstract stimuli divorced from any situational or social context, the anticipated prospective event is essentially something that the volunteers will never have encountered before. For instance, forming a PM to purchase a comic would require specific information about the retrieval cue (namely the supermarket) to indicate the moment at which the intention should be realized, such as its relative location (left/right roadside) and context (e.g., light/heavy traffic), both of which are affected by what route the driver actually travels.

By taking a different route home, the driver may encounter the supermarket from a different perspective than had been imagined when forming the intention. There is a body of work that has examined this discrepancy between initial cue encoding and what is perceived at retrieval. Several studies (e.g., Cook, Marsh, & Hicks, 2005; Logie & Maylor, 2009; Maylor & Logie, 2010; Nowinski & Dismukes, 2005) have manipulated context via the initial instructions given to the participants. These papers agreed in their conclusion that the probability of successful cue detection is affected by how information is processed at encoding and subsequently interacts with the perceived PM cue at the point of retrieval. In situations where the disparity between encoding and retrieval was high, performance was always impaired in these studies (see also, Ellis & Milne, 1996; McDaniel, Robinson-Riegler, & Einstein, 1998). However, as far as we are aware, no laboratory PM study has explored how changes initiated by the participant could induce an encoding/retrieval PM disparity on cue appearance. All manipulations to date have been experimenter driven, with the timing and appearance of the PM cue insensitive to the actions of the participant as they perform the ongoing task.

Our primary aim, therefore, was to investigate the impact of participant-driven actions on PM performance and to do so by highlighting the relationship between planning and successful PM. This relationship was explored by Kliegel, McDaniel, and Einstein (2000) who showed the importance of plan elaboration and plan following on successful PM performance. Participants could change cue presentation time by performing tasks in a different order than originally planned. However, the tasks used by Kliegel et al. (2000) were always located on a table in front of the participant. Clearly the order in which participants performed the tasks would not affect the relative location and appearance of the cue, whereas participant-driven task order is typical of many real-life PM scenarios. The new approach that we adopt in the present study contrasts with conventional PM methods, in that context manipulations are generated by the participant, not by the experimenter. Differential predictions can be made depending upon whether or not one assumes that the pretest plan generated by the participant creates a contextual relationship between each errand and its related information such as expected cue appearance from a given viewpoint based upon task order. If there is such a contextual relationship, we would expect that participants who adhere closely to their plan will exhibit fewer PM errors than those who do not. If there is no such relationship, and participants have only a loose order planned, then spontaneous changes to actual completion of the task order in response to PM cues when they happen to appear should have little negative effect on PM errors. In this case, performance may even be better because the PM cue prompts enactment of an intention at the time the cue is encountered rather than the participant performing the tasks in the planned order regardless of when they encounter each cue.

A second aim was to explore the relationship between working memory and successful PM performance. The role of working memory in PM has
typically been examined either by making the ongoing task harder or by giving participants an additional task to perform concurrently. For example, Marsh and Hicks (1998) conducted several experiments showing that only tasks that placed a demand on the central executive adversely affected PM. Moreover, several studies since have highlighted the relationship between individual differences of working memory capacity and PM (Brewer, Knight, Marsh, & Unsworth, 2010; Einstein, McDaniel, Manzi, Cochran, & Baker, 2000; Smith, 2003; Smith & Bayen, 2005; West & Craik, 2001). However, all of these studies have used verbal working memory tasks as their estimator of individual working memory capacity. To the authors’ knowledge, this is the first time PM performance has been explored from a domain–specific working memory process perspective (Baddeley & Logie, 1999; Logie & Baddeley, in press) by indexing both verbal and visuospatial working memory capacity as predictors in a regression model. Independent measures of retrospective memory, planning, and PM were also used as predictors in the regression model.

As noted, in most laboratory paradigms for studying PM, cue presentation is predefined by the experimenter. A range of studies have used more naturalistic settings, many of which have focused on the age–prospective memory paradox in which older people appear to outperform younger people on PM tasks in the naturalistic setting but not in a laboratory setting (e.g., Bailey et al., 2010; Rendell & Craik, 2000). In these settings, the participant’s actions do determine when and how a PM cue is encountered. However, genuine naturalistic settings are very complex and lack experimental control, so results may be driven by factors of which the experimenter is not aware or cannot influence. Realistic scenarios in the laboratory have been explored using video recordings of real-world scenes (e.g., Farrimond, Knight, & Titov, 2006), or laboratory simulations (e.g., Craik & Bialystok, 2006; Rendell & Craik, 2000; Paraskevaides et al., 2010). However, the Farrimond et al. (2006) simulated shopping task lacks an ongoing task, and the authors acknowledge that limitation. Although a study by Kinsella, Ong, and Tucker (2009) specifically addressed this limitation by asking participants to monitor the shopping video for “specials offers” while performing their virtual shop, all of these paradigms restrict when cues are encountered and/or the order in which participants perform actions. Therefore, a third aim was to introduce a novel PM methodology in a controlled, laboratory setting but where cue presentation is determined by the movement sequences chosen by the participant as they undertake a range of tasks. As such, the relationship between encoding and retrieval can be disrupted, virtually step by step, by the choices made by the participant in the intervening retention phase. Our approach is based upon analysis of the route the participant takes as they attempt a series of errands in a large virtual building using the Edinburgh Virtual Errands Task (EVET) (Logie, Trawley, & Law, 2010). This combines a simulation of a realistic setting with control of the environment, the range of cues that the participant will encounter, and the range of actions that the participant may perform. In the EVET, each errand has a specific location within the virtual building, spread over 38 rooms and four floors. Access to each floor is provided by two sets of stairs, one for travelling up and the other for travelling down. By allowing participants to roam freely in this virtual space, we were able to examine the effect of cue encoding/retrieval disparity as a consequence of the participant’s self-determined route. Participants could encounter PM cues (such as a room number or a stairwell) from a variety of directions, presenting several possible PM cue perspectives. Moreover, the context in which these cues are encountered is also variable, such as when they are encountered (early or late in the test) and what tasks are currently active (number of items carried). For example, one errand involved collecting a keycard on the left-hand side of the second floor, but as part of a different errand the participant might be carrying a package to be delivered elsewhere in the building. Prior to starting the test, every participant indicated their optimum errand order and, therefore, by definition, their direction of travel to each errand. During the test, however, each participant has several possible navigational routes to the keycard, such as entering the left side of the second floor via...
the stairs or crossing the second-floor concourse from the right. Furthermore, when they decide to perform this errand during the test, it may be when they happen to encounter a particular cue (e.g., a specific room number), and this may occur earlier than envisaged in the original plan, with several tasks already completed or left to do. This variation provides the basis for the encoding/retrieval discrepancy. In summary, the aim of this paper is to investigate how PM failures in the EVET were related to planning, participant-driven actions, and independent measures of cognitive functioning, including tests of verbal and spatial working memory capacity.

Method

Participants

An initial total of 165 participants were recruited for the experiment. However, 12 participants were unable to finish the independent tests of PM and of planning because of technical problems, so their data were excluded from subsequent analysis. A final total of 153 participants (95 women and 58 men) were included in all subsequent analysis. We describe below the rationale and the procedure followed to generate scores.

Tests and procedure

All testing was conducted over a two-hour session, which was split evenly between the EVET procedure in the first hour and the individual differences measures in the second hour. Except for the word recall test, all tasks were viewed on a 42-cm colour monitor and were run on a Dell XPS PC with an Intel Core Quad 2.33 GHz processor and 1GB ATI graphics card. Viewing distance from monitor was approximately 50 cm.

The Edinburgh Virtual Errands Task (EVET). The virtual environment was developed with the Valve Hammer Editor, a 3-D map creation programme freely available with the computer game Half Life 2™. The test building was rectangular in shape, with 38 rooms spread over four storeys. Each floor was accessed by two sets of internal stairs located on either side of this space. Figure 1 shows a screen shot of the concourse on the ground floor (Floor zero) and a birds-eye view of the virtual building. Where appropriate, glass wall panels were used to facilitate learning of the building structure and to make navigation easier for participants.

The participant explored the virtual environment using the keyboard and mouse. With this control method, the keyboard was used for forward/lateral/backward movement (keys “a”, “d”, “s”, and “w”) and physical actions such as picking up objects (key “e”). The mouse provided control over visual pitch (up and down) and yaw (spin left and right) perspectives. Participant position and movement within the virtual building were automatically recorded as a series of XYZ spatial coordinates, at a sampling rate of approximately 10 Hz. In addition, any actions made by the participants were recorded with a time stamp. Participants were given 8 minutes in which to complete a list of eight errands. Two different lists were used (half the participants completed one list, and half completed the other), but both lists followed the same structure. These lists are shown in Table 1. Three of the errands had two stages, for example “Pick up brown package in T4 and take to G6”. One errand was an open-ended task that asked participants to sort as many red and blue file-binders as possible into separate boxes. Participants had to decide for themselves how long they could devote to this task without compromising their overall goal of completing all the errands. The remaining four errands were simple one-step tasks (e.g., turn off lift on ground floor), and two of these had to be completed at or before a particular time. These last two tasks were removed from the analysis as they were time-based not event-based PM tasks.

Participants who used List A started the task on the ground floor, while people who used List B started on the top floor. The errands were listed in an inefficient order for completion, but participants had the opportunity to make a plan of their preferred order before they began the test.

Participants were first given the EVET instruction sheet, which detailed the nature of the task,
building layout, and rules (which they were explicitly asked to follow throughout the entire test period). The building rules required participants only to use the left stairs for travelling down and the right stairs for travelling up, to avoid entering any non-task-related rooms, and to avoid picking up any non-task-related objects. Next, participants completed the EVET practice session (approximately 5 minutes), which required each participant to follow a series of onscreen errand commands. The practice errands were to collect an object and deliver it, press a button on a wall within the environment, unlock the stairwell door with a key-code, and sort some red and blue folders into separate boxes. These practice errands were similar to, but not the same as, those used in the main testing session.

Next, participants studied their allocated errand list (Set A or B) for two minutes, after which they were given a free-recall test of the list, and the number of errands correctly recalled was recorded. This was followed by five minutes of further study then a test of cued recall, and, again, each participant was scored on the number of errands correctly recalled. After these measures of list recall were taken, participants were provided with a schematic building map and a copy of the errand list. They were asked to indicate the order in which they planned to perform the errands to achieve maximum efficiency, but they were also told that

![Figure 1. Screen shot of Edinburgh Virtual Errands Task (EVET) concourse area on the ground floor (left) and birds-eye view of the building (right) showing details of the top floor. To view a colour version of this figure, please see the online issue of the Journal.](image)

<table>
<thead>
<tr>
<th>Table 1. EVET errand lists</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Errand list</strong></td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>4</td>
</tr>
<tr>
<td>5</td>
</tr>
<tr>
<td>6</td>
</tr>
<tr>
<td>7</td>
</tr>
<tr>
<td>8</td>
</tr>
</tbody>
</table>

Note: EVET = Edinburgh Virtual Errands Task.
they could change their plan during the actual test. Upon completion of their plan, which took each participant approximately five minutes, the task list was removed along with their written plan, and they were asked again to verbally recall the errand list and building rules. Any mistakes were corrected, and this process was repeated until recall of the list was at 100% (this required approximately a further two minutes of study time). This minimized the risk that participants would fail to complete errands simply because they could not recall them. Any participants who failed to recall all of the errands after all of these procedures had been followed were asked to perform the EVET anyway, but their data were not included in subsequent analysis (this happened very rarely, and these data were not part of the original sample of 165). Including the initial learning phase, plan creating, and final checking, each participant spent approximately 14 minutes working with the errand list before starting the EVET. The EVET test lasted for 8 minutes (neither task list nor plan was present during the test). Afterwards they were scored on their free recall of all of the errands regardless of whether all had been completed.

Independent tests of cognitive resources. The Word Recall Task was based on the Capitani, Della Sala, Logie, and Spinnler (1992) general procedure and was used as an independent measure of retrospective memory. It consisted of five lists of 12 words that were read out by the experimenter at a rate of 1 per second. At the end of each list, participants were prompted to recall the words in any order. The dependent variable was total score out of a maximum of 72.

Working Memory Verbal Span required participants to verify a series of unconnected sentences while memorizing the last word of each sentence based on Baddeley, Logie, Nimmo-Smith, and Brereton (1985; Duff & Logie, 2001). All sentences were presented in sets, starting with a set of two and finishing with a maximum set size of seven. Regardless of participant performance, each set was repeated three times. All sentences were presented for three seconds and were preceded by a fixation cross for one second. Total correct recall of the sentence-final words was calculated as a proportion of maximum possible recall score (81 maximum). Sentence presentation was controlled by E-Prime 2 (Psychology Software Tools).

Working Memory Spatial Span was based on a task devised by Shah and Miyake (1996). Participants were shown a series of block capital letters that appeared consecutively on a computer monitor. They had to judge whether the letter was shown in its normal configuration or as a mirror image. Additionally the letters were shown in different orientations within a circular area, and participants had to memorize these orientations and recall them at the end of the set. The task began with a set-size of two letters and increased by one letter each time to a maximum of five. All participants completed three repetitions at each set size regardless of whether they had performed previous trials successfully. Letters remained on the screen for 3 s (preceded by a 1-s fixation cross). Total correct recall was calculated as a proportion of the maximum possible score (70). Presentation was controlled by E-Prime 2 (Psychology Software Tools).

The Travelling Salesperson Task (TST) required participants to imagine they were a salesperson who had to visit several target locations in the shortest distance possible. As this task involved the planning of routes between specific locations, we used this as an index of planning ability. In our version, cities were represented by a 5 × 5 array of coloured shapes (created using Matlab 7.1). At the bottom of each array was an information bar that contained nine coloured shapes, with the first labelled “Start/End” and the rest “Target Locations”. Participants were asked to plan the shortest route that connected all the destinations (assuming straight-line distances) and to use the mouse to click on each of these target locations in turn. When participants clicked on a location it disappeared from the information bar at the bottom of the screen, leaving only those that had yet to be visited. Participants completed two practice arrays before the main test, the first containing only targets (no distractors) and the second with the full array. They were then given 10 test arrays, each of which only had one optimum solution for the set of target locations—
this was calculated using an algorithm for travelling salesman problems (Kirk, 2007). Performance was scored as the proportion of distance longer than the optimum, averaged across the 10 arrays.

The Breakfast Task was devised by Craik and Bialystok (2006), who kindly provided a copy of the computer program. It was a simulation of the task of cooking breakfast, with different screens showing different foods and a main screen where participants had to set a table by using the mouse to drag and drop items of cutlery into place settings. Each food was required to be cooked for a particular length of time (2 minutes to 5.30 minutes), and it was the participant’s task to make sure they were all ready at the same time. Therefore, they first had to click on an icon of the food with the longest cooking time (i.e., 5.30 minutes) as shown beside the virtual table. This took them to the screen showing the food along with a timer bar. They clicked on the food icon to start the timer, which showed the progression of cooking. They then had to return to the main screen and continue to move cutlery to the virtual place settings until it was time to start the food with the next longest cooking time. This continued until the time at which all the foods should be ready. Participants then had to visit each screen to stop the cooking of each food. Prior to the actual test, participants were given a simple practice scenario involving only two breakfast foods. The outcome measure was the average deviation between the actual start time for each food and the time that it should have been started. As the task primarily involved prospective memory (for starting each of the foods at the correct time while engaged in another task, table setting), it was taken as a measure of PM ability that was independent from the EVET.

Results

Results for overall performance on errand completion are reported elsewhere (Logie et al., 2010). Here, we focus on prospective memory data that were not included in that previous report. We describe below the rationale and the procedure followed to generate the four main outcome measures.

PM error scoring

The PM error measure relied on the common EVET situation of participants walking past a room that they should have entered to complete an errand. If at the cued recall at the end of the session the participant could still successfully recall that intention, then this was marked as a PM error. Although this approach to PM assessment appears very different from that used in the typical Einstein–McDaniel paradigm, the two are equivalent in all important respects. In both cases, the participant has been asked to form an intention, with a specific action to perform upon encountering a specific cue. During this retention period, the participant is engaged in an ongoing task (navigation) that demands attention. Furthermore, all participants were checked for failures of retrospective memory for the tasks they were asked to perform. However, by allowing the participant free movement we are attempting to create realistic PM scenarios, in contrast to the more common practice of the experimenter prescribing the exact cue context from the start. PM error score was calculated as the number of errors divided by the number of cues encountered.

EVET travel time

This indicated the total amount of time each participant spent travelling in the EVET building. Time spent in a room was excluded (i.e., completing a specific errand), so it was predicted that this measure would directly index each participant’s ability to efficiently navigate their path through the virtual building.

Errand follow score

This score was designed to highlight the overlap between planned and actual errand performance for each participant. Furthermore, it indexes the relationship between encoding and retrieval that is a function of the choices made by the participant during the test. The correspondence between these errand orders was based on allocating one point for each errand that was conducted in the same position or sequence as planned. The follow score was calculated by dividing total overlap points by number of tasks completed.
Plan efficiency
We identified the optimum plan by calculating the minimum distance required to complete all eight errands, while following the building rules and working within the time constraints imposed by the two time-based errands. This calculated optimum plan was validated by finding it matched with the average task rank order of the 5 highest performing subjects (see Logie et al., 2010).

Descriptive statistics and intercorrelations among all these measures are shown in Table 2. Planning efficiency correlated with better overall PM performance—that is, participants with efficient plans tended to have fewer PM errors. However, this relationship did not hold when controlling for whether participants actually followed their plan \( r = -0.09, p = 0.29 \). In contrast, the partial correlation between the plan-following measure and PM performance (when controlling for plan efficiency) was significant \( r = -0.33, p < 0.001 \). This relationship suggests that participants who stuck with their original plan (good or bad) tended to have fewer PM errors than participants who changed their plan online, even if the change resulted in a plan that was closer to the optimum. The role of spatial working memory is highlighted through a significant relationship with EVET travel time \( r = -0.19, p < 0.02 \), whereas no significant relationship was found between EVET travel time and verbal working memory capacity. This is consistent with domain-specific spatial working memory resources linked with navigation around the building.

Additional analyses focused on examining which of the five independent measures contributed unique variance to the prediction of PM errors. This was carried out using multiple regression techniques, and the results of multiple linear regression with backwards stepwise elimination measures are shown in Table 3. The regression model showed that independent measures of planning ability (TST), PM (breakfast task), and verbal working memory span were reliable predictors, while neither spatial working memory nor the word recall task had any unique relationship with number of PM errors. The failure of spatial working memory performance to act as a reliable predictor argues for domain-specific working memory processes and highlights a role for verbal working memory capacity in successful prospective memory as assessed by the multiple errands methodology.

Discussion
The aim of this paper was threefold: first, to explore the role of planning in successful PM performance;
Table 3. Results of multiple regression with backwards stepwise elimination to assess the contribution to common variance between prospective memory errors and scores on five different measures of mental ability

<table>
<thead>
<tr>
<th>Variable</th>
<th>B</th>
<th>SE B</th>
<th>Beta</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>TST</td>
<td>62.05</td>
<td>17.30</td>
<td>.28</td>
<td>3.59</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Verbal working memory</td>
<td>−0.19</td>
<td>0.09</td>
<td>−.17</td>
<td>−2.20</td>
<td>.001</td>
</tr>
<tr>
<td>Breakfast Task</td>
<td>−0.13</td>
<td>0.07</td>
<td>−.13</td>
<td>−1.73</td>
<td>.09</td>
</tr>
</tbody>
</table>


second, to investigate whether domain-specific or domain-general working memory processes are at play in PM; and, finally, to validate a novel approach to PM assessment using cost-effective virtual reality software. With regard to the first aim, the importance of the planning task in the regression model is in line with Kliegel et al. (2000) who highlighted the role that planning has in successful PM. One novel finding here is the correlation between plan-following and PM performance in the new paradigm. It would appear that following a plan, rather than changing the plan online, provides some PM retrieval support. Although previous research has demonstrated this relationship (Kliegel et al., 2000), our study is the first to show this effect of planning on PM performance in participants whose choices in the environment affect the match or the discrepancy between the context for encoding and the context for retrieval.

With regard to our second aim, the significance of verbal working memory highlighted in our regression model is in accordance with previous work that has demonstrated a link between working memory span and higher PM performance (e.g., Brewer et al., 2010; Einstein et al., 2000; Smith, 2003; Smith & Bayen, 2005; West & Craik, 2001). Meilinger, Knauff, and Bülthoff (2008) reported a study where participants learned specific routes through a virtual city while performing a verbal, visual, or spatial concurrent task. During a subsequent test phase, participants who had performed a verbal or spatial task during the learning phase were more likely to get lost. This finding is in line with our results, which demonstrated a relationship between lower PM errors and higher verbal working memory capacity. The nonsignificance of spatial working memory capacity as a predictor could be interpreted as evidence that PM is primarily a cognitive process that is represented in the verbal domain, and no spatial representations are required for successful performance. Alternative explanations are possible; the first relates to the PM error measure itself. By only examining behaviour around the PM errand location itself, we are, in effect, ignoring the navigational effort it took to get there. This interpretation is supported by the significant relationship that spatial, but not verbal, working memory capacity had with our index of movement efficiency (EVET travel time). The removal of navigational effort from our measure of PM performance addresses the disparity with the Marsh and Hicks (1998) finding of interference from both spatial and verbal concurrent tasks on PM performance. However, it is important to clarify the distinction between the Marsh and Hicks study and our approach. In addition to our PM error measure not indexing spatial ability, the absence of spatial working memory as a predictor of PM performance does not indicate there is no functional relationship. Rather, it states that spatial working memory capacity cannot predict PM performance in this version of the EVET task. However, it may be that only a minimal level of spatial working memory is required for the task. This would make a measure of the maximum spatial working ability of each participant insensitive to variations in EVET performance (see Logie & Baddeley, in press, for a discussion). A different virtual environment—for example, one familiar to the participant—could result in spatial working memory being a better predictor, if planned errand order is based on a route rather than solely a list of errands. An everyday example here might be planning for shopping in a familiar supermarket where the locations of specific goods are known. Also of note is that we have used a two-dimensional spatial task as an independent
measure of working memory ability. Considering the three-dimensional nature of the EVET, a spatial wayfinding task might be more suitable for use in future studies (e.g., see Wolbers & Buchel, 2005).

Similarly, based on the above premise, the failure of the retrospective memory measure (the recall task) as a significant predictor was not unexpected, given that our index of PM performance only considered tasks that the participant could successfully recall after the EVET. Specifically, we actively attempted to separate prospective from retrospective failures. This is in line with the standard prospective memory research methodology in which participants are asked post test to recall their instructions. Therefore, since PM failures cannot be attributed to retrospective failures, it is not surprising that the retrospective memory was not a significant predictor of PM performance. A further manipulation could incorporate concurrent task methodology that may highlight the resources required. For example, an interesting question for a future study is whether concurrent performance of a verbal or spatial orientated task would selectively interfere with PM performance during the EVET.

The final aim of this paper was the development and validation of a novel methodology, which has been demonstrated, in part, by the planning effects reported above. By allowing free movement, we are creating a larger and more complete picture of the factors contributing to PM performance. It is hard to envision how the standard laboratory paradigm could address the relationship between encoding and retrieval as conceptualized in this paper. EVET incorporates advantages of a naturalistic PM paradigm with experimental control of the environment. It also allows for very much shorter testing time than is possible with naturalistic paradigms that may take several hours (e.g., Shallice & Burgess, 1991), or several days (e.g., Rendell & Craik, 2000). Like the typical laboratory PM task, EVET has an ongoing task of navigation around the virtual building. However, our experimental platform is sufficiently flexible that, in future studies, it could readily be used to investigate other research questions such as the impact on PM of different additional ongoing embedded tasks (e.g., Scullin, McDaniel, & Einstein, 2010; Smith, Hunt, McVay, & McConnell, 2007). A potential caveat might be whether this novel multiple task approach to the study of PM can be compared with results from studies that measure PM using more traditional single-task PM methodologies. As noted, one of our aims was to introduce a new kind of paradigm that can address questions about PM performance that cannot readily be addressed by traditional PM laboratory paradigms. A further aim was to incorporate the experimental control that is missing from naturalistic PM paradigms. It is worth noting that Burgess, Veitch, de Lacy Costello, and Shallice (2000) accounted for multitasking impairments that are sequellae to frontal lobe damage, in part by partitioning specific measures of PM contributions to multitasking performance. This work shows that not only is PM a key component of successful multitasking, but that it can be indexed separately from other cognitive processes. Similarly, Kliegel et al. (2000) indexed PM performance on their complex PM task, which required participants to perform multiple tasks. Therefore we see our results as being complementary, but adding to those obtained from typical laboratory paradigms.

A potential implementation of this methodology would be to create virtual analogues of real-world locations and explore the effect of location familiarity on PM. Titov and Knight (2001) have shown that familiarity with an environment improves prospective memory performance. These authors developed a video paradigm that attempted to replicate an everyday shopping experience that manipulated context by using two films; both show very similar shopping streets, one familiar and the other unfamiliar. The familiar location produced significantly more successful PM responses than the unfamiliar. They argued that, although the two videos were in principle identical, location familiarity (and consequent availability of contextual cues) enhanced planning and organization of the PM tasks. However, using video material of actual locations results in several methodological issues. In addition to the difficulties involved with identifying suitable intentions, cues, and responses
from video material, the suggestion of “movement” is dictated by the serial order of video clip presentation. With such passive video presentation, variability in navigational strategies between individuals cannot be assessed. Moreover, research has shown that navigator movement strategies (Hölscher, Büchner, Brösamle, Meilinger, & Strube, 2007; Hölscher, Meilinger, Vrachliotis, Brösamle, & Knauff, 2006) and target orientation (Frankenstein, Meilinger, Mohler, & Bülthoff, 2009) are heavily influenced by their degree of familiarity with the environment. This new approach may allow researchers to explore how and when a priori knowledge of the structural and functional aspects of a location is used in PM. The important role of “cue specificity” in PM, as highlighted by Ellis and Milne (1996), provides a theoretical framework for future studies into the role of location familiarity and PM performance.

The role of planning in successful performance was highlighted by the significance of planning (TST) in the regression model and the importance of plan following for PM error behaviour. We have not yet addressed the question of what processes are involved when people form a plan for a future activity (in contrast to the plan-following measure discussed above). From our current data set we can see that among the independent measures of cognition, only word recall had a significant correlation with planning efficiency. The absence of a correlation with either the working memory or planning tasks is unexpected. One explanation centres around the difficulty of creating an EVET plan and opportunities for elaborating plans (see Kliegel et al., 2000). In future studies with EVET, planning difficulty could be manipulated by allowing the participant to determine their preferred level of plan elaboration.

In conclusion, by using a novel methodology for examining PM in a healthy young adult population, the data demonstrate how participant-driven plans are implemented and the how their implementation affects PM performance. Furthermore, our results are consistent with domain-specific cognitive resources, not a global attentional resource, for successful PM performance.

REFERENCES


