Open access • Journal Article • DOI:10.1037/0882-7974.19.1.27

A meta-analytic review of prospective memory and aging. — Source link

Julie D. Henry, Mairi S. MacLeod, Louise H. Phillips, John Robertson Crawford

Institutions: University of Aberdeen

Published on: 01 Mar 2004 - Psychology and Aging (American Psychological Association)

Topics: Retrospective memory, Prospective memory, Recall, Free recall and Cognition

Psychology and Aging
2004, Vol. 19, No. 1, 27–39

0882-7974/04/$12.00 DOI: 10.1037/0882-7974.19.1.27

A Meta-Analytic Review of Prospective Memory and Aging

Julie D. Henry, Mairi S. MacLeod, Louise H. Phillips, and John R. Crawford
University of Aberdeen

A meta-analysis of prospective memory (PM) studies revealed that in laboratory settings younger participants outperform older participants on tests of both time- and event-based PM (*r*s = −.39 and −.34, respectively). Event-based PM tasks that impose higher levels of controlled strategic demand are associated with significantly larger age effects than event-based PM tasks that are supported by relatively more automatic processes (*r*s = −.40 vs. −.14, respectively). However, contrary to the prevailing view in the literature, retrospective memory as measured by free recall is associated with significantly greater age-related decline (*r* = –.52) than PM, and older participants perform substantially better than their younger counterparts in naturalistic PM studies (*r*s = .35 and .52 for event- and time-based PM, respectively).

Much research on cognitive aging has focused on retrospective memory, or recollection of past events (for a review, see Light, 1991), and almost invariably it has been reported that substantial deficits in this aspect of cognition are associated with normal aging. However, interest has increasingly shifted to investigating *prospective memory* (PM), that is, memory for future intentions. Relative to retrospective memory, PM is believed to be more dependent on internal control mechanisms (Craik, 1983, 1986). This is because, according to Craik's (1986) theoretical model, the act of recollection is dependent on reconstructing events in memory, and it is suggested that this process must be guided either by external cues, or in their absence, self-initiated cues. In retrospective memory tasks explicit prompts to recall are provided by the experimenter, whereas in PM tasks the cue is not an explicit request for action, but instead it requires either interpretation of a cue or an internal impetus. It has often been argued that this requirement for self-initiated remembering means that PM tasks should be more susceptible to the effects of adult aging than retrospective memory tasks (e.g., Craik, 1986; Maylor, 1995; McDaniel & Einstein, 2000).

A distinction has been made between time- and event-based PM (TBPM and EBPM, respectively; Einstein & McDaniel, 1990; Einstein, McDaniel, Richardson, Guynn, & Cunfer, 1995). Whereas the former requires the participant to perform a specified behavior at a particular time, for the latter the required behavior is prompted by an external cue. Of the two types of PM, TBPM is therefore believed to be the most reliant on internal control mechanisms because, assuming no external mnemonic aid is used, TBPM is more dependent on self-initiated mental activities, such as active time monitoring (d'Ydewalle, Bouckaert, & Brunfaut, 2001). Thus, of the two types of PM, it has been argued that TBPM should be especially sensitive to age-related decline (Einstein et al., 1995; Maylor, 1995).

However, although many studies have found evidence of age-related decline on tests of PM, this is not sufficient to infer the presence of a *differential* deficit in this aspect of cognition, particularly because it has been suggested that PM tasks also involve a retrospective component (Cohen, West, & Craik, 2001; McDaniel & Einstein, 1992). Successfully performing a PM task requires not only recall of something that is to be done in the future, but also retrieval of what it is that needs to be done, and this latter component clearly implicates retrospective memory. Indeed, Cherry et al. (2001) reported that two measures of retrospective memory accounted for 68% of the age-related variance in PM performance. Thus, if PM deficits are to be regarded as differential deficits, then it is necessary to demonstrate that they exceed deficits for tests of retrospective memory.

Empirical studies do not suggest that older adults are impaired in all aspects of PM, and in particular, older adults tend to perform as well or better than their younger counterparts in TBPM tasks that are carried out in naturalistic rather than laboratory settings. These tasks include measures in which the participant is required to telephone the experimenter at a specific time over 4 weeks (Devolder, Brigham, & Pressley, 1990), 3 weeks (Poon & Schaffer, 1982), 2 weeks (Moscovitch, 1982), and 5 days (Maylor, 1990); mail postcards to the experimenter (Patton & Meit, 1993); and periodically log the time on an electronic organizer (Rendell & Thomson, 1993, 1999; Sawyer, 1988). In addition, older adults tend to show better TBPM for attending appointments (Martin, 1986).

However, young and old adults may differ in their motivation to successfully complete PM tasks outside of the laboratory (Patton & Meit, 1993; Rendell & Craik, 2000). An important determinant of the magnitude and direction of age effects on PM tasks is likely to be the presence or absence of external aids to cue the PM event. In naturalistic tasks, older adults are usually able to set up external cues to act as reminders. The reliance of older adults on external aids is perhaps not surprising as they tend to report more everyday memory failures and more concern about this (Cavanaugh, Grady, & Perlmutter, 1983). Thus, when required to make prearranged phone calls, for instance, older adults use "conjunction cues," such as placing the action to be remembered with another routine event such as having a meal (Maylor, 1990). Naturalistic studies offer

Julie D. Henry, Mairi S. MacLeod, Louise H. Phillips, and John R. Crawford, School of Psychology, King's College, University of Aberdeen, Aberdeen, United Kingdom.

Correspondence concerning this article should be addressed to Julie D. Henry, School of Psychology, King's College, University of Aberdeen, AB24 3HN United Kingdom. E-mail: j.d.henry@abdn.ac.uk

little control over the use of such external aids, and when they are prevented, age-related benefits are typically reduced (Maylor, 1990) and in some cases no longer significant (Patton & Meit, 1993), although it is important to note that d'Ydewalle and Brunfaut (1996) found an age benefit on naturalistic PM tasks even when restrictions on external aids were imposed.

Despite greater experimental control compared with naturalistic studies, laboratory studies of PM have yielded more inconsistent results. Most laboratory-based PM tasks are in fact dual tasks, where the participant carries out an ongoing task and must occasionally respond to PM cues. Thus, in EBPM and TBPM laboratory studies, participants are typically required to exhibit a particular behavior in response to a specific event or at specific time intervals, respectively, while simultaneously performing an ongoing task.

Most laboratory EBPM studies report that older adults are substantially impaired relative to their younger counterparts (Cherry et al., 2001; Dobbs & Rule, 1987; Kidder, Park, Hertzog, & Morrell, 1997; Mantyla & Nilsson, 1997; Maylor, 1993, 1996; Park, Hertzog, Kidder, Morrell, & Mayhorn, 1997; West & Covell, 2001), although some have revealed no evidence of age-related decline (Einstein & McDaniel, 1990; Einstein et al., 1995; Li & Blackburn, 1994). In comparison to EBPM, studies of TBPM have more consistently reported age-related deficits (d'Ydewalle et al., 2001; Einstein et al., 1995; Park et al., 1997), and this may reflect the presumed greater reliance of TBPM on self-initiated retrieval processes.

Indeed, it has been reported that during performance of TBPM tasks, older adults tend to monitor the clock less often than their younger counterparts (Einstein et al., 1995; Park et al., 1997), and this may be attributable to deficits in attentional resources, or poorer estimation of time intervals (Einstein et al., 1995). Maylor (1998) also found that older adults reported thinking less about the PM component of such tasks than middle aged or younger adults, suggesting that younger people outperform their older counterparts because they can keep the intention of carrying out the PM task at a state of higher activation (Koriat, Ben-Zur, & Nussbaum, 1990).

As noted above, much of the literature has focused on differences in age effects on TBPM relative to EBPM. Where both TBPM and EBPM have been examined in the same participants, age-related deficits are more consistently associated with TBPM (Einstein & McDaniel, 1996; Einstein et al., 1995) and are typically more pronounced for TBPM conditions (McDaniel & Einstein, 1992; Park et al., 1997). However, in some studies, older adults perform better on TBPM relative to EBPM (d'Ydewalle, Luwel, & Brunfaut, 1999; d'Ydewalle, Utsi, & Brunfaut, 1996).

Along with the distinctions between TBPM–EBPM and naturalistic–laboratory tasks, a number of other dimensions have been highlighted as important in determining the size and direction of age effects on PM tasks. Particularly influential is McDaniel and Einstein's (2000) multiprocess framework, in which it is argued that event-based prospective remembering can be supported by either strategically monitoring the environment for the presence of the prospective cue, or by relying on the prospective cue to automatically prompt the target action. Because aging is presumed to be associated with deficits in attentional capacities, this framework therefore predicts that the magnitude of age effects on EBPM tasks will be determined by the extent to which the task depends on automatic processing versus controlled resource-demanding processing. McDaniel and Einstein (2000) suggested that the following factors may increase the strategic, controlled demands of EBPM paradigms, and thus may increase any age deficits: (a) nondistinctive PM cues, (b) a weak association between the cue and the intended action, (c) a highly attention-demanding or engaging ongoing task, or (d) processing of the PM cue being peripheral to the processing carried out in the ongoing task.

Other authors have also emphasized the importance of various factors in determining the magnitude of aging effects on PM. It has, for example, been suggested that the nature of the ongoing task is probably important (d'Ydewalle, 1995; d'Ydewalle et al., 1999; Einstein, Smith, McDaniel, & Shaw, 1997; Martin & Schumann-Hengsteler, 2001), with age effects likely to be reduced if the cognitive demands of the ongoing task are relatively low (see Maylor, 1995). The self-initiated processing requirements at retrieval also vary across studies in terms of the stimulus properties of the prospective cue, and this may account for some inconsistencies (e.g., whether high typicality vs. low typicality prompts are used; Cherry et al., 2001; Mantyla, 1993, 1994). Moreover, it has been found that older adults with high educational achievement and verbal abilities (Cherry & LeCompte, 1999) and higher fluid intelligence (Cockburn & Smith, 1991) tend not to be as impaired on EBPM tasks as those with lower abilities. Thus, it is perhaps not surprising that Cherry et al. (2001) argued that "Comparisons across studies in which different methodologies were used invite interpretative caution" (p. 191).

In this article, we address four issues relating to age effects on PM tasks, using meta-analytic techniques. First, it has been argued that age deficits should be greater for TBPM compared with EBPM tasks, because the former is more dependent on strategic processing and internal cues to action (Einstein et al., 1995; Maylor, 1995). The first aim is therefore to compare the effect sizes of age differences on TBPM and EBPM. Second, there is evidence to suggest age improvements in naturalistic PM tasks, but age declines on laboratory-based PM tasks. Thus, the second aim is to quantify the relative effect size of these age effects. A third issue relates to the predictions made by McDaniel and Einstein (2000) that age-related deficits should be greater where an EBPM task imposes greater demands on strategic processing. The third aim is therefore to test this prediction by comparing the magnitude of age-related deficits on EBPM tasks for studies that have used experimental manipulations to vary the level of strategic demand associated with tests of EBPM. Finally, we attempt to address the prediction that the effects of aging on PM tasks should be greater than those on retrospective memory tasks because of the greater demands PM tasks place on self-initiation and controlled processing (see Maylor, 1995). Thus, the magnitude of age-related effects on PM and retrospective memory measures are compared.

Method

Literature Search

A computer-based search involving the *Web of Science, PsycLIT CD-ROM,* and *Science Direct* databases was conducted. The key conceptual terms used as search parameters were *prospective memory, intentional memory, memory for intentions,* and *future memory.* In addition, a backward citation search was also undertaken (i.e., references in each of the articles retrieved were checked). The search was completed in December 2001.

Inclusion Criteria

Studies were included if they had a research design that compared healthy young and older groups. Only groups for which the mean age of older adults exceeded 55 were considered older. Although for younger groups the mean age was permitted to vary between 18 and 59, we also additionally required that the mean age of the older group against which they were compared was a minimum of 15 years older. Also, the older participants must have been community dwelling, and each study must have included a measure of EBPM and/or TBPM. Finally, the study must have presented precise statistics convertible to the effect-size correlation (e.g., mean, standard deviation, standard error, *F* ratio, *t* test, standardized difference [*Z*]), been published, been in English, and been written in a journal. This latter criterion avoids the potential problem of the same or highly related data being reported in journals and book chapters.

PM tasks were defined as event based if participants were required to perform the PM action immediately following a prompt. Time-based tasks had to be conducted at a particular time, which for some studies was at a specified time interval after a cue. In Einstein, McDaniel, Smith, and Shaw's (1998) study, for instance, participants were required to perform an action whenever they began a new task, but not within the first 30 s of the task. In the present study, this was considered to be a measure of TBPM because the participant was required to remember to perform the target action within a specified time interval. In contrast, Einstein, McDaniel, Manzi, Cochran, and Baker's (2000) study in which participants were required to perform an action if they encountered a particular word, but to delay that action until they reached the trivia question phase of the trial, was classified as event based. Although as with Einstein et al.'s (1998) study there was a delay between the PM cue and the PM target action, only in Einstein et al.'s (2000) study was there a second external "prompt" that signified that the action was to be immediately performed (i.e., encountering the trivia phase of the trial). It should be noted that altering either of these classificatory decisions would not alter any of the conclusions reached in the present study.

Naturalistic studies were defined as those in which the PM task was carried out during the participants' everyday life, such as phoning an experimenter from their home at specified times or attending appointments.

The strategic load analyses required identifying studies in which experimental manipulations were used to vary the level of controlled strategic demands in EBPM tasks, with McDaniel and Einstein's (2000) multiprocess framework used to guide classificatory decisions. Where studies included more than one manipulation, only the conditions with the highest and lowest levels of strategic demand were permitted to contribute to these analyses. These will be referred to as high-demand EBPM and low-demand EBPM, respectively, throughout. In addition, only studies in which the participants in the high- and low-cognitive demand groups were independent of one another were included in these analyses.

Briefly, Cherry et al. (2001) subdivided participants in Study 1 and Study 2 on the basis of the specificity of the prospective cue. Participants in the general condition were told that the cue was a word from a particular taxonomic category, whereas those in the specific condition were told that the cue was a specific word. These conditions were categorized as high- and low-demand EBPM, respectively. In Study 3, in addition to cue specificity, cue typicality was manipulated (i.e., how representative particular target words are of their respective target categories). General atypical cues and specific typical cues were regarded as imposing high and low levels of strategic demands, respectively.

In Study 1, Einstein, Holland, McDaniel, and Guynn (1992) manipulated the PM load by varying the number of cues used to prompt an EBPM task embedded in a short-term memory task. The requirement to perform a PM action in response to one specific target was classified as low in terms of strategic demand, whereas the condition in which the PM action was prompted by four distinct targets was classified as high. In Study 1, Einstein et al. (1997) varied the attentional demands of the background task (standard vs. demanding); these conditions were coded as low and high, respectively. In Study 2, level of attentional demand during encoding and retrieval were independently manipulated. The combination of a standard ongoing task with standard retrieval was classified as low in strategic demands, whereas a demanding ongoing task with demanding retrieval was classified as high.

Kidder et al. (1997) independently varied the level of difficulty of the ongoing task (two vs. three words to be recalled) and PM load (one and three target patterns served as cues). Where the difficulty of the ongoing task and the PM load was low (i.e., two words had to be recalled, and there was only one PM cue), this was considered to be low in strategic demands. The high-demand condition was therefore when both the difficulty of the ongoing task and the PM load were increased. Finally, Park et al. (1997) focused on PM performance as a function of event density. Participants were required to perform a PM task every time a PM prompt appeared. For both 6- and 12-event conditions, the task took 12 min. However, for the former the PM prompt randomly appeared once within every 2-min interval, whereas for the latter, the PM prompt appeared once within every 1-min interval. Greater frequency will presumably maintain activation of the PM task, and thus in terms of level of strategic demands, the 6-event and 12-event conditions were considered to be high and low, respectively.

Finally, effect-size estimates for free recall, recognition, and vocabulary were derived from studies that also reported PM results. It should be noted that for the majority of studies the items that were to be recalled or recognized were unrelated to the PM task. Thus, typically a set of items independent of the PM task were presented, with participants informed that they would be required to later recall these items or discriminate these target items from a larger group of items (e.g., Cherry et al., 2001; Devolder et al., 1990; Einstein et al., 1992; Rendell & Thomson, 1993). On only one occasion was either free recall or recognition used as the ongoing task in which the PM task was embedded (d'Ydewalle et al., 1999): In this study, free recall was used as the ongoing measure. It should be noted that excluding the d'Ydewalle et al. study in these analyses did not alter any of the conclusions reached, and thus the reported analyses include this study.

Recorded Variables

For each study, year of publication, and the number, age, gender, and education of the participants were recorded for both the younger and older groups. Twenty-six studies published between 1986 and 2001 met the inclusion criteria specified. These studies are presented in the Appendix.

Statistical Analysis

Meta-analysis is a rigorous, quantitative alternative to the traditional review process, as it involves statistical integration of results. The basis of this methodology is the effect size, a standardized statistic that quantifies the magnitude of an effect. In the present study the Pearson product–moment correlation was used, which in mathematical terms corresponds with the degree of correlation between the two variables of interest. It should be noted that because the correlation coefficient is associated with a slight bias, Fisher (1928) derived a transformation of the Pearson product–moment correlation that Snedecor and Cochran (1989) have recommended should be used during statistical analyses in preference to the untransformed statistic. However, this transformed estimate is itself associated with a bias, and in a Monte Carlo analysis Field (2001) reported that for random-effects meta-analytic models, transformed effect-size estimates produced substantial upward biases of a larger magnitude than the corresponding downward biases associated with untransformed correlation coefficients. Thus, in the present study, untransformed correlation coefficients have been used for statistical analyses.

For each construct, effects were pooled to derive an estimate of the mean, with each effect weighted for sample size to correct for sampling error. To do so, the random-effects meta-analytic model (Shadish & Haddock, 1994) was selected in preference to the more commonly used

Table 1
Demographic Summary Statistics for Prospective Memory Studies

	Younger group					Older group				
Variable	*MDN*	*M*	*SD*	Range	*N*	*MDN*	*M*	*SD*	Range	*N*
Age (in years)	20.2	22.5	7.6	19–59	1,426	69.8	70.7	4.9	59–84	1,462
Education (years)	13.9	13.8	1.0	11–16	513	12.8	13.3	2.0	10–17	546
Male (%)	38.2	37.9	15.1	0–56	584	40.0	37.5	11.5	0–50	573

fixed-effects model as it yields more generalizable parameter estimates. This is because in the fixed-effects model the mean is presumed to reflect a common underlying effect parameter that gives rise to the sample observations. However, in the random-effects model the mean represents a hyperparameter, as it allows for substantive differences beyond sampling error that differentiate the effects contributing to each respective mean. Statistically, the crucial difference between these methodologies is in the calculation of standard errors and confidence intervals (CIs), which for the random-effects model are typically much larger. For this reason the National Research Council (1992) suggested that the latter model is preferable, as fixed-effects models may lead to inappropriately strong conclusions.

It was also important to test whether the difference in the magnitude of mean effects between, for instance, TBPM and EBPM was statistically significant. However, there is no agreed method for statistically comparing mean effects by using the random-effects meta-analytic model. One issue is whether the degrees of freedom in such analyses should be based on the number of participants or the number of studies (Glass, 2000). Given that a relatively small number of studies contributes to each mean in the present work, it was considered important to maximize the sensitivity to detect differences in mean effects. Therefore, the standardized difference between the two mean effects of interest was calculated using total number in the sample as the degree of freedom. On some occasions the same participants contributed toward the mean effect for both variables to be compared (e.g., EBPM and TBPM in laboratory conditions). Although in these circumstances each participant only contributed once when determining the total sample used for inferential statistics, it should be noted that including particular subject groups more than once violates the assumption of statistical independence of effect sizes. There is, unfortunately, no elegant way to deal with this problem.

To interpret how important a particular effect was in practical terms, Cohen's (1977) guidelines were adopted. These guidelines suggest that a correlation of .1 should be regarded as small, .3 as medium, and .5 as large. In addition, effect sizes were squared and multiplied by 100 as these represent the percentage of the variance (*PV*) on a measure of interest that is accounted for by group membership (i.e., being young as opposed to being older).

Multiple effect sizes for the same construct were permitted from the same study in circumstances where more than one experiment was included in the study, or subgroups were created within a particular experiment so long as the groups differed from one another in terms of the participants sampled.[1] Thus, it is possible for a particular summary statistic (including effect-size estimates) to be based on a number of cases (i.e., where cases refer to different groups of participants) that exceeds the total number of studies (i.e., 26).

Table 1 summarizes the medians, means, standard deviations, and ranges for age, education, and gender. In total, data from 1,470 young and 1,506 older healthy volunteers were incorporated in the results. The mean age difference between the young and older groups was 48.5 years ($SD = 7.78$), with an age range from 17 to 63 years. It can be seen that the two groups are closely matched on gender but that there is a slight tendency for the younger group to be better educated. Although the upper end of the age range for the young group is quite high (59), using a more restrictive cutoff for the upper end of the young age group would not alter the conclusions of the present work.

Results

A total of 152 study-level effects were calculated: 23 for naturalistic TBPM, 9 for laboratory TBPM, 4 for naturalistic EBPM, 47 for laboratory EBPM, 20 for vocabulary, 22 for recognition, and 27 for free recall. Where more than one effect was presented for the same construct from the same group of participants, the mean of these effects was calculated, and it was this value that was entered when the overall mean for the construct of interest was calculated. This is why, for instance, although there are 23 study-level effects for naturalistic TBPM, the overall mean effect size for this construct was calculated from 18 effects. All effects were calculated either from precise descriptive statistics (i.e., means, standard deviations, or standard errors) or from the F or t values for analysis of variance (ANOVA) or t tests, respectively. Kolmogorov-Smirnov tests for each of the cognitive measures revealed that the distributions of effects did not deviate significantly from a normal distribution (Zs ranged from .37 to .92, all ps $> .05$), and thus it is acceptable to use the mean as a measure of central tendency.

Table 2 presents estimates of the mean effects and their variability. A plus sign means that older participants have performed better than younger participants, a minus sign indicates the reverse. First, it should be noted that with the exception of low-demand EBPM, all of the mean effects were highly significant ($p < .01$). All of the mean effects were at least small in magnitude, with the *PV* accounted for by group membership, ranging from 1.9% to 27.4%.

With respect to the distinction among prompt types, in naturalistic conditions younger participants were more impaired than their older counterparts on TBPM relative to EBPM, but in laboratory settings older participants were more impaired than their younger counterparts on TBPM relative to EBPM. These differences did not attain significance ($Z = 1.37$, $p = .170$, and $Z = 1.27$, $p = .203$, respectively).

However, the direction of the effects was systematically related to experimental location. Whereas both TBPM and EBPM were positively related to age in naturalistic conditions (rs $= .52$ and

[1] Partially overlapping groups were permitted on four occasions; Rendell and Craik (2000), Rendell and Thomson (1999), and Maylor (1998) compared the same young group against independent older groups; Einstein et al. (1995) compared the same older group against independent young groups.

Table 2
Mean Effect-Size Estimates and Associated Confidence Intervals (CIs), Statistical Significance, and Practical Importance

Variable	*M*	*K*	*N*	*SE*	95% CIs Lower	95% CIs Upper	*Z*	Size of *r*	PV
TBPM									
Naturalistic	.52	18	699	.03	.47	.57	20.7*	Large	27.4
Laboratory	−.39	8	613	.05	−.49	−.30	−8.2*	Medium	15.4
EBPM									
Naturalistic	.35	2	48	.11	.14	.57	3.2*	Medium	12.5
Laboratory	−.34	43	1,705	.03	−.41	−.28	−10.4*	Medium	11.7
High strategic demands	−.40	9	268	.05	−.51	−.30	−7.9*	Medium	16.4
Low strategic demands	−.14	9	268	.09	−.31	.04	−1.6	Small	1.9
Vocabulary	.40	20	1,553	.05	.30	.49	8.3*	Medium	15.7
Retrospective Memory									
Recognition	−.37	22	1,205	.04	−.45	−.29	−9.1*	Medium	13.7
Free recall	−.52	27	1,433	.03	−.58	−.46	−17.0*	Large	26.9

Note. Negative effect sizes indicate older adults performing worse; positive effect sizes indicate older adults performing better. PV = percentage of the variance; TBPM = time-based prospective memory; EBPM = event-based prospective memory.
* $p < .01$.

.35, respectively), these relationships were negative in laboratory conditions (*r*s = −.39 and −.34, respectively).

Consistent with McDaniel and Einstein's (2000) multiprocess framework, tasks that placed relatively high strategic demands were associated with significantly larger age deficits than those that imposed lower strategic demands (*r*s = −.40 and −.14, respectively; $Z = 3.33$, $p < .001$). Although the mean effect for the former did not differ significantly from laboratory TBPM ($Z = 0.18$, $p = .859$), low-demand EBPM was significantly less impaired than TBPM ($Z = 3.76$, $p < .001$). It should be noted that two studies were excluded from these analyses because the same participants were tested in both the high- and low-strategic demand conditions (Mantyla, 1994; Tombaugh, Grandmaison, & Schmidt, 1995). However, in both of these studies, higher strategic demands were also associated with larger age deficits relative to lower strategic demands (see the Appendix).

Of the retrospective memory measures, free recall was significantly more impaired than recognition (*r*s = −.52 vs. −.37; $Z = 3.52$, $p < .001$). Next, a comparison of effect sizes on free recall and PM revealed that age deficits were significantly greater for free recall than for all types of PM (all *p*s < .05). Finally, older participants performed substantially better on tests of vocabulary relative to their younger counterparts ($r = .40$).

Heterogeneity of Effect Sizes

To estimate the degree of heterogeneity of the effects contributing to each mean, we estimated the homogeneity statistic (*Q*) and the random-effects variance (σ^2_θ). Table 3 presents these values, as well as the standard deviation of random effects, and the 95% CI within which random effects can be expected to fall.

For the PM measures, only EBPM in laboratory conditions and low-demand EBPM were associated with significant heterogeneity, and thus the effects contributing to each of these means differ substantively. The random-effects variance was estimated to be .008, 0, and 0 for TBPM in laboratory, TBPM in naturalistic, and EBPM in naturalistic conditions, respectively. For high- and low-demand EBPM mean effects, the corresponding values were 0 and .04, respectively.

For both measures of retrospective memory (recognition and free recall) and vocabulary, significant values of the homogeneity statistic (*Q*) were observed, and the associated standard deviations of mean effects were .14, .11, and .19, respectively. However, for neither of the retrospective measures was the upper CI positive nor was the lower CI for vocabulary negative. Thus, retrospective memory is consistently associated with an age-related deficit, and vocabulary is consistently associated with an age-related advantage.

Testing for Publication Bias

A number of validity threats have been identified that may lead to imprecise conclusions in both nonquantitative and meta-analytic reviews. Particularly problematic is "the file drawer problem" (Rosenthal, 1979), which refers to the fact that significant results are more likely to be published than nonsignificant results (Easterbrook, Berlin, Gopalan, & Mathews, 1991; Sterling, 1959). To assess whether this bias posed a threat to the results of the present study, we constructed funnel-plot diagrams. In these diagrams, sample size is plotted against the corresponding study-level effect. If statistically nonsignificant results have been discriminated against, then there should be a relative absence of studies with small sample sizes that report weak effects. This was done for all the variables except for EBPM in naturalistic conditions, as only two effects contributed to this mean (see Figure 1; it should be noted that the intersection of the *x*-ordinate and *y*-ordinate indicates the mean effect size for each variable of interest). Although mean effects based on a relatively small number of studies can be regarded as less robust to the potential problem of publication bias than those based on larger number of studies, it can be seen that for none of the variables is there any actual evidence of this confound operating in the funnel plots constructed.

Table 3
Heterogeneity of Effect Sizes Contributing to Each Mean

Variable	*M*	*Q*	*df*	σ_θ^2	*SD*	95% CIs Lower	95% CIs Upper
TBPM							
Naturalistic	.52	14.9	17	—	—	—	—
Laboratory	−.39	13.1	7	.008	.09	−.57	−.22
EBPM							
Naturalistic	.35	0.2	1	—	—	—	—
Laboratory	−.34	105.3*	42	.025	.16	−.65	−.03
High strategic demands	−.40	4.4	8	—	—	—	—
Low strategic demands	−.14	18.7*	8	.040	.20	−.53	.25
Vocabulary	.40	106.5*	19	.035	.19	.03	.77
Retrospective Memory							
Recognition	−.37	54.9*	21	.019	.14	−.64	−.10
Free recall	−.52	64.4*	26	.013	.11	−.74	−.30

Dashes indicate that random effects variance is zero. σ_θ^2 = random-effects variance; TBPM = time-based prospective memory; EBPM = event-based prospective memory.
* $p < .05$.

Discussion

Quantifying PM Mean Effects as a Function of Prompt Type and Location

Across all PM conditions, highly significant effects were revealed, although the direction of the effect was dependent on location. Thus, whereas laboratory locations were associated with substantial age-related deficits (*r*s = −.39 and −.34 for TBPM and EBPM, respectively), naturalistic locations were associated with substantial age advantages (*r*s = .52 and .35). This suggests that even if aging is associated with a decline in the basic processes involved in PM (which is probable given the greater experimental control associated with laboratory studies), this does not translate to deficits in everyday life. This result corresponds with findings from studies of planning, which suggest that age deficits are seen on novel laboratory tasks but not on more naturalistic measures of planning ability (Garden, Phillips, & MacPherson, 2001; Phillips, MacLeod, & Kliegel, in press).

Indeed, in naturalistic conditions older participants appear able to not only compensate for any age-related decline in basic processing mechanisms but also to substantially outperform their younger counterparts. The superior performance of older adults in naturalistic conditions may reflect more experience with time management, knowledge of their memory's fallibility, fewer distractions, greater opportunity to plan how they will remember to execute the tasks, and more efficient use of PM cues.

In relation to the factors that McDaniel and Einstein (2000) proposed that will decrease age deficits on PM tasks, naturalistic tasks are likely to have relatively strong associations between the PM cue and intended action. Better task motivation among older adults (especially where the younger comparison group is comprised of undergraduates) may also be an important factor, as TBPM performance is positively related to perceived task importance (Kliegel, Martin, McDaniel, & Einstein, 2001). However, it is unclear whether motivation only affects a subset of PM studies, as it has been found that greater task importance does not improve EBPM (Kliegel et al., 2001). Alternatively, it may be as McDaniel and Einstein (2000) suggested, that the EBPM task in Kliegel et al.'s study was supported by relatively automatic processes.

Indeed, Kliegel, Martin, McDaniel, and Einstein (in press) found that there was no significant difference among high- versus low-importance PM tasks for an EBPM task presumed to rely on relatively automatic processes, irrespective of whether the ongoing task was normal (r = .24) or demanding (r = .05) in terms of its strategic load. However, when an EBPM task was used that was presumed to impose substantial demands on the strategic allocation of attentional monitoring resources, participants in the high-importance condition performed significantly better than participants in the low-importance condition. This was true of participants given either a normal or a demanding ongoing task, although the magnitude of the effect for participants given the latter type of ongoing task was larger (*r*s = .45 and .72, respectively). Thus, where PM tasks impose substantial demands on monitoring resources, it may be that the level of motivation or importance attributed to the task is important in determining the magnitude, and possibly even the direction, of any age effects observed.

Although in our study TBPM in laboratory conditions was more impaired than EBPM, this difference did not attain significance. Thus, although it is widely presumed that the former type of task imposes greater demands on self-initiation (Craik, 1986) and will thus be associated with larger age effects, the present study did not provide support for this perspective. However, a large number of variables have been identified as potential moderators of the magnitude of the age-related effect associated with EBPM (McDaniel & Einstein, 2000). These include participant characteristics such as level of verbal intelligence (Cherry & LeCompte, 1999) and aspects of methodological design such as the level of engagement of the ongoing task (d'Ydewalle, 1995; d'Ydewalle et al., 1999; Einstein et al., 1997; Martin & Schumann-Hengsteler, 2001), and as has been discussed, the level of task importance (Kliegel et al., in press).

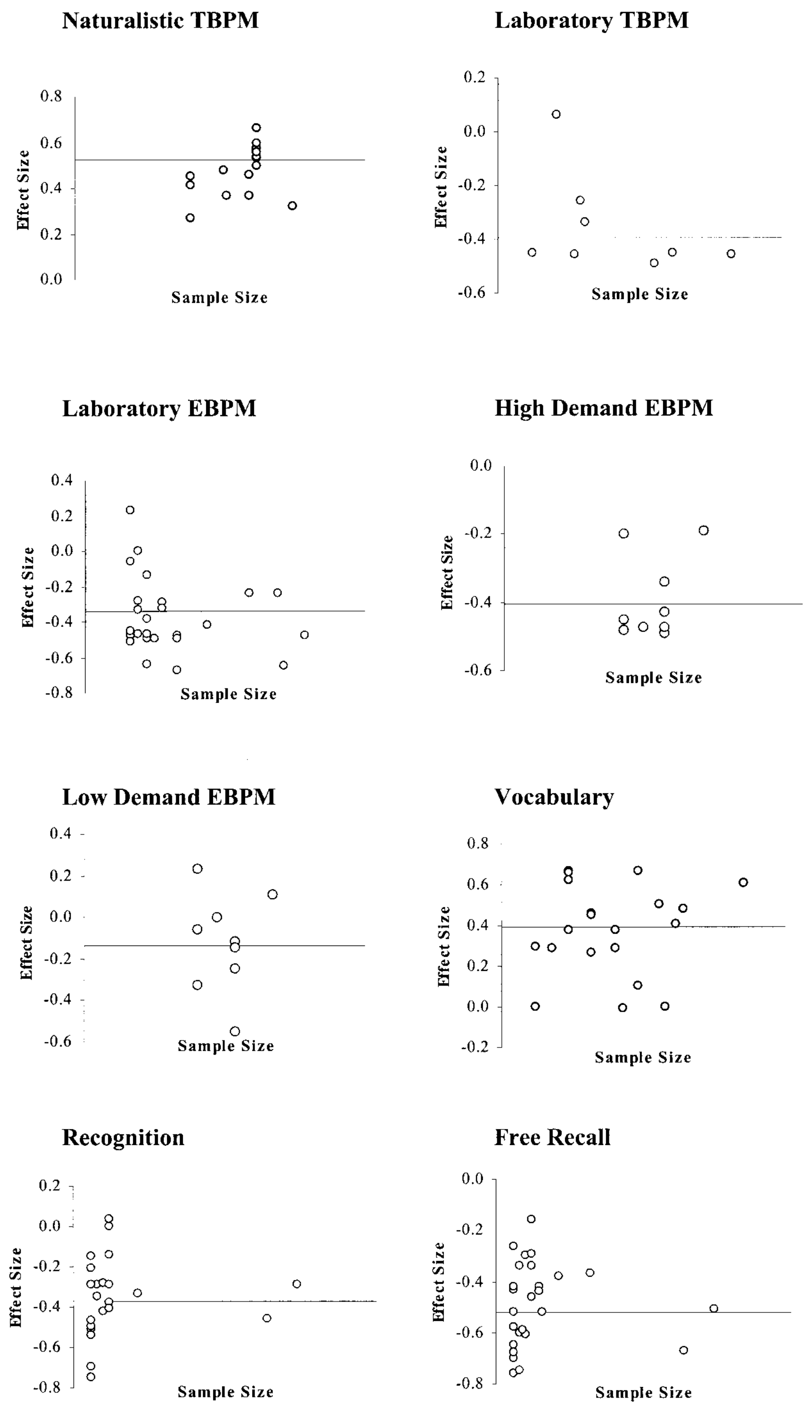

Figure 1. Funnel plot diagrams of effect sizes against sample sizes. TBPM = time-based prospective memory; EBPM = event-based prospective memory.

Too few studies have been carried out using similar manipulations of such variables as, for example, typicality of PM cues or ongoing task engagement, to allow a meta-analysis of the contribution of these particular types of experimental manipulations on the magnitude of age effects. However, we attempted to provide an overall test of the proposal (McDaniel & Einstein, 2000) that increases in the strategic demands of PM tasks, which would increase the size of age effects. This analysis provided clear evidence that in EBPM tasks age effects are partially determined by the level of strategic task demands. Consistent with McDaniel and Einstein's (2000) multiprocess framework, there appear to be EBPM tasks that are relatively automatic—and for which age differences are minimal—and other EBPM tasks that impose heavy strategic demands—and for which the age differences are correspondingly larger.

The mean deficit for EBPM tasks that imposed a high level of controlled processing ($r = -.40$) was significantly larger than the mean effect for those tasks that placed fewer demands on this aspect of cognition ($r = -.14$). Moreover, although the magnitude of the age deficit associated with TBPM ($r = -.39$) did not differ from high-demand EBPM, it was substantially in excess of the deficit for low-demand EBPM. Thus, the level of strategic demand is an important moderator of the age effect observed, and it may account for apparent discrepancies among individual studies in regard to the relative magnitude of age-related deficits for tests of EBPM and TBPM. As noted previously, whereas some studies have found that age-related deficits are more pronounced for TBPM than EBPM tasks (McDaniel & Einstein, 1992; Park et al., 1997), in other studies there is evidence of the reverse (d'Ydewalle et al., 1996, 1999). Indeed, in the present study, the homogeneity statistic, Q, for EBPM in laboratory conditions was found to be highly significant ($p < .001$). However, when subdivisions were made according to level of strategic demands, the mean effect for the high-demand subgroup was statistically homogenous, and for the low-demand subgroup, the degree of heterogeneity was substantially reduced ($p = .02$).

However, when the number of studies (k) is relatively small, it is difficult to determine whether there is significant heterogeneity in effect-size estimates (Hedges & Olkin, 1985). Thus, it may be that, although much of the variance has been removed by stratifying EBPM effects according to level of strategic demand, substantive differences between the effects contributing to the mean effects for high-demand and low-demand EBPM remains. Future research is therefore needed to more clearly delineate which particular manipulations are particularly effective moderators of the age effects observed.

It should also be noted that although several other PM mean effects were associated with nonsignificant estimates of the homogeneity statistic (Q), for TBPM in laboratory conditions and EBPM in naturalistic conditions, again the number of effects contributing to each of these means was relatively small (eight and two, respectively). Thus, it may be premature to conclude that the effects contributing to each of these respective means measure a common underlying parameter. It can be confidently concluded that, only for TBPM in naturalistic locations, which was calculated from 18 effects, the predominant source of variance between study-level effects is artifactual.

PM Relative to Retrospective Memory

Although all of the mean effects in the present study with the exception of low-demand EBPM were highly significant ($p < .01$), free recall was associated with the largest age-related deficit of all of the measures assessed ($r = -.52$). Recognition was also moderately negatively correlated with age ($r = -.37$), whereas for vocabulary a substantial age advantage was found ($r = .40$). Thus, all the mean effects were classified as at least small in magnitude according to Cohen's (1977) criteria, with correlations ranging from $-.14$ to $.52$ ($PV = 1.9\%$ to 27.4%).

Failure to find a larger deficit on either TBPM or EBPM, relative to retrospective memory, is not consistent with the prevailing view in the literature that of these types of memory, PM is particularly susceptible to age-related decline (see Craik, 1986; Maylor, 1995). Indeed, if anything, our study provides support in the reverse direction: Free recall was significantly more impaired than all measures of PM.

It might be argued that estimates of retrospective memory are larger in this study because often these measures are not entirely independent of the PM task. Although it was noted earlier that for only one study was the measure of free recall used as the ongoing task in which the PM task was embedded, and for most studies the free recall and recognition items were unrelated to the PM task, for a minority of studies, although free recall did not actually constitute the ongoing task, following completion of the dual-task PM measure participants were then unexpectedly asked to recall words that had been included in the ongoing task (Einstein et al., 1997; Mantyla, 1994). Moreover, in many studies the PM task requirements are revealed to participants prior to being asked to complete tests of retrospective memory, and this may interfere with performance on the latter. As McDaniel and Einstein (2000) noted, to maintain activation of the cue–intention association, it may be that executive resources are allocated to periodically bringing the intended action to mind.

Comparisons of the present study's results with other meta-analytic reviews of age effects on retrospective memory do suggest that the effect sizes for retrospective memory were particularly high in our study. Three other meta-analytic reviews have revealed that (expressed as the standardized differences between younger and older groups) mean age effects for free recall were $-.97$ (La Voie & Light, 1994), -.99 (Verhaeghen, Marcoen, & Goossens, 1993), and -1.01 (Spencer & Raz, 1995). In the present study, the mean effect for free recall expressed as a standardized difference was -1.22. However, it is important to note that in the current study the PM deficits expressed as standardized differences were -0.86, -0.73, -0.88, and -0.28 for laboratory TBPM, total EBPM, high-demand EBPM, and low-demand EBPM, respectively. Thus, even when the results from other meta-analytic studies are referred to, there is no evidence that PM is more impaired than retrospective memory as measured by free recall. This evidence goes against the hypothesis that age effects will be greater on tests of prospective memory (e.g., Craik, 1986; McDaniel & Einstein, 2000).

It is not entirely clear why free recall is associated with larger, or at least comparable, age deficits than measures of PM. It is possible that the calibration between retrospective and PM tasks is not very close, and thus direct comparison of the age effects for free recall and PM may not be a fair one. There are typically methodological differences between retrospective and prospective memory paradigms, and in particular, the former is typically associated with list lengths that are substantially longer than those used for the latter. This is potentially important, for it has been found that "list length" moderates the magnitude of the age-related deficits associated with tests of PM (Einstein et al., 1992). This may preclude fair comparisons at the level of the individual study as well at the level of meta-analysis. However, note that, unlike tests of retrospective memory, PM tasks are typically conducted in dual-task conditions. Despite this, the mean deficits for tests of PM found in this study were comparable or smaller than the deficits associated with free recall in the present as well as in other meta-analyses.

It is suggested that the strategic demands of free recall may have been underestimated. There is, for instance, evidence that age-related deficits in retrospective memory reflect executive impairment (e.g., Parkin & Walter, 1991), and Crawford, Bryan, Luszcz, Obonsawin, and Stewart (2000) found that executive function accounted for age-related variance in free recall and recognition, even after controlling for general cognitive ability. Moreover, increased retrospective memory load has been demonstrated experimentally to reduce PM performance in older adults (Einstein et al., 1992), thus demonstrating that there is a degree of overlap between age effects on the two forms of memory. Indeed, as noted previously, Cherry et al. (2001) reported that two measures of retrospective memory accounted for 68% of the age-related variance in PM performance, and Einstein et al. (1995) found evidence that "self-initiated retrieval processes are an important component of age-related differences *across both retrospective and prospective memory tasks* [italics added]" (p. 996).

Moreover, the present results provide clear evidence that at least some PM tasks do not place such heavy demands on effortful or executive processes as has been presumed and are thus not subject to disproportionate age-related decline. EBPM tasks that imposed low-strategic demands according to McDaniel and Einstein's (2000) multiprocess framework were associated with significantly smaller age effects than EBPM tasks that imposed relatively high-strategic demands. Our study therefore indicates that it is erroneous to regard all tests of PM, and particularly EBPM, as tests that draw heavily on self-initiated retrieval. The characteristics of the EBPM task in terms of the stimulus properties of the prospective cue, ongoing task and retrieval demands appear to be fundamental in determining the degree to which automatic versus effortful processes are invoked, and thus the magnitude of the age deficits observed. The present results are thus consistent with McDaniel and Einstein's (2000) multiprocess model that suggests that event-based prospective remembering can be supported by either strategically monitoring the environment for the presence of the prospective cue or by relying on the prospective cue to automatically prompt the target action.

References

References marked with an asterisk indicate studies included in the meta-analysis.

American Psychological Association. (2001). *Publication manual of the American Psychological Association* (5th ed.). Washington, DC: Author.

Cavanaugh, J. C., Grady, J. G., & Perlmutter, M. (1983). Forgetting and use of memory aides in 20- to 70-year-olds' everyday life. *International Journal of Aging and Human Development, 17,* 113–122.

*Cherry, K. E., & LeCompte, D. C. (1999). Age and individual differences influence prospective memory. *Psychology and Aging, 14,* 60–76.

*Cherry, K. E., Martin, R. C., Simmons-D'Gerolamo, S. S., Pinkston, J. B., Griffing, A., & Gouvier, W. D. (2001). Prospective remembering in younger and older adults: Role of the prospective cue. *Memory, 9,* 177–193.

Cockburn, J., & Smith, P. T. (1991). The relative influence of intelligence and age on everyday memory. *Journals of Gerontology, Series B: Psychological Sciences and Social Sciences, 46,* P31-P36.

*Cohen, A. L., West, R., & Craik, F. I. M. (2001). Modulation of the prospective and retrospective components of memory for intentions in younger and older adults. *Aging, Neuropsychology and Cognition, 8,* 1–13.

Cohen, J. (1977). *Statistical power analysis for the behavior sciences* (Rev. Ed.). New York: Academic Press.

Craik, F. I. M. (1983). On the transfer of information from temporary to permanent memory. *Philosophical Transactions of the Royal Society, Series B, 302,* 341–359.

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

Crawford, J. R., Bryan, J., Luszcz, M. A., Obonsawin, M. C., & Stewart, L. (2000). The executive decline hypothesis of cognitive aging: Do executive deficits qualify as differential deficits and do they mediate age-related memory decline? *Aging, Neuropsychology and Cognition, 7,* 9–31.

*Devolder, P. A., Brigham, M. C., & Pressley, M. (1990). Memory performance awareness in younger and older adults. *Psychology and Aging, 5,* 291–303.

Dobbs, A. R., & Rule, B. G. (1987). Prospective memory and self-reports of memory abilities in older adults. *Canadian Journal of Psychology, 41,* 209–222.

d'Ydewalle, G. (1995). Age-related interference of intervening activities in a prospective memory task. *Psychologica Belgica, 35,* 189–203.

d'Ydewalle, G., Bouckaert, D., & Brunfaut, E. (2001). Age-related differences and complexity of ongoing activities in time- and event-based prospective memory. *American Journal of Psychology, 114,* 411–423.

d'Ydewalle, G., & Brunfaut, E. (1996). Are older subjects necessarily worse in prospective memory tasks? In M. Georgas, E. Manthouli, E. Besevegis, & A. Kokkevi (Eds.), *Contemporary psychology in Europe: Theory, research and applications* (pp. 161–172). Gottingen, Germany: Hogrefe & Huber.

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

d'Ydewalle, G., Utsi, S., & Brunfaut, E. (1996). *Time- and event-based prospective memory as a function of age: The importance of on-going concurrent activities* (Psych. Rep. No. 201). Leuven, Belgium: Katholieke Universiteit Leuven, Laboratory of Experimental Psychology.

Easterbrook, P. J., Berlin, J. A., Gopalan, R., & Mathews, D. R. (1991, April). Publication bias in clinical research. *Lancet, 337,* 867–872.

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

Einstein, G. O., & McDaniel, M. A. (1990). Normal aging and prospective memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 717–726.

Einstein, G. O., & McDaniel, M. A. (1996). Retrieval processes in prospective memory: Theoretical approaches and some new empirical findings. In M. Brandimonte, G. O. Einstein, & M. A. McDaniel (Eds.), *Prospective memory: Theory and applications* (pp. 115–141). Mahwah, NJ: Erlbaum.

*Einstein, G. O., McDaniel, M. A., Manzi, M., Cochran, B., & Baker, M. (2000). Prospective memory and aging: Forgetting intentions over short delays. *Psychology and Aging, 15,* 671–683.

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

*Einstein, G. O., McDaniel, M. A., Smith, R. E., & Shaw, P. (1998). Habitual prospective memory and aging: Remembering intentions and forgetting actions. *Psychological Science, 9,* 284–288.

*Einstein, G. O., Smith, R. E., McDaniel, M. A., & Shaw, P. (1997). Aging and prospective memory: The influence of increased task demands at encoding and retrieval. *Psychology and Aging, 12,* 479–488.

Field, A. P. (2001). Meta-analysis of correlation coefficients: A Monte Carlo comparison of fixed- and random-effects models. *Psychological Methods, 6,* 161–180.

Fisher, R. A. (1928). *Statistical methods for research workers* (2nd ed.). London: Oliver & Boyd.

Garden, S., Phillips, L. H., & MacPherson, S. E. (2001). Midlife aging, open-ended planning, and laboratory measures of executive function. *Neuropsychology, 15,* 472–482.

Glass, G. V. (2000). *Meta-analysis at 25.* Retrieved July 2002, from http://glass.ed.asu.edu/gene/papers/meta25.html

Hedges, L., & Olkin, I. (1985). *Statistical methods for meta-analysis.* New York: Academic Press.

*Kidder, D. P., Park, D. C., Hertzog, C., & Morrell, R. W. (1997). Prospective memory and aging: The effects of working memory and prospective memory task load. *Aging Neuropsychology and Cognition, 4,* 93–112.

Kliegel, M., Martin, M., McDaniel, M. A., & Einstein, G. O. (2001). Varying the importance of a prospective memory task: Differential effects across time- and event-based prospective memory. *Memory, 9,* 1–11.

Kliegel, M., Martin, M., McDaniel, M. A., & Einstein, G. O. (in press). Importance effects on performance in event-based prospective memory tasks. *Memory.*

Koriat, A., Ben-Zur, H., & Nussbaum, A. (1990). Encoding information for future action: Memory for to-be-performed tasks versus memory for to-be-recalled tasks. *Memory & Cognition, 18,* 568–578.

La Voie, D., & Light, L. L. (1994). Adult age differences in repetition priming: A meta-analysis. *Psychology and Aging, 9,* 539–553.

Li, K. Z. H., & Blackburn, A. B. (1994). *Event-based prospective memory in young and elderly adults.* Unpublished master's thesis, University of Toronto, Toronto, Ontario, Canada.

Light, L. L. (1991). Memory and aging: Four hypotheses in search of data. *Annual Review of Psychology, 42,* 333–376.

*Mantyla, T. (1993). Priming effects in prospective memory. *Memory, 1,* 203–208.

*Mantyla, T. (1994). Remembering to remember: Adult age differences in prospective memory. *Journal of Gerontology, 49,* 276–282.

Mantyla, T., & Nilsson, L. G. (1997). Remembering to remember in adulthood: A population-based study on aging and prospective memory. *Aging, Neuropsychology, and Cognition, 4,* 81–92.

*Martin, M. (1986). Aging and patterns of change in everyday memory and cognition. *Human Learning, 5,* 63–74.

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

Maylor, E. A. (1990). Age and prospective memory. *The Quarterly Journal of Experimental Psychology, 42A,* 471–493.

*Maylor, E. A. (1993). Aging and forgetting in prospective and retrospective memory tasks. *Psychology and Aging, 8,* 420–428.

Maylor, E. A. (1995). Prospective memory in normal ageing and dementia. *Neurocase, 1,* 285–289.

*Maylor, E. A. (1996). Age-related impairment in an event-based prospective-memory task. *Psychology and Aging, 11,* 74–78.

*Maylor, E. A. (1998). Changes in event-based prospective memory across adulthood. *Aging, Neuropsychology, and Cognition, 5,* 107–128.

McDaniel, M. A., & Einstein, G. O. (1992). Aging and prospective memory: Basic findings and practical applications. *Advances in Learning and Behavioral Disabilities, 7,* 87–105.

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

Moscovitch, M. (1982). A neuropsychological approach to memory and perception in normal and pathological aging. In F. I. M. Craik & S. Trehub (Eds.), *Aging and cognitive processes* (pp. 55–78). New York: Plenum Press.

National Research Council. (1992). *Combining information: Statistical issues and opportunities for research.* Washington, DC: National Academy Press.

*Park, D. C., Hertzog, C., Kidder, D. P., Morrell, R. W., & Mayhorn, C. B. (1997). Effect of age on event-based and time-based prospective memory. *Psychology and Aging, 12,* 314–327.

Parkin, A. J., & Walter, B. M. (1991). Aging, short term memory, and frontal dysfunction. *Psychobiology, 19,* 175–179.

*Patton, G. W., & Meit, M. (1993). Effect of aging on prospective and incidental memory. *Experimental Aging Research, 19,* 165–176.

Phillips, L. H., MacLeod, M., & Kliegel, M. (in press). Adult aging and cognitive planning. In G. Ward & R. Morris (Eds.), *The cognitive psychology of planning.* Hove, UK: Psychology Press.

Poon, L. W., & Schaffer, G. (1982). *Prospective memory in young and elderly adults.* Washington, DC: American Psychological Association.

*Rendell, P. G., & Craik, F. I. M. (2000). Virtual week and actual week: Age-related differences in prospective memory. *Applied Cognitive Psychology, 14,* S43-S62.

*Rendell, P. G., & Thomson, D. M. (1993). The effect of ageing on remembering to remember: An investigation of simulated medication regimens. *Australian Journal of Ageing, 12,* 11–18.

*Rendell, P. G., & Thomson, D. M. (1999). Aging and prospective memory: Differences between naturalistic and laboratory tasks. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 54,* P256-P269.

Rosenthal, R. (1979). The "file drawer problem" and tolerance for null results. *Psychological Bulletin, 86,* 638–641.

Sawyer, G. (1988). *Remembering to remember.* Unpublished master's thesis, Monash University, Victoria, Australia.

Shadish, W. R., & Haddock, C. K. (1994). Combining estimates of effect size. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 261–281). New York: Russell Sage Foundation.

Snedecor, G. W., & Cochran, W. G. (1989). *Statistical methods* (8th ed.). Ames: Iowa State University Press.

Spencer, W. D., & Raz, N. (1995). Differential effects of aging on memory for content and context: A meta-analysis. *Psychology and Aging, 10,* 527–539.

Sterling, T. D. (1959). Publication decisions and their possible effects on inferences drawn from tests of significance-or vice versa. *Journal of the American Statistical Association, 54,* 30–34.

*Tombaugh, T. N., Grandmaison, L. J., & Schmidt, J. P. (1995). Prospective memory: Relationship to age and retrospective memory in the Learning and Memory Battery (LAMB). *Clinical Neuropsychologist, 9,* 135–142.

Verhaeghen, P., Marcoen, A., & Goossens, L. (1993). Facts and fiction about memory aging: A quantitative integration of research findings. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 48,* P157-P171.

*West, R., & Covell, E. (2001). Effects of aging on event-related neural activity related to prospective memory. *Neuroreport, 12,* 2855–2858.

*West, R., & Craik, F. I. M. (1999). Age-related decline in prospective memory: The roles of cue accessibility and cue sensitivity. *Psychology and Aging, 14,* 264–272.

Appendix

Studies Included in Quantitative Review

Study	Young		Old		Type of TBPM task	Where PM Task is EB		Notes	Effect size (*r*)				
	M age	*N*	*M* age	*N*		Ongoing task	Demand		EBPM	TBPM	Recall	Recognition	Vocabulary
Cherry & LeCompte	19	24	71	24		STM		Low-ability adults	−.31		−.16	−.29	.62
(1999)	21	24	69	24		STM		High-ability adults	.01		−.46	−.14	.66
Cherry et al. (2001)													
Study 1	22	16	69	16		STM	Low	Specific PM cue	−.55		−.34	−.35	
	22	16	69	16		STM	High	General PM cue	−.34		−.60	−.29	
	22	32[a]	69	32[a]									.46[b]
Study 2	20	20	70	20		STM	Low	Specific PM cue	.11		−.61	−.42	
	20	20	70	20		STM	High	General PM prompt	−.19		−.61	−.28	
	20	40[a]	70	40[a]									.38[b]
Study 3	21	12	69	12		STM	Low	Specific typical cue	−.32		−.58	−.29	
	21	12	69	12		STM		Specific atypical cue	−.04		−.26	−.21	
	21	12	69	12		STM		General typical cue	−.27		−.43	−.51	
	21	12	69	12		STM	High	General atypical cue	−.20		−.52	−.15	
	21	48[a]	69	48[a]									.66[b]
Cohen et al. (2001)													
Study 1	22	24	71	24		Paired associates			−.48[b]				
Study 2	22	24	73	24		As study 1			−.67[b]				
Devolder et al. (1990)													
Study 2	28	24	69	24	Naturalistic			Prediction condition		.46	−.29	.00	
	28	24	69	24	Naturalistic			Postdiction condition		.37	−.34	.04	
d'Ydewalle et al. (1999)	19	30	63	30		Questions			−.49		−.42[c]		
	19	30[a]	63	30[a]		Identify faces			−.36				
	19	30	63	29	Laboratory	Questions				−.27	−.44[c]		
	19	30[a]	63	29[a]	Laboratory	Identify faces				−.24			
Einstein et al. (1992)													
Study 1	21	12	69	12		STM	Low	1 cue; short interval	−.06		−.42	−.50	
	21	12	69	12		STM	Low	1 cue; long interval	.23		−.76	−.54	
	21	12	69	12		STM	High	4 cues; short interval	−.48		−.70	−.70	
	21	12	69	12		STM	High	4 cues; long interval	−.45		−.68	−.47	
	21	48[a]	69	48[a]									.11[b]
Study 2	20	12	72	12		STM		4 cues	−.51		−.65	−.75	.00[b]
Einstein et al. (2000)													
Study 1	19	20	69	20		Questions		No delay	−.29[b]				
	19	20	69	20		Questions		Delay	−.32[b]				
	19	40[a]	69	40[a]									.29[b]
Study 2	21	24	71	24		Questions			−.49**				.38[b]
Einstein et al. (1995)													
Study 1	18–21	12	61–78	12	Laboratory					−.45[b]			.29[b]
Study 3	20	36	66	26	Laboratory					−.34[b]			
	43	28	66	26[a]	Laboratory					−.46[b]			
Einstein et al. (1998)	20	63	71	60	Laboratory					−.45[b]			.41[b]
Einstein et al. (1997)													
Study 1	19	16	73	16		Word-rating	Low	Standard task	−.25				
	19	16	73	16		Word-rating	High	Demanding task	−.43				
	19	32[a]	73	32[a]							−.52[b]	−.46[b]	.27[b]

(appendix continues)

Appendix (*continued*)

Study	Young		Old		Type of TBPM task	Where PM Task is EB		Notes	Effect size (*r*)				
	M age	*N*	*M* age	*N*		Ongoing task	Demand		EBPM	TBPM	Recall	Recognition	Vocabulary
Einstein et al. (1997) (*continued*)													
Study 2	20	16	73	16		Word-rating	Low	Standard task/retrieval	−.12				
	20	16	73	16		Word-rating		Standard task/demanding retrieval	−.31				
	20	16	73	16		Word-rating		Demanding task/standard retrieval	.09				
	20	16	73	16		Word-rating	High	Demanding task/retrieval	−.49				
	20	64[a]	73	64[a]							−.37[b]	−.37[b]	.48[b]
Kidder et al. (1997)[d]	20	15	71	13		WM	Low	1 PM cue; low WML	.00				
	20	15	71	13		WM		1 cue; high WML	−.28				
	20	15	71	13		WM		3 cues; low WML	−.33				
	20	15	71	13		WM	High	3 cues; high WML	−.47				
	20	90[a]	71	80[a]									.61
Mantyla (1993)	20–33	16	65–79	16		Free association			−.64[b]				
Mantyla (1994)	25	18	73	18		Free association	Low	Typical PM prompt	−.28		−.59		.29
	25	18[a]	73	18[a]		Free association	High	Atypical PM prompt	−.69				
Martin (1986)	23	30	66	30	Naturalistic					.32[b]			
Martin and Schumann-Hengsteler (2001)	24	90	69	75	Laboratory					−.46[b]			
Maylor (1993)	57	43	75	43		Identify faces			−.24[b]		−.38	−.33	−.01
Maylor (1996)	59	56	76	59		Identify faces			−.48[b]				.00
Maylor (1998)	20	45	59	56		Identify faces			−.24				
	20	45[a]	76	59		Identify faces			−.65				
Park et al. (1997)													
Study 1	19	16	70	16		WM	High	1 prompt, 2 min	−.47				
	19	16	70	16		WM	Low	1 prompt, 1 min	−.14				
Study 2	20	56	70	55	Laboratory					−.49			.50
Patton & Meit (1993)													
Study 1	20	24	69	17	Naturalistic			Postcards returned		.41	−.30		
	20	24[a]	69	17[a]	Naturalistic			Returned on time		.54			
	20	24[a]	69	17[a]	Laboratory			Shut off movie		.06			
Study 3	19	22	70	20	Naturalistic			Postcards returned		.35			
	19	22[a]	70	20[a]	Naturalistic			Returned on time		.39			
Rendell & Craik (2000)													
Study 2	23	16	68	16	Naturalistic	Naturalistic		Regular PM tasks	.25	.35			
	23	16[a]	68	16[a]	Naturalistic	Naturalistic		Irregular PM tasks	.36	.54			
	23	16[a]	79	16	Naturalistic	Naturalistic		Regular PM tasks	.38	.45			
	23	16[a]	79	16[a]	Naturalistic	Naturalistic		Irregular PM tasks	.41	.36			
Rendell & Thomson (1993)	21	16	84	16	Naturalistic			1-a-day condition		.23	−.75		
	21	16[a]	84	16[a]	Naturalistic			4-a-day condition		.32			
Rendell & Thomson (1999)													
Study 1	20	30	65	20	Naturalistic			Same regular		.57			
	20	30[a]	83	20	Naturalistic			Same regular		.66			
	20	30	65	20	Naturalistic			Same irregular		.53			
	20	30[a]	83	20	Naturalistic			Same irregular		.54			
	20	30	65	20	Naturalistic			Different regular		.57			
	20	30[a]	83	20	Naturalistic			Different regular		.58			
	20	30	65	20	Naturalistic			Different irregular		.56			
	20	30[a]	83	20	Naturalistic			Different irregular		.50			

Appendix (*continued*)

Study	Young		Old		Type of TBPM task	Where PM Task is EB		Notes	Effect size (*r*)				
	M age	*N*	*M* age	*N*		Ongoing task	Demand		EBPM	TBPM	Recall	Recognition	Vocabulary
Rendell & Thomson (1999) (*continued*)													
Study 2	20	30	65	20	Naturalistic			Regimen alarm		.60			
	20	30	65	20	Naturalistic			Regimen choice		.60			
Study 3	20	175[a]	65	120[a]							−.51	−.29	
	20	175[a]	83	80[a]							−.67	−.46	
Tombaugh et al. (1995)	28	31	62	33		RM	High	No PM prompt	−.45				.46[b]
	28	31[a]	62	33[a]		RM		General prompt	−.46				
	28	31[a]	62	33[a]		RM	Low	Specific prompt	−.34				
West & Covell (2001)	19–21	16	62–84	16		Semantic test			−.38[b]				
West & Craik (1999)													
Study 1	20	24	69	24		Semantic test			−.33				.67
Study 2	24	12	72	12		Semantic test			−.31				

Note. TBPM = time-based prospective memory; PM = prospective memory; EB = event based; EBPM = event-based PM; STM = short-term memory; WM = working memory; WML = WM load; RM = retrospective memory.

[a] Participants already included in table. [b] Effect size calculated via directly from *F, t,* or *Z.* [c] Free recall was the ongoing task. [d] The mean effect for vocabulary was calculated from the total sample (90 for young participants, 80 for old participants); of these only 60 young and 52 old were administered a PM measure. *N* for each condition is an estimate, as all that is stated is that "The number of participants in each cell ranged from 13 to 16" (p. 97, Kidder et al., 1997).

Received August 12, 2002
Revision received March 17, 2003
Accepted March 19, 2003 ■