THE QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY, 2004, *57A* (2), 241–267

# Evidence for disproportionate dual-task costs in older adults for episodic but not semantic memory

Leigh M. Riby
*Glasgow Caledonian University, Glasgow, UK*
Timothy J. Perfect
*University of Plymouth, Plymouth, UK*
Brian T. Stollery
*University of Bristol, Bristol, UK*

Previous research demonstrates that older adults are poor at dual tasking, but there is less agreement on whether their decrement is worse than that predicted from single-task performance. This study investigated whether task domain moderates dual-task costs in old age. In two experiments, young and older adults retrieved either previously learned associates (episodic retrieval) or overlearned category members (semantic retrieval) under single or working-memory load conditions, using cued recall (Experiment 1) and recognition (Experiment 2) procedures. In both experiments the proportional costs of dual tasking were age invariant for semantic retrieval but were particularly marked for episodic retrieval, although the size of the age effect was reduced in recognition compared to cued recall. The data suggest that age effects in dual tasking may be domain specific.

Previous research has consistently found age-related differences in dual-task performance when absolute measures are used (e.g., Baron & Mattila, 1989; Craik & McDowd, 1987). In fact, Craik (1977, p. 391) suggests that "one of the clearest results in the experimental psychology of aging is the finding that older subjects are more penalized when they must divide their attention". Older adults' difficulties with dual tasking, whether in the laboratory or in applied settings (McDowd, Vercruyssen, & Birren, 1991), are not surprising because they have difficulties on a variety of cognitive tasks when performed alone. Thus, what is of interest is whether older adults have a deficit greater than one would predict from age changes in single-task performance (Perfect & Maylor, 2000).

Correspondence should be addressed to Leigh M. Riby, Department of Psychology, Glasgow Caledonian University, 70 Cowcaddens Road, Glasgow, G4 0BA, UK. Email: L.Riby@gcal.ac.uk


http://www.tandf.co.uk/journals/pp/02724987.html DOI:10.1080/02724980343000206

Somberg and Salthouse (1982) were among the first to examine dual-task costs in older adults as a function of single-task performance. In their study, younger and older adults performed a manual reaction time task concurrently with a repetitive keying task. In addition to age-group differences in reaction time in both single- and dual-task conditions, there was an age by condition interaction supporting the view that older adults have particular difficulty with dual-tasks. However, once the effects of single-task performance were controlled by dividing the absolute difference between task reaction time under full- and divided-attention conditions by the single-task reaction time, the age difference was removed. Thus, Somberg and Salthouse argued that there was no particular problem in dual-task performance for the elderly; their poorer performance under dual-task conditions is entirely in line with their poorer performance in the single-task condition.

Other studies have provided support for the claim that older adults have no disproportional deficit of dual tasking. For instance, Tun and Wingfield (1993) reviewed the dual-task literature regarding language processing and found that dual-tasks that involve sentence verification, sentence recall, or text recall show age equivalence even when absolute measures are used. Implicit memory seems to be another domain immune to the influence of concurrent task under certain conditions. For example, Isingrini, Vazou, and Leroy (1995) compared the influence of age and division of attention on explicit and implicit memory. Participants were required to either carry out a cued-recall task or a repetition priming task (category exemplar generation) with a secondary letter detection task at learning. There were clear effects of age and division of attention for cued recall but not for category exemplar generation. A similar conclusion was reached by Light and Prull (1995) in a study of repetition priming using a word-naming task when participants were required to concurrently carry out a number addition task.

Although this evidence suggests that older adults do not have a disproportional dual-task deficit, what characterizes these tasks is that they are relatively simple (e.g., simple perceptual-tasks) or rely on relatively automatic processing (e.g., language processing and implicit memory). Under these circumstances, the requirement to divide attention does not seem especially problematic for older adults. Tsang and Shaner (1998) suggest that after taking into account age differences in single-task performance, older adults only have difficulty in dual tasking when the attentional demands of the tasks are high. In that study, participants performed a horizontal-axis tracking task combined with one of the following tasks: a vertical-axis tracking task, two versions of a spatial orientation task (two levels of difficulty), or a short-term memory task with either manual or vocal responses. In all task combinations, there was a disproportionate cost of dual tasking for the older adults. Of particular interest were the exaggerated dual-task costs as the task demands increased. Work by McDowd and Craik (1988) also provides strong evidence for disproportionate age-related dual-task costs and suggests that dividing attention is just one of several ways of increasing task complexity or task difficulty.

The issue of task difficulty has been at the heart of previous discussions as to why there are discrepancies in the results for dual-task costs in older adults. However, others have argued against the notion of difficulty. For example, Tun, Wingfield, Stine, and Mecsas (1992) investigated the effects of age and division of attention on rapid speech processing. In the single-task condition it was found that when speech rates were increased, older adults were differentially affected when they were required to immediately recall a spoken passage. However, the further requirement to divided attention did not exaggerate the effect. The authors concluded that increasing the demands of the task by varying speech rate and increasing the demands by the requirement to divide attention are independent and do not necessarily bring about the same effect.

Considering the task difficulty account of disproportionate dual-task costs for older adults, there is disagreement as to why an age effect is only observed in certain conditions. The aim of the present research is to investigate an alternative explanation for the variability in dual-task costs across the age range. Instead of focusing on task difficulty, we explore task domain as a moderator variable. That is, the present study investigated whether older adults have difficulty in combining tasks in particular domains of cognition as well as at differential levels of difficulty. The present studies focus on episodic and semantic memory.

Previous research on the effects of age and division of attention on episodic and semantic memory retrieval is inconclusive because there are no studies that have investigated both tasks in the same experiment using the dual-task cost measure. Craik and McDowd (1987) provided evidence to suggest that older adults are particularly disadvantaged in episodic memory recall compared to recognition, but they did not measure retrieval from semantic memory. In cued recall it was found that older adults were particularly disadvantaged when they were required to carry out a concurrent four-choice reaction time task. Anderson, Craik, and Naveh-Benjamin (1998) examined the effects of age and divided attention on encoding and retrieval in free recall, cued recall, and recognition. There was little or no effect of dual tasking on episodic memory retrieval performance for both groups. However, secondary-task reaction times were considerably slowed for the older adults at retrieval. Other studies have also found that when there is a requirement to divide attention, memory retrieval is not largely affected, but secondary-task costs arise (e.g., Macht & Buschke, 1983). From these results, Anderson et al. (1998) concluded that retrieval is obligatory for all age groups but it makes greater attentional demands for older adults, as indexed by a secondary-task performance decrement. The authors suggest that retrieval might not require attentional control but it does draw on attentional resource.

However Anderson et al.'s (1998) conclusion that "retrieval makes greater demands on attentional resources for older than younger adults" (p. 419) is too generalized because they did not study retrieval from semantic memory. Perfect and Rabbitt's (1993) data argue against the idea that retrieval may be difficult for older adults because they found that divided-attention costs were age invariant when using a semantic retrieval task (category member generation). This remained the case for both easy and difficult categories, thus arguing against the idea of task difficulty. That is, retrieval from difficult categories was slower and less accurate for all participants, but did not produce higher divided-attention costs for the older adults. Thus, it may be the case that it is retrieval from episodic memory that is particularly difficult for older adults, rather than retrieval in general. However, because Perfect and Rabbitt did not include an episodic retrieval condition, this hypothesis remains to be tested.

## EXPERIMENT 1

The present experiment examined two kinds of memory retrieval task: an episodic paired-associate task and a semantic category exemplar generation task. In addition, a working-memory requirement was introduced to both tasks, and the difficulty of this was manipulated to investigate the effect of increasing task demands. In traditional dual-task studies, participants are required to make a response to two separate tasks. This can lead to difficulties of interpretation because participants will differ in the trade-off they make between tasks. A number of other paradigms have been used to investigate older adults' ability to coordinate multiple tasks or processes. For instance, in the language domain age differences in dual-task costs have been investigated using the working-memory paradigm used by Daneman and Carpenter (1980). In this procedure participants are required to undertake a sentence verification task while concurrently keeping in mind a series of words. While verifying sentences participants are required to keep in mind the final word of the sentence for later recall. This is the approach developed here.

In order to minimize task trade-off effects, the working-memory element of the present experiment was integrated into the memory retrieval task by requiring participants to keep a running record of the last two or three trials. On each test trial a cue was presented to indicate which of the past *n*-trials participants should attempt to respond to. Thus completion of the secondary task was a requirement for successful performance of the main task, rather than an independent burden. (See Kirchner, 1958; Watter, Geffen, & Geffen, 2001, for a perceptual equivalent to this *n*-back task and justification for the use as a dual-task.)

In summary, younger and older participants attempted retrieval from episodic and semantic memory under no load, low-load (up to 2-back), and high-load (up to 3-back) conditions. If domain is the critical factor we expect that age differences in dual-task costs would be observed only for the episodic paired-associate task, because deficits in episodic abilities have been consistently found. Semantic memory is relatively preserved in old age, and the effect of a concurrent task on retrieval performance is likely to be less dramatic. On the other hand, if task difficulty is critical, as the majority if the literature suggests, there will be greater age-related deficits in dual-task costs in the high-load conditions, and task domain will be irrelevant.

## Method

### *Participants*

A total of 36 younger adults (mean age 22.6 years, *SD* 3.6 years) and 36 older adults (mean age 72.7 years, *SD* 5.1 years) participated in the experiment. The young volunteers were undergraduate students from the University of Bristol. The older adults were volunteers selected from a pool of older adults registered with the Experimental Psychology Department. A total of 18 participants from each age group were randomly assigned to the episodic and semantic test conditions. On average the younger adults were better educated than the older adults (16.4 vs. 13.5 years of education respectively, $t(70) = 4.5$, $p < .01$). Older adults scored more highly on the National Adult Reading Test: raw scores of 42.1 and 39.0, respectively, $t(70) = 2.5$, $p < .05$.

### *Materials*

The paired associates were constructed from high-imagery, high-frequency nouns, between four and seven letters in length, selected from the MRC Psycholinguistic Database (Coltheart, 1981). Nine word pair lists were constructed, randomly pairing such items (e.g., trial–water). The word pairs were unrelated, and each list was randomly assigned to no-load, low-load, and high-load conditions. There was one practice block and two experimental blocks for each of the three episodic conditions. The categories and category members were selected form the Belfast Category Norms (Brown, 1978). The top six ranked responses for 21 categories were used to create nine lists of 14 category–member pairs (e.g., furniture–bed).[1] Each list was randomly assigned to no-load, low-load, and high-load conditions at the beginning of each testing session. There was one practice block and two experimental blocks for each of the three semantic conditions.

## Procedure

### *Semantic memory task*

Prior to the experimental session each participant received detailed instructions with regard the nature of the tasks. A printed example with screen shots of the no-load and the load (2–back and 3–back) conditions accompanied the instructions. Participants were required to demonstrate to the experimenter that they understood the instructions by completing the printed examples.

In order to ensure that participants understood the requirements of the task, a practice session of one block for each condition was given before the start of the experimental block. In the experimental block there were two sessions each of the no-load, the low-load, and the high-load tasks, presented in counter-balanced order across participants. In each session there were 10 test trials.

*No-load condition*. In the no-load condition, the category generation task consisted of the presentation of a category name and a category exemplar cue. The category name (e.g., Animal) appeared on the computer screen with a partially completed category exemplar cue (e.g., Z****) underneath (see top panel of Figure 1 for a pictorial representation of the task). The first letter of the category exemplar was provided, and asterisks replaced each of the missing letters. Participants were instructed to respond verbally with an appropriate category member as quickly and as accurately as possible with a maximum time limit of 7 s. Participants had to produce a category member that started with the specified letter and was exactly the correct length. Category cues were presented at 7-s intervals. No feedback was given, and the final response made during the 7-s interval was recorded by a pushbutton operated by the experimenter.

*Low-load (2-back) condition*. In the low-load condition, a category was presented for 7 s. After the category disappeared a prompt was presented on the computer screen for 7 s. The prompt consisted of one of the following signals: an asterisk (*), 1–back, or 2–back. If the * prompt appeared this indicated that no response was required on this occasion. When either the 1–back or the 2–back prompt appeared the partially completed category exemplar appeared at the same time above the prompt (see lower panel of Figure 1 for an example). If the 1–back prompt appeared the participant responded by producing a member of the category that had just disappeared from the computer screen. If the 2–back prompt appeared participants were required to use the category from two positions back. Thus, participants were required to keep in mind the previous two categories presented in order to respond appropriately when the prompt appeared. The 1–back signal was included so that participants could not prepare a response to the 2–back signal in advance. The no-back signals (*) were included as fillers so that each cue was only tested once and again their inclusion made it difficult for participants to prepare a response. A total of four no-response prompts were used in each experimental session. A total of 80% of the prompts were the 2–back cue and 20% were the 1–back cue. The experimenter recorded the participants' responses.

*High-load (3-back) condition*. In the high-load condition, the task requirements were the same as those in the low-load condition except that the participants were required to keep in mind the previous

[1]A full list of materials used in the present study is available on request from the first author.

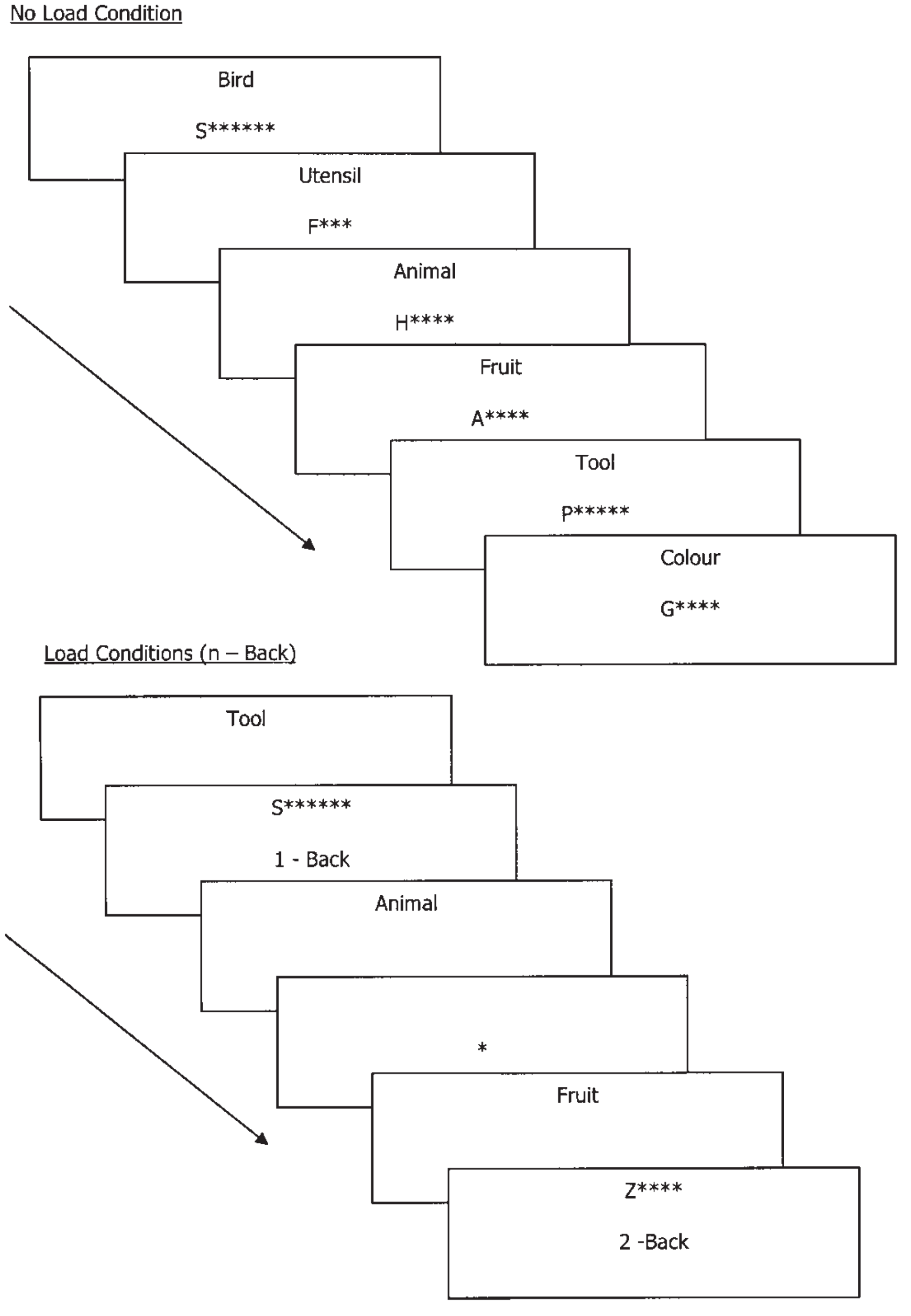


**Figure 1.** Example of the procedure. Semantic memory retrieval, no-load, and load conditions.

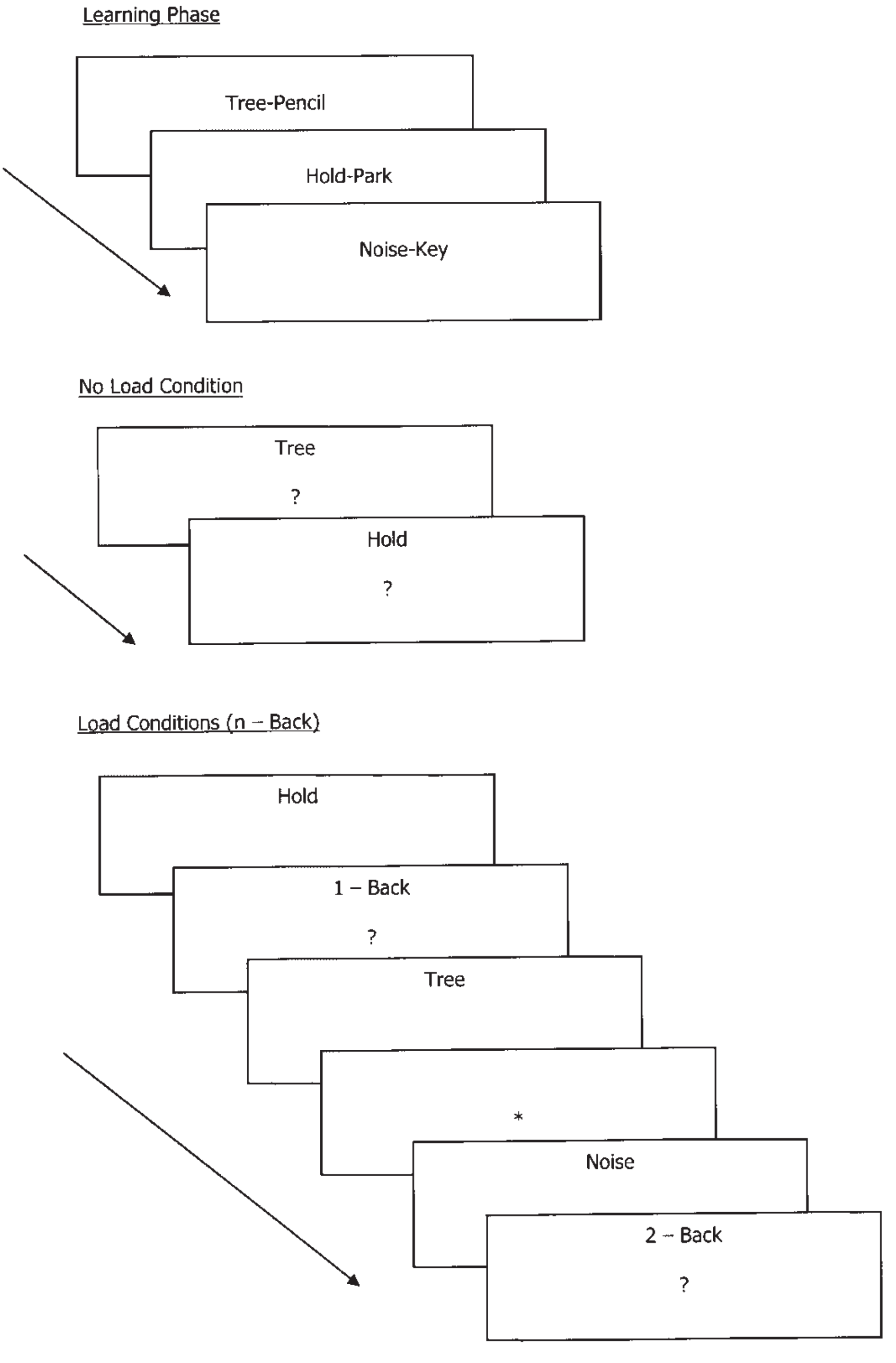


**Figure 2.** Example of the procedure. Episodic memory retrieval, no-load, and load conditions.

three categories that had been presented. On this occasion, as well as the no-response (*), the 1–back, and the 2–back conditions, participants were also required to respond to a 3-back condition. Again, there were four no-response prompts used as fillers. A total of 80% of the test trials were for the 3–back cue, with 10% for each of the 2–back and 1–back cues.

*Episodic memory task*

The procedure for the episodic memory task was identical to that for the semantic memory task, except that episodic retrieval cues were used instead of category retrieval cues (see Figure 2 for example). The episodic retrieval cues referred to a set of paired associates that were learned in a pretest training session. It should be noted that the experimenter stressed to the participant that they were required to keep in mind the cues and only attempt to retrieve the paired associate at the test trial. This was done to ensure that the task parameters remained constant across tasks. Potentially in the episodic condition participants could generate possible responses as the cues occurred rather than think *n* back at the test phase and then generate a response, although this strategy might prove more difficult.

*Pretest training*. In the pretest learning phase participants were trained on each word list before each condition. On each trial, 14 word pairs (e.g., apple–card) were presented, at 8-s intervals, on the computer screen. Participants were asked to memorize each word pair for later recall and were encouraged to use visual imagery mnemonics to aid their recall. Before the final retrieval phase, the experimenter read aloud the first word of each word pair, and participants were required to respond, in their own time, with the second word of the word pair. This training procedure continued for six trials or until participants were able to recall at least 10 of the 14 word pairs: a criterion of 71% correct. The word pairs were re-presented, in the same order, on up to six training trials until the criterion was achieved. All participants reached criterion within six presentations of the word lists. The criterion of 10 out of the 14 word pairs was selected on the basis of older adults' performance in a pilot study carried out prior to the present study.

## Results

### *Analytic strategy*

Response time and error data were collected for all load conditions for both memory tests. We begin by reporting the analyses of the mean response times and error rates. However, the main focus of the present work is on the divided-attention costs of each memory task. Thus, our main analysis focuses on the proportional increase in response time due to the memory loads. Although it is numerically possible to conduct an equivalent analysis based on error rates, it makes no psychological sense to do so, for the following reasons. Foremost is the fact that the majority of errors were time-out errors (see details below). That is, the participants did not respond within the given response deadline. Thus, these errors are highly related to the measure of response time, since slower responding would automatically lead to higher rates of time-out errors. A second reason for not wishing to focus on errors is that the basis of an error is unclear. If a participant fails to generate the correct answer on a given trial involving a memory load this could be because they are recalling from an inappropriate cue, or because they have failed to retain in mind the cues from the previous trials. Thus, they may not be operating under load when errors are made. In contrast, if the correct response is provided then we know that participants must be operating under load since they have retrieved the correct answer to the correct cue. Thus, the clearest analyses are those that are based on response latencies for correct retrievals.

Nonetheless, it remains a possibility that speed–error trade-off could account for some or all of our pattern of findings in the response time measure. Therefore, we then report analyses designed to examine whether the notion of speed–error trade-off is compatible with the findings we report. Finally, in line with other studies involving comparison of younger and older adults, we report analyses designed to determine whether the reported pattern in divided-attention costs is due to differences in baseline response time.

Because the age groups differed in education and vocabulary ability, these measures were initially entered as covariates in the analyses. However, as neither covariate was significant in any analysis, they are not discussed further. Post hoc analyses were conducted using the Bonferroni procedure with a significance level of $p < .05$ unless otherwise stated.

### *Analysis of absolute response time and error rate data*

The absolute response time and error data for the no-load and load conditions for both the episodic and semantic memory retrieval tasks are shown in Table 1.

Memory retrieval response times were analysed in a 2 (age group) × 2 (memory task) × 3 (memory load) analysis of variance. This analysis revealed that response times were generally slower for the older adults, $F(1, 65) = 58.14$, $MSE = 728{,}962$, $p < .01$, were longer in the semantic memory task, $F(1, 65) = 45.94$, $MSE = 728{,}962$, $p < .01$, and increased in line with working-memory load, $F(2, 130) = 59.18$, $MSE = 221{,}694$, $p < .01$. The interaction between memory task and age group demonstrated that older adults' response times were particularly high in the episodic task, $F(1, 65) = 9.02$, $MSE = 728{,}962$, $p < .01$. The interaction between working-memory load and age group demonstrated that older adults were particularly impaired when carrying out a memory retrieval task with a concurrent load, $F(2, 130) = 15.73$, $MSE = 221{,}694$, $p < .01$. An analysis of simple main effects, using the Bonferroni procedure, found a significant difference in the younger adults' response times between the no-load and both the low-load and the high-load conditions. The difference between the high-load and low-load response times approached significance. For the older adults there was a significant

TABLE 1
Experiment 1 response latency[a] and errors[b] for the episodic paired-associate task and the semantic category exemplar generation task, under no-load, low-load[c], or high-load[d] conditions

| | | *Retrieval task* | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | *Paired associates* | | | | *Category exemplar generation* | | | |
| | *Condition* | *Response latency* | | *Errors* | | *Response latency* | | *Errors* | |
| *Adults* | *Load* | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* |
| Younger | None | 2,123 | 281 | 3.4 | 3.1 | 3,156 | 410 | 6.8 | 2.4 |
| | Low | 2,246 | 428 | 3.5 | 2.8 | 3,546 | 531 | 9.7 | 3.1 |
| | High | 2,493 | 528 | 5.1 | 4.9 | 3,653 | 453 | 11.5 | 3.5 |
| Older | None | 2,580 | 407 | 6.3 | 1.8 | 3,476 | 426 | 8 | 2.1 |
| | Low | 3,767 | 934 | 10.3 | 3.4 | 4,283 | 725 | 15.6 | 2.1 |
| | High | 4,310 | 1,163 | 13.7 | 3.2 | 4,246 | 618 | 16.4 | 2.5 |

[a]In ms. [b]Out of 20. [c]2-back. [d]3-back.

difference between the no-load and both the load conditions, but the two-load conditions did not differ. The interaction between memory task and age group showed that the older adults were particularly disadvantaged on the episodic memory task, $F(1, 65) = 9.02$, $MSE = 728{,}962$, $p < .01$. There was also an interaction between memory task and working-memory load, $F(2, 130) = 3.94$, $MSE = 221{,}694$, $p < .05$. An analysis of simple main effects, using the Bonferroni procedure, found that in the episodic memory condition response times increased with an increase in working-memory load. In the semantic memory condition there was a significant difference between the no-load and both the load conditions. Finally, there was a three-way interaction between age group, memory task, and working-memory load, $F(2, 130) = 5.78$, $MSE = 221{,}694$, $p < .01$. Investigation of the interaction, using the Bonferroni procedure, revealed that in the episodic memory condition there was a significant difference between the younger adults' response times for the no-load condition and those for both the load conditions. For the older adults in the episodic memory condition the response times increased with working-memory load. In the semantic memory condition the interaction between working-memory load and age group was not significant.

One possible difficulty with the absolute data is that in some cases mean response times were calculated from very few responses. This occurred because mean response times were calculated only for correct responses, and older adults' performance in particular was quite poor. In order to determine whether this unduly biased the results two further analysis were conducted. First, an analysis was conducted after excluding the data where the mean response time was based on less than five correct responses. While no participants were excluded on the basis of accuracy in the episodic task, one younger and eight older adults were excluded on the basis of their accuracy in the semantic task. Excluding these data points did not affect the pattern of results reported above, and the results of this reanalysis are not reported here.

We also looked at this issue in a second way. If low numbers of observations had rendered the response time data unreliable, then the data might be expected to lack reliability across the two occasions that each condition was run. To investigate this we entered block as a variable in the above analyses. However, entering block as a variable in this manner made no difference to any of the previous analyses, and block was never significant, either as a main effect or as an interaction term. Thus we are confident in the robustness of our findings, despite the low number of observations in some conditions. In Experiment 2 the problem with low accuracy was overcome by using a recognition procedure rather than cued recall.

The next analysis was carried out on the memory retrieval errors. It should be noted that there were two types of error—out-of-time errors and retrieval errors—and these were combined for this analysis. In fact, the out-of-time errors were the most common errors across task and age group. In the episodic task for the younger adults 76%, 78%, and 76% of errors were no responses in the no-load, low-load, and high-load conditions, respectively. For the older adults 84%, 77%, and 69% of errors were no responses in the no-load, low-load, and high-load conditions, respectively. In the semantic task for the younger adults 93%, 90% and 89% of errors were no responses in the no-load, low-load, and high-load conditions, respectively. For the older adults 94%, 84%, and 87% of errors were no responses in the no-load, low-load, and high-load conditions, respectively.

A 2 × (memory task) 2 × (age group) × 3 (working-memory load) analysis of variance was carried out on the error data. This analysis showed that the older adults produced more errors, $F(1, 68) = 76.48$, $MSE = 17.91$, $p < .01$, error rates were greater in the semantic memory condition, $F(1, 65) = 55.16$, $MSE = 17.91$, $p < .01$, and accuracy was impaired as working-memory load increased, $F(2, 136) = 120.09$, $MSE = 4.75$, $p < .01$. Post hoc comparisons using the Bonferroni procedure confirmed that accuracy declined as working-memory load increased. The interaction between working-memory load and age group, $F(2, 136) = 25.93$, $MSE = 4.75$, $p < .01$, showed that older adults' accuracy was particularly impaired when carrying out a memory retrieval task with concurrent load. In addition, the interaction between memory task and working-memory load, $F(2, 136) = 10.05$, $MSE = 4.75$, $p < .01$, demonstrated that error rates increased mainly in the semantic memory condition with the presence of a concurrent task. Finally, the interaction between memory task and age group, $F(1, 68) = 3.36$, $MSE = 17.91$, $p > .05$, and the working-memory load by age group by memory task interaction, $F(2, 136) = 2.08$, $MSE = 4.75$, $p > .05$, were not significant.

One potential explanation of the difficulty that older adults had with the episodic task is that it might stem from poorer initial learning of the material, despite our training to criterion. To investigate this possibility, we re-analysed the error data, using number of trials to training as a covariate. If those who struggle to learn have poorer encoding, then this covariate might emerge as a factor at retrieval. However, it did not, and it made no difference to the pattern of results reported above.

### *Divided-attention costs analysis*

In order to control for baseline differences in memory retrieval response times a dual-task costs analysis was carried out. Older adults invariably perform more poorly on a variety of tasks, and as a result these differences must be considered when we evaluate older adults' performance on dual tasks. There have been two different metrics widely used to index dual-task costs in the literature—that is absolute differences (dual-task score minus single-task score) and proportional differences: dual-task score minus single-task score)/single-task score. It is arguable which method is more appropriate in the analysis of dual-task performance, and most researchers adopt a particular method without justification. In this research we were concerned with whether older adults have a disproportionate deficit in dual tasking. A significant age by condition interaction (single vs. dual) gives some suggestion of a disproportionate age effect, but since this interaction is based on absolute differences in performance, this can lead to interpretative difficulties when there are baseline differences in performance (see Perfect & Maylor, 2000). Therefore, we use proportional difference scores to control for baseline differences in performance, as advocated by Somberg and Salthouse (1982); for an in depth discussion on metric selection see Guttentag (1989). The use of proportional changes in performance provides a conservative estimate of divided-attention costs, and so any observed age differences in this index are likely to reflect a genuine age-related change. Figure 3 shows the mean divided-attention costs for younger and older adults for this experiment.

Dual-task costs on response time were analysed in a 2 (age group) × 2 (memory task) × 2 (memory load) analysis of variance. The main effect of age group showed greater proportional cost for older adults, $F(1, 65) = 21.27$, $MSE = 0.12$, $p < .01$, dual-task costs were greater in the episodic memory condition, $F(1, 65) = 6.59$, $MSE = 0.12$, $p < .05$, and there were greater costs as working-memory load increased, $F(1, 65) = 12$, $MSE = 0.02$, $p < .01$. The interaction between age group and memory load was unreliable, $F(1, 65) = 0.33$, $MSE = 0.008$, $p > .05$, as was the three-way interaction between working-memory load, age group, and memory task,

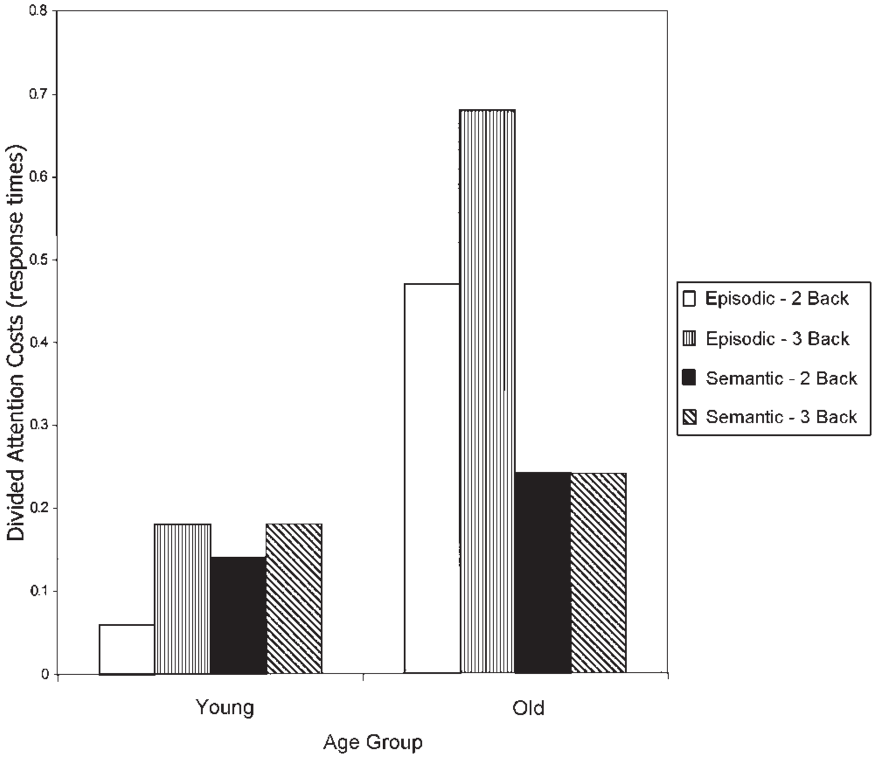


**Figure 3.** The proportional increase in memory retrieval, response times for younger and older adults.

*F*(1, 65) = 1.82, *MSE* = 0.04, $p > .05$. However, the Memory Task × Age Group interaction showed that older adults were particularly disadvantaged in the episodic memory tasks, *F*(1, 65) = 10.39, *MSE* = 0.12, $p < .01$. The interaction between working-memory load and memory task showed that increasing the memory load had the greatest effect in the episodic memory task, *F*(1, 65) = 7.54, *MSE* = 0.02, $p < .01$.

### *Speed–error trade-off*

The data so far have indicated that older adults are particularly slowed at retrieval for episodic memory compared to semantic memory, but they make fewer errors on the episodic task. Are these two facts related? Is the poorer performance of older adults on the semantic task a result of a differential speed–error trade-off across tasks? If participants are sacrificing response speed to maintain accuracy, then it follows that the change in error rates from no-load to load conditions should be negatively correlated with divided-attention costs. However, this was not the case. For the semantic task, the correlations were $r = .08$ for the low-load condition, and $r = .14$ for the high-load condition. For the episodic task, the equivalent correlations were $r = .61$ and $r = .57$. Thus, no negative relations were seen, and for the episodic task the relations were strongly positive. This is contrary to any notion of speed–error trade-off. Instead it indicates that those who were slowed by load were also made less accurate.

Additionally, the number of time-out errors is quite informative. For younger adults, as load increased from no load to high-load, time-out errors remained constant (76% of errors in the semantic task for no-load and high-load, and 93% and 89% for the equivalent conditions in the episodic task). For the older adults, proportion of time-out errors reduced as load increased for the episodic task (84% to 69%) and also for the semantic task, although less dramatically (94% to 87%). Thus, particularly for the task that older adults have more difficulty with, increased load shows the lowest proportion of time-out errors. This is the reverse of what one would predict if older adults were sacrificing speed to maintain accuracy in the demanding conditions, since time-out errors are errors of slow responding. If anything the data suggest that as demand increases, older adults are sacrificing accuracy to maintain speed. Thus, the observed divided-attention costs in our speed measure are more likely to be underestimates than overestimates.

### *Differential baselines*

One potential interpretative difficulty with our data is that the baseline performance of the two groups is not matched, as is common in studies of cognitive ageing. As has been frequently acknowledged, this makes interpretation of interaction terms involving age problematic (Perfect & Maylor, 2000). Thus, the apparent difficulty that older adults have with dual tasks involving episodic memory may be a consequence of the age differences in baseline performance on that task, perhaps stemming from poorer initial learning by the older adults. To investigate this further, we decided to rerun the dual-task analysis above, with matched groups of young and old. This was achieved by selecting the best performing older adults and the worst performing younger adults on the basis of baseline response time. For the episodic task this produced 9 younger and 9 older adults, with mean baseline response times of 2338 and 2293 ms, respectively. For the semantic task there were 17 younger adults (mean response time 3208 ms) and 10 older adults (mean response time 3176 ms). There were no age group differences in mean response time for either task ($t < 1$ in all cases). Fortunately, this selection process also eliminated mean accuracy differences between age groups. For the episodic task, younger adults were correct 76% of the time compared to 71% accuracy by older adults. This difference was not significant, $t < 1$. For the semantic task, the equivalent figures were 66% and 58%, which again did not differ, $t(25) = 1.41$, $p < .15$.

The question of interest is whether the previously observed interaction between age group and memory task remains in the dual-task analysis. The answer is, it does. The analysis based upon the matched subsets still showed both an overall age effect in dual-task cost, $F(1, 39) = 18.5$, $p < .001$, and a significant age by task interaction, $F(1, 39) = 4.58$, $p < .05$. In fact, the means are remarkably similar to those for the whole sample. For the semantic task, the young subsample showed a mean divided-attention cost of .12, and the older sample's mean divided-attention cost was .26. The equivalent figures for the whole sample were .16 and .24 for younger and older adults, respectively (collapsed across the manipulation of difficulty in each case). For the episodic task, the young subsample showed a mean divided-attention cost of .09, and the older adults' mean was .54. The equivalent figures for the full sample were .12 and .57. Thus, even with baseline performance matched across age groups, the significant age by task interaction remains, and these older adults show a particular problem with dual-tasks involving episodic memory.

## Discussion

The current study set out to investigate the circumstances where older adults have difficulties in combining tasks. We were particularly interested in age differences between semantic and episodic memory retrieval and the effects of increasing working-memory load. While dual-task performance in episodic memory tasks (e.g., Anderson et al., 1998) and dual-task performance in semantic memory tasks (Perfect & Rabbitt, 1993) have been examined in earlier studies, neither of these studies has provided a direct comparison between the two tasks.

Consider first the effects observed using absolute measures of reaction time and error rates. In addition to slower speed and poorer accuracy in the episodic and semantic tasks under single-task conditions, the present data demonstrated that older adults are impaired by concurrent processing demands. This was particularly pronounced in the episodic memory task. These results, based on absolute measures of performance, are consistent with a wide range of previous research on dual-task performance in older adults. However, not all studies have considered controlling for baseline differences in performance—that is, assessing dual-task costs. Of more importance, no earlier studies have assessed the dual-task costs associated with performance in episodic and semantic tasks in the same study.

As noted earlier, the analysis of an age by condition interaction, without controlling for baseline differences in performance, has been heavily and justly criticized (e.g., Somberg & Salthouse, 1982). In response to this, several methods of assessing dual-task costs have been advocated (Salthouse, Fristoe, Lineweaver, & Coon, 1995), but the most widely accepted method is the one originally proposed by Somberg and Salthouse (1982). This method of assessing dual-task costs was therefore used in the present study. Reaction time data were transformed into dual-task costs and compared between older and younger adults. For the semantic retrieval task, the results from the analysis were consistent with those of Perfect and Rabbitt (1993): After controlling for the effects of single-task performance there was no differential effect of dual tasking. The pattern of results for the episodic retrieval task was quite different. There was a clear effect of age in dual-task costs, which demonstrates that when episodic memory retrieval was required the older adults were more penalized by the concurrent processing demands. Anderson et al. (1998) also investigated episodic retrieval and found that, based on a dual-task cost analysis, older adults had a particular difficulty with retrieving episodic information. Thus the data presented here are consistent with Anderson et al.'s dual-task cost analysis of episodic retrieval, although the use of the *n*-back procedure eliminates the need to consider the trade-off that can occur using a traditional dual-task paradigm.

Moreover, the above conclusions hold even when one matches the two age groups on baseline performance, in terms of both speed and accuracy. Thus, the pattern observed is not an artefact of different baselines in the two groups. In addition, whilst speed–error trade-off is a possible confound, our analyses indicate that the direction of changes in response rate and accuracy were not consistent with the interpretation that the cause of the increased divided-attention costs in the older adults, for episodic memory, was a sacrifice of speed for accuracy in that condition. Thus, we are relatively confident that our data are robust.

The most important aspect of the data reported here is the suggestion that older adults do not show a general retrieval deficit in dual tasking. Rather, based on dual-task costs, the results reveal a deficit only when episodic retrieval is involved. This raises the question of why there should be a different pattern of dual-task costs across the two kinds of memory retrieval. One possible explanation that is frequently favoured is that of task difficulty. The idea here is that the more difficult each of the component tasks is made the greater the likelihood of disproportionate age effects when tasks are examined in combination. However, difficulty cannot explain the discrepancy between older adults' performance on the semantic and episodic memory tasks for two reasons. First, when difficulty was manipulated on the semantic task it did not produce an age effect in dual-task costs. While increasing the memory load did increase overall memory retrieval times, it affected both groups in the same way. This result is consistent with Perfect and Rabbitt's (1993) finding that increasing the difficulty of category generation produced slowing, but did not bring about an age difference in dual-task costs. Second, it might be argued that even in the high-load condition the semantic task was less demanding than the episodic task. However, the evidence does not support such a view because under single-task conditions, speed and accuracy were poorer for category exemplar retrieval than for paired-associate retrieval. That is, the conventional measures of demand—response times and accuracy—seem to indicate that the semantic retrieval task was more demanding for participants, and yet it was the episodic task that produced the disproportionate divided-attention costs for the older adults. It should also be noted that increasing difficulty (working-memory load) for both tasks affected both age groups to same extent.

In this experiment two demanding tasks were used, as indexed by the single-task performance scores, but the crucial question raised by the results is why does the requirement to hold and update material in working-memory only affect episodic memory retrieval disproportionately? Adding a load is certainly not just adding to the demands and complexity of the overall situation because this should have affected both tasks equally. Our results are more consistent with the notion that ageing in accompanied by a decline in "executive control". Successful episodic remembering requires considerable monitoring and strategic processing in order to retrieve item and contextual information. This contrasts with semantic retrieval where the information to be retrieved is highly familiar. Consequently, when "effortful" controlled/executive processes are already being employed (during episodic retrieval), concurrent processing (another putative executive ability) is especially problematic. Significantly, older adults are able to draw on a lifetime of learning to optimize their memory performance in the semantic dual tasks. Under these favourable retrieval conditions there would be less need for "executive control".

## EXPERIMENT 2

The previous experiment provided convincing evidence that older adults have difficulties in dual-tasking only in certain cognitive domains (episodic compared to semantic memory). We argued that the findings were difficult to reconcile with a task difficulty account. One key issue remains from Experiment 1 in that whilst it shows that episodic memory causes problems for older adults, we do not know whether this is because episodic memory is involved or just episodic recall. The present experiment addresses this issue by employing recognition versions of the dual tasks with the aim of replicating the pattern of findings.

Previous research investigating the attentional demands of encoding and retrieval from episodic memory provides strong evidence that episodic memory in general is problematic for older adults (e.g., Anderson et al., 1998; Craik & McDowd, 1987). There is broad consensus that an age-related decline in episodic memory is inevitable, but there is variability in the magnitude of the age effects across episodic tasks. Craik (1986) described the distinction between those tasks that draw on self-initiated processes (e.g., free and cued recall) and those that draw on environment cues (e.g., recognition) to aid performance. In free recall in particular, older adults are found to perform more poorly than younger adults, especially in the context of a dual-task (e.g., Anderson et al., 1998). Free recall requires largely effortful control processes and self-initiated retrieval operations. However, the burden can be minimized if external cues are provided, creating less need to draw on these effortful processes. In recognition tests, the necessary information is re-presented at test, which minimizes the need for effortful self-initiated processes. Typically, age differences on recognition tasks are observed but the magnitude of the effect is smaller than that seen in recall tasks. Regardless, these studies suggest that episodic memory is the crucial factor, and therefore in the present experiment we expect to find comparable age effects for our episodic recognition task.

Our prediction, based on the previous experiment, was that domain (episodic vs. semantic) would be more critical in producing age-related differences in divided-attention costs than would load. However, given the previous literature, our expectations were also that the general magnitude of such effects would be smaller for the recognition test than for the recall test. Our design also enabled us to test another possibility. Experiment 1 found that older adults are slow in episodic recall. If a recognition version of the dual-task paradigm fails to replicate this pattern, then this would imply that it is not episodic memory per se that is difficult for the elderly, but only recall from episodic memory.

## Method

### Participants

A total of 18 younger adults (mean age 20.6 years, *SD* 1.6 years) and 18 older adults (mean age 70.9 years, *SD* 5.9 years) participated in the experiment. The young volunteers were undergraduate students from the University of Bristol. The older adults were volunteers selected from a pool of older adults registered with the experimental psychology department. A total of 18 participants from each age group were randomly assigned to the episodic and semantic test conditions. On average the younger adults were better educated than the older adults: 16.1 vs. 12.7 years of education respectively, $t(34) = 6.3$, $p < .01$. Older adults scored more highly on the National Adult Reading Test: raw scores of 39.7 and 36.6, respectively, $t(34) = 2.3$, $p < .05$. These scores were entered as covariates in the main analyses.

### Materials

Six lists of nine words were constructed from high-imagery, high-frequency nouns, between four and seven letters in length, selected from the MRC Psycholinguistic Database (Coltheart, 1981). Three sets of three words were constructed randomly for each of the six lists. The six lists were randomly assigned to no-load, low-load, and high-load conditions at the start of each session. There was one practice block and one experimental block for each of the three episodic conditions. The categories and category members were selected form the Belfast Category Norms (Brown, 1978). A master list of the top six ranked responses for 21 categories was used to create six lists of 10 category–member pairs (e.g., furniture–bed). Each list was randomly assigned to no-load, low-load, and high-load conditions at the start of each experimental session. There was one practice block and one experimental block for each of the three semantic conditions.

## Procedure

*Semantic memory task*

Prior to the experimental session each participant received detailed instructions regarding the nature of the tasks. A printed example with screen shots of the no-load, low-load, and high-load conditions accompanied the instructions. Participants were required to demonstrate to the experimenter that they understood the instructions by completing the printed examples.

In order to ensure that participants understood the requirements of the task, a practice session of one block for each condition was given before the start of the experimental block. In the experimental block there was one session each of the no-load, the low-load, and the high-load tasks presented in counterbalanced order across participants. In each session there were 10 test trials.

*No-load condition.* In the no-load condition, the category generation task consisted of the presentation of a category name and a category exemplar. The category name appeared on the computer screen with a category exemplar beside (e.g., animal–zebra; see top panel of Figure 4 for a pictorial representation of the task). Participants were instructed to respond manually "yes" on a computer keyboard if the category exemplar belonged to the category, or "no" if it did not as quickly and as accurately as possible with a maximum time limit of 5 s (there were an equal number of "yes" and "no" responses). Category names and category exemplars were presented at 5-s intervals. No feedback was given and the computer recorded the final response made during the 5-s interval. There were a total of 10 test trials where participants were required to make a response.

*Low-load condition.* In the low-load condition, two words were presented simultaneously for 5 s (e.g., chisel–grass). After these items disappeared a category name was presented on the computer screen for 5 s (e.g., tool?). At this point participants had to decide whether either of the last two items that had just disappeared from the computer screen previously was a member of that category (see lower panel of Figure 4 for an example). Participants were instructed to respond on the computer keyboard as quickly and as accurately as possible. Thus, participants were required to keep in mind the previous two items presented in order to respond appropriately when the category name appeared. There were 10 test trials in total.

*High-load condition.* In the high-load condition, the task requirements were the same as those in the low-load condition except that the participants were required to keep in mind four items that were presented prior to a test trial. There were 10 test trials in total.

*Episodic memory task*

The procedure for the episodic memory task was identical to that for the semantic memory task, except that set names were used instead of category names (see Figure 5 for an example). The set names referred to sets of three words that were learned in a pretest training session (see details below). For example, Set A consisted of the words, *coin, screen, lady*. Prior to a test trial, participants would see either two or four words from different sets (e.g., *fish, coin*). On the test trial, participants were asked to decide whether one of the words belonged to a particular memory set, prompted by a set cue (e.g., Set A?).

*Pretest training.* In the pretest learning phase participants were trained on three sets of three words. A set of words would appear on the computer screen for 6 s (e.g., Set A: coin–screen–lady). After each of the three sets had been presented on the computer screen the experimenter asked the participant to recall each word from each set. This training procedure continued for 10 trials or until participants were able to recall all the words on two consecutive occasions. All participants reached criterion within 10 presentations of the word sets.

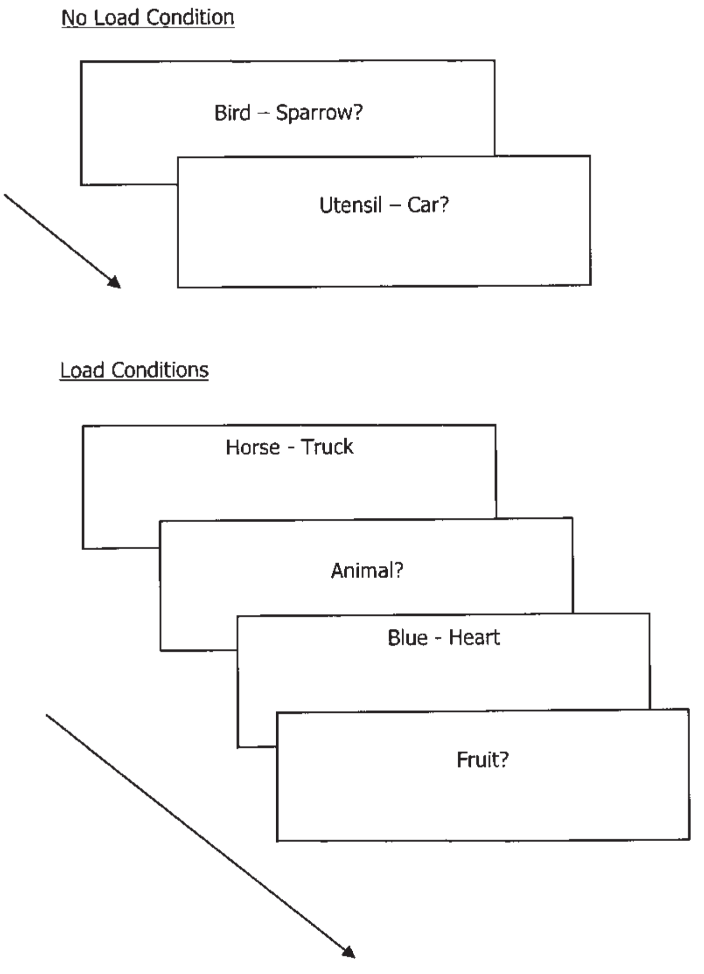


**Figure 4.** Example of the procedure. Semantic memory retrieval, no-load, and load conditions.

## Results

The analytic strategy used in Experiment 1 is replicated here. As before, because the age groups differed in education and vocabulary ability, these measures were initially entered as covariates in the analyses. However, as neither covariate was significant in any analysis, they are not discussed further. Post hoc analysis were conducted using the Bonferroni procedure with a significance level of $p < .05$ unless otherwise stated.

### *Analysis of absolute response time and error rate data*

The absolute response time and error data for the no-load and load conditions for both the episodic and semantic memory retrieval tasks are shown in Table 2.

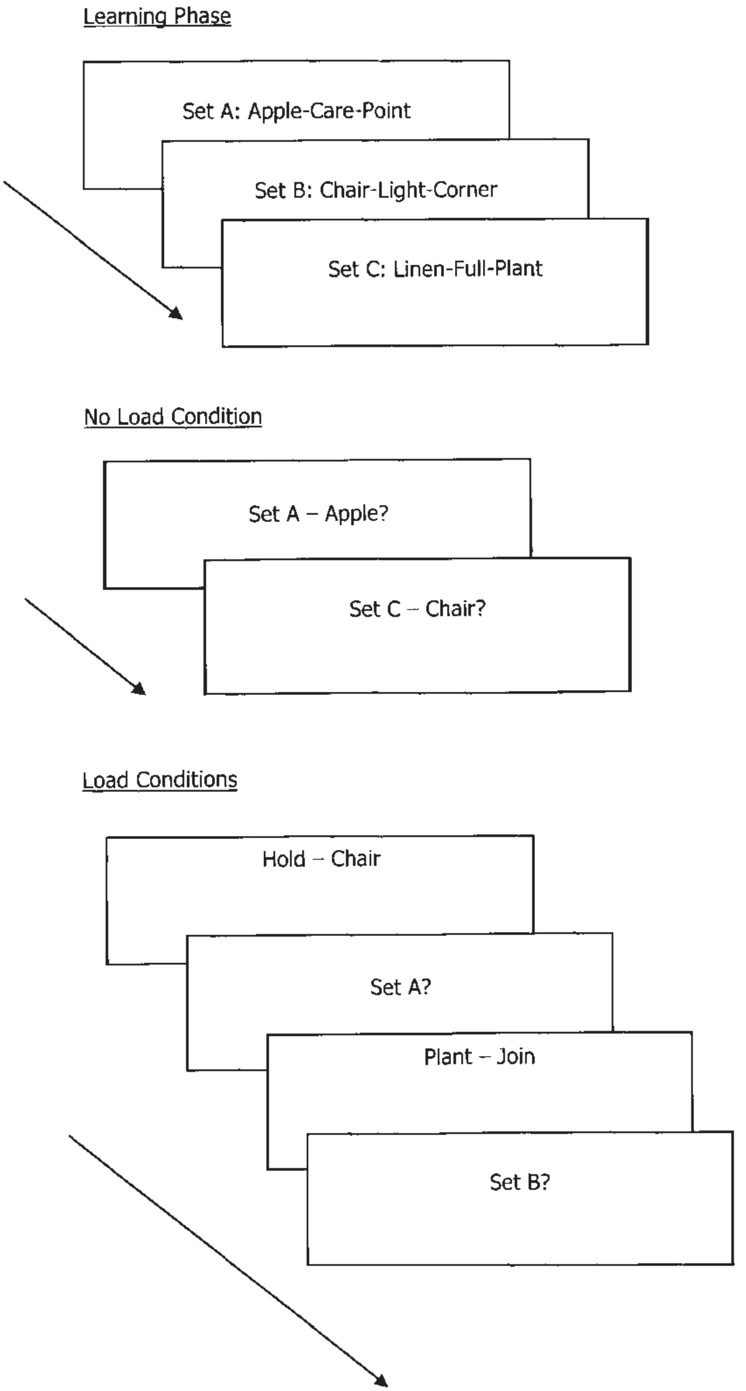


**Figure 5.** Example of the procedure. Episodic memory retrieval, no-load, and load conditions.

TABLE 2
Experiment 2 response latency[a] and errors[b] for the episodic and semantic recognition tasks under no-load, low-load, and high-load conditions

| | | *Recognition memory task* | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | *Episodic memory* | | | | *Semantic memory* | | | |
| | *Condition* | *Response latency* | | *Errors* | | *Response latency* | | *Errors* | |
| *Adults* | *Load* | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* |
| Younger | None | 1,179 | 211 | 0.5 | 0.9 | 1,424 | 295 | 0.3 | 0.5 |
| | Low | 1,168 | 188 | 0.5 | 0.7 | 1,086 | 185 | 1.1 | 0.5 |
| | High | 1,263 | 401 | 1.2 | 0.6 | 1,588 | 290 | 0.7 | 0.7 |
| Older | None | 2,680 | 471 | 2.6 | 2.0 | 2,293 | 524 | 0.6 | 0.9 |
| | Low | 3,025 | 583 | 3.2 | 2.4 | 2,245 | 618 | 5.2 | 2.4 |
| | High | 3,391 | 346 | 2.3 | 2.1 | 2,456 | 442 | 1.7 | 1.3 |

[a]In ms. [b]Out of 10.

Response times for correct answers to the recognition tests were analysed in a 2 (age group) × 2 (memory task) × 3 (memory load) analysis of variance. This analysis revealed that response times were generally slower for the older adults, $F(1, 34) = 390.4$, $MSE = 253{,}951$, $p < .01$, were longer in the episodic memory task, $F(1, 34) = 11.4$, $MSE = 244{,}175$, $p < .01$, and greatest in the high working-memory load condition, $F(2, 68) = 16$, $MSE = 100{,}817$, $p < .01$. The interaction between memory task and age group demonstrated that older adults' response times were particularly high in the episodic task, $F(1, 34) = 49.6$, $MSE = 244{,}176$, $p < .01$. The interaction between working-memory load and age group was significant, $F(2, 68) = 4.5$, $MSE = 100{,}817$, $p < .05$. An analysis of simple main effects, using the Bonferroni procedure, found that older adults' response times were slower in the high-load condition compared to either of the no-load or low-load conditions, which did not differ. For younger adults response times were again slower in the high-load condition. The response times were surprisingly slower in the no-load condition than in the low-load condition. Finally, there was a three-way interaction between age group, memory task, and working-memory load, $F(2, 68) = 3.13$, $MSE = 141{,}320$, $p < .05$. Investigation of the interaction, using the Bonferroni procedure, revealed that only in the episodic condition was there an interaction between working-memory load and age.

The next analysis was carried out on the recognition errors. A 2 × (memory task) 2 × (age group) × 3 (working-memory load) analysis of variance was carried out on the error data. This analysis showed that the older adults produced more errors, $F(1, 34) = 68.47$, $MSE = 2.81$, $p < .01$. A main effect of load demonstrated that there were fewer errors in the no-load condition than in the load conditions, $F(2, 68) = 22.01$, $MSE = 1.98$, $p < .01$. Post hoc comparisons using the Bonferroni procedure surprisingly found that there were fewer errors in the high-load conditions than in the low-load condition. The interaction between working-memory load and age group, $F(2, 68) = 15.98$, $MSE = 1.98$, $p < .01$, showed that older adults' accuracy was particularly impaired when carrying out a recognition memory task with concurrent load. There was an interaction between memory task and working-memory load, $F(2, 68) = 21.24$, $MSE = 1.38$, $p < .01$. Examining this interaction using the Bonferroni procedure showed that for the episodic condition the error rates were equivalent for each level of load. For the semantic memory condition error rates were the greatest in the load conditions. However, error rates were surprisingly higher in the low-load condition than in the high-load condition. Finally, the working-memory load by age group by memory task interaction was significant, $F(2, 68) = 8.47$, $MSE = 1.38$, $p < .01$. An analysis of simple interaction effects showed that in the episodic condition, for older adults the error rates were equivalent at the three levels of load. For younger adults the error rates were the greater in both the load conditions than in to the no-load condition. In the semantic memory task, for older adults the error rates were the greatest in the load conditions. Surprisingly, error rates were greater in the low-load condition than in the high-load condition. For younger adults the error rates were the highest in the low-load condition. There was no difference between error rates in the no-load and high-load conditions.

### *Dual-task costs analysis*

In order to control for baseline differences in performance, dual-task costs analysis of response times was carried out. Figure 6 shows the mean divided-attention costs for younger and older adults for this experiment. However, because the low-load condition in this experiment did not produce any slowing in performance, and hence no apparent cost, this condition was omitted from the following analysis. (We discuss possible reasons for this pattern and

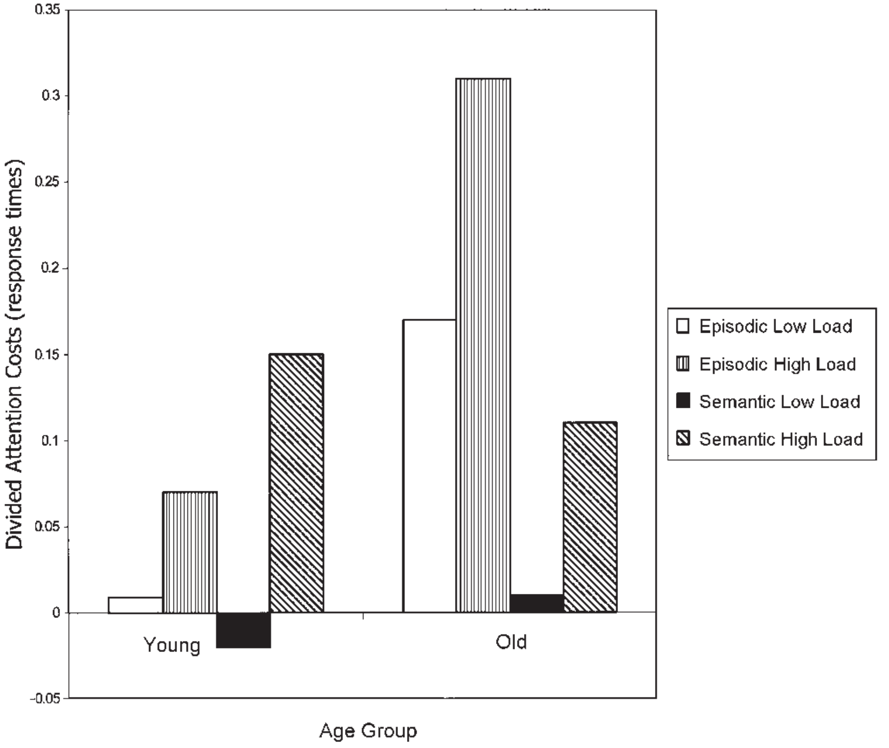


**Figure 6.** The proportional increase in memory retrieval response times for younger and older adults.

justify this analytic strategy further in the Discussion section below.) Thus divided-attention costs were analysed for the high-load conditions in the episodic and semantic memory tasks only, using a 2 (age group) × 2 (memory task) repeated measures analysis of variance. There were no main effects of either age group, $F(1, 34) = 2.58$, $MSE = 0.07$, $p < .12$, or memory task, $F < 1$. However, there was a significant interaction between the two factors, $F(1, 34) = 4.14$, $MSE = 0.08$, $p < .05$. Younger adults showed equivalent divided-attention costs for the episodic (divided-attention cost = .07) and semantic (divided-attention cost = 0.15) tasks. However, whilst older adults showed similar divided-attention costs to the younger adults for the semantic task, (divided-attention cost = .11), for the episodic task they were considerably worse (divided-attention cost = .31).

### *Speed–error trade-off*

Since we were concerned with both the speed of memory retrieval and the errors we investigated whether these two measures were related. In order to examine the speed–error trade-off between participants we correlated the raw response time and error scores for all conditions. If participants are sacrificing response time to maintain accuracy, negative correlations would be expected. There was no evidence to suggest that the speed of retrieving from memory was related to the error rates as the correlations were either not significant or positive. However, for the episodic task the correlations were positive ($r = 0.46$, $p < .01$ for no-load, $r = .78$, $p < .01$ for low-load, and $r = .34$, $p < .05$ for the high-load). For the semantic task, the correlations were positive ($r = .39$, $p < .05$ for no-load, $r = .53$, $p < .05$ for low-load, and $r = .28$, $p > .05$ for the high-load).

A second analysis examined strategic changes between no-load and load conditions. The reasoning is as follows. Participants may maintain speed by sacrificing accuracy as load increases. If this is the case then increase in error rate should negatively correlate with divided attention costs. Whilst this was the case for episodic memory, the correlation was weak ($r = -.19$, $p < .26$) and did not alter the relationship between age and divided-attention costs in the episodic task. For semantic memory, there was no evidence of a relationship between increase in error rate and divided-attention costs ($r = -.03$).

### *Differential baselines*

As with Experiment 1, we were interested to know whether older adults' increased divided-attention costs in the episodic task were due to initial baseline differences in response time. However, unlike in Experiment 1, it was not possible to select matching groups of younger and older adults on the basis of baseline response time, since the two response time distributions were nonoverlapping.

In order to explore this issue, we did a median split for baseline response time for each memory task, for each age group, and examined the resultant divided-attention costs. This procedure, whilst not producing groups with matched baselines, does produce two groups who differ less than the entire sample. Thus, if baseline response times are causal in producing the increased divided-attention costs for the older adults, these restricted samples should show reduced age differences. For the episodic task, the fastest older adults produced an average divided-attention cost of .54, whilst the slowest younger adults showed an average divided-attention cost of .09. Thus, the mean age difference in divided-attention cost is numerically higher, which is the reverse of what would be anticipated if the increased divided-attention costs were due to poorer baseline performance. For the semantic task, the same procedure led to divided attention costs of .02 for the young and .25 for the old. So, increasing the similarity of the baseline responses for the two age groups once again numerically increased the age differences in divided-attention costs. This is contrary to what would be expected if baseline differences in response time caused the ageing pattern observed in divided-attention costs. Thus, whilst the data do not allow as elegant an analysis as the first study, the pattern in the data suggests that baseline differences in response time are not the cause of the differential age effect on divided-attention costs in episodic memory. We therefore remain convinced that the increased divided-attention costs for older adults on the episodic memory task are not an artefact of baseline response time differences.

## Discussion

Experiment 2 had two main aims. First, the aim was to replicate the pattern of age-related divided attention costs being related more to domain (greater in episodic than in semantic memory) rather than memory load. Second, the aim was to investigate whether increased environmental support (i.e., using recognition) moderated the dual-task age effect within the episodic domain.

Consider first the effects observed using the absolute reaction times and error rates. As expected, older adults' response times and error rates were higher. Of particular interest was that for response times there was an age by load interaction, which demonstrated that older adults overall found dual tasking problematic. This effect also interacted with task, showing that it was only in the episodic memory retrieval task that older adults had problems dual tasking, replicating the pattern seen in Experiment 1.

The analysis of the error rates did not demonstrate any particular problems for older adults. We have already discussed the problems with the error rate analysis for the *n*-back tasks. Again, with the tasks used in this study it is difficult to know whether a retrieval error has occurred or whether the error or no response is due to forgetting the items presented in the load conditions. The recognition procedure provided favourable retrieval conditions, and both groups were able to maintain their accuracy reasonably well in the load conditions. A comparison of Tables 1 and 2 demonstrates that the use of the recognition rather than cued recall reduced the error rates for both groups. Furthermore, in Experiment 1 there were a large number of no-response errors, and this procedure has eliminated this problem. By maintaining accuracy for both groups the focus could be on the correct response data where we could be reasonably sure that load is appropriate. In other words, the participant has remembered the load items and has recognized that one of the items belongs to a particular category or set.

One pattern that was unexpected was the superior performance seen in the low-load conditions. In three out of four cells the responses are faster in the low-load condition than in the control condition. One possible reason for this is that the reading demands are less in the load test trials. In the no-load condition, participants read and verify category–exemplar pairs, whereas in the load conditions they first read the items and then respond only to the category cue. Thus, whilst the load may have slowed responding in the low-load condition, this is offset by the reduction in reading time. The same would apply to high-load trials, but clearly the extra demands of the high load were sufficient to demonstrate costs notwithstanding the reading time differences. Thus, the divided-attention costs measured are probably underestimates for both load conditions.

Another possibility to consider is the likelihood of speed–error trade-offs across different load conditions. As well as showing faster responses in the low-load conditions, there was also evidence that the low-load conditions produced unexpectedly high error rates. In three out of four cells, the low-load condition produced more errors than the high-load conditions. This raises the question of whether speed–error trade-off is a major confound in the present experiment. Whilst our individual difference analyses provided no evidence for this, it is possible that all participants changed strategy from no-load to load (sacrificing accuracy for speed) to the same degree. However, this would only be a selective reading of the data, since with higher load, accuracy improved, relative to low load. To explain the entire pattern in terms of speed–error trade-offs would require first that participants sacrifice accuracy for speed in the low-load condition, but then speed for accuracy in the high-load condition. Such a pattern is possible, but there is no a priori reason to expect it. Given the lack of any evidence for differential strategy use by different individuals, we do not favour such a speed–error trade-off account of our data. Nonetheless, the performance on the low-load condition does appear unexpected. The only sensible response to this, given the lack of clear evidence of divided-attention costs in the low-load condition, is to focus only on the high-load condition, which is precisely what we did.

As with the previous experiment a dual-task costs analysis investigated whether older adults had problems dual tasking over and above the problems they had in the full-attention condition. For semantic memory retrieval older adults' response times in terms of dual-task costs were equivalent, and increasing load had a similar effect for both groups. However, for episodic memory retrieval older adults found dual tasking particularly problematic. An analysis was not carried out on the error rate data as in some instance participants made no errors and their divided costs could not be calculated. Moreover, we noted how the interpretation of the error rate data is misleading.

The age effect was greater for episodic recall in Experiment 1 than for recognition in Experiment 2 ($d = 1.52$ compared to $d = 0.71$). Under dual-task conditions older adults are more penalized when using a cued-recall procedure. Older adults have been shown to benefit from supportive retrieval conditions, and although the age effect remained for the recognition version of the dual-task used in the present experiment it was substantially smaller. Cued recall does require additional self-initiated retrieval processes compared to recognition that has a large familiarity component. Therefore, with concurrent processing demands a more detrimental effect for older adults in cued recall is observed. This is consistent with previous research that has compared retrieval from free recall, cued recall, and recognition (e.g., Anderson et al., 1998).

## GENERAL DISCUSSION

The primary goal of the present study was to investigate age-related differences in the effects of concurrent tasks on memory retrieval. These data provide strong support for the hypothesis that task domain moderates dual-task costs in older adults. A clear age effect of dual tasking was found for episodic memory retrieval. In contrast, for semantic memory retrieval, age invariance was observed for dual-task costs, and it is suggested that this may be due to semantic memory being preserved in normal ageing. The results also indicated that when a supportive retrieval environment was introduced (recognition dual-task) the age effect was reduced. No doubt task difficulty and complexity have a role in dual-task performance, but these results are best explained by a domain-specific account. That is, age difference in dual-task performance may be observed only in particular domains of cognition, rather than at different levels of difficulty.

Throughout the literature, three main accounts have been proposed to explain poor dual tasking in old age: (1) a reduction of attentional resource; (2) a decline in the speed of processing; and (3) a specific deficit in executive control. The resource deficit account proposes that cognitive resources decrease with age, resulting in less resource to deploy to the component tasks. The cognitive-slowing view argues that concurrent tasks reduce the amount of time available for processing, and since speed of processing declines with age, older adults will be particularly impaired on dual tasks. Our data do not support these accounts as disproportionate dual-task costs were found for episodic memory retrieval. Furthermore, consistent with previous research (e.g., Korteling, 1991), large age differences were observed in the less "complex" or demanding episodic than semantic tasks (although in Experiment 2, for older adults, the reaction times and error rates were greater for the episodic task). In addition, the attentional resource and speed of processing account would predict that increasing the demands of the dual task would exaggerate the age effect. In the present study, manipulating the working-memory load affected both groups in the same way. This is against the work of Tsang and Shaner (1998), who found that the age effect was exaggerated when the demands were manipulated within task. In sum, the resource deficit model of ageing (e.g., Craik, 1986) and an account based on generalized slowing (e.g., Somberg & Salthouse, 1982) are insufficient in accounting for all instances of age-related differences in dual-task performance. Here we argue that older adults' difficulties are best explained in term of a specific deficit in executive control. Importantly, in certain task domains older adults are able to capitalize on intact cognitive function, and consequently there is less need to schedule and coordinate multiple processes while dual tasking (Tun & Wingfield, 1993).

Tun et al. (1992) argued strongly that the particular tasks involved in the dual-task activity are crucial in whether we observed age differences in performance. They found that increasing the demands within a task and increasing the demands by the requirement to divided attention do not necessarily bring about the same effect. The authors suggest that dual-tasking impairment only occurs in domains of cognition that are impaired in normal ageing. However, in domains like language processing we might expect to see age invariance of dual-task cost. Likewise, in the semantic memory domain it has been argued that the processes involved are largely "automatic" and are carried out without awareness. Therefore, competing activities are unlikely to cause much interference for the young or the old. Consequently, there was no significant difference between younger and older adults' dual-task costs for semantic memory in either experiment. With semantic memory it was found that the material was overlearned, and therefore its retrieval drew on processing mechanisms relatively intact in ageing (e.g., familiarity). In the category exemplar generation tasks, when a cue is presented semantic activation of candidate exemplars might occur, and retrieval follows in a relatively automatic manner. This enables more attention to be devoted to the concurrent activity. We propose that the semantic processing domain is relatively independent and resilient to the effects of ageing and divided attention.

Research findings regarding episodic memory are quite different. In Experiment 1, the episodic memory version of the task involved the effortful retrieval of words that had previously been learned. Although the task was cued recall, and some environmental support was provided at test, older adults were required to engage in consciously controlled processing involved in generating retrieval strategies, organizing of a retrieval search, and the like. Significantly, the retrieval of contextual information is required in episodic retrieval tasks, and declines have been observed in normal ageing (e.g., Newman, Allen, & Kaszniak, 2001). Therefore, the processes involved in episodic retrieval and the possible involvement of controlled/executive processes in task coordination would tend to exaggerate the age effect. Competition for such processing mechanisms would be heightened for older adults. However, the burden of task coordination for older adults is reduced if there is less need to draw on controlled or executive processes involved in memory retrieval, which are known to decline in ageing.

Familiarity of the material and therefore less need to draw on effortful retrieval of context enables older adults to optimize their dual-task performance when semantic memory is involved. Familiarity can also moderate dual-task cost in older adults within the episodic domain. Craik (1986) described the distinction between those tasks that draw on self-initiated processes and those that draw on environmental cues to aid performance. With supportive retrieval environments, older adults may be able to compensate for their poorer retrieval performance. In free recall in particular, older adults are found to perform more poorly than younger adults, particularly in the context of a dual-task (e.g., Anderson et al., 1998). Free recall requires largely effortful control processes and self-initiated retrieval operations. The burden can be minimized if external cues are provided so that there is less need to draw on those effortful processes. In our episodic dual-tasks we found that manipulating the amount of self-initiated processes (recognition vs. cued recall) affected older adults' dual-task performance. The age effect was greater in the cued recall task, than in the recognition version of the task ($d = 1.52$ and $d = 0.71$, respectively). These effects are similar to those found by Anderson et al.; $d = 1.64$ and $d = 0.89$. Familiarity, and therefore, less reliance on effortful retrieval make an impact on dual-task performance. Particularly, age differences in dual-task costs can be reduced by relying less on processing mechanisms that are impaired in normal ageing.

With the specific tasks used in this study, task domain seems to be the critical moderator variable and replicates earlier work on category exemplar generation (Perfect & Rabbitt, 1993) and episodic memory (e.g., Anderson et al., 1998). Further attention should be focused on particular task combinations in the dual-task setting, especially those studies that consider task domain to be an important mediator. To be clear, we believe that under certain conditions a deficit in executive control prevents older adults effectively monitoring and scheduling concurrent activities. However, in domains where older adults show expertise, are well learned, or use skilled processing it is expected that coconcurrent processing should not be especially problematic.

## REFERENCES

Anderson, N. D., Craik, F. I. M., & Naveh-Benjamin, M. (1998). The attentional demands of encoding and retrieval in younger and older adults: 1. Evidence from divided attention costs. *Psychology and Aging*, *13*(3), 405–423.

Baron, A., & Mattila, W. R. (1989). Response slowing of older adults: Effects of time-limit contingencies on single-task and dual task performances. *Psychology and Aging*, *4*(1), 66–72.

Brown, W. P. (1978). *Belfast Category Norms, 1971–77.* Department of Psychology, Queen's University of Belfast.

Craik, F. I. M. (1977). Age differences in human memory. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (pp. 384–420). New York: Van Nostrand Reinhold.

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). New York: Elsevier Science.

Craik, F. I. M., & McDowd, J. M. (1987). Age-differences in recall and recognition. *Journal of Experimental Psychology: Learning Memory and Cognition*, *13*(3), 474–479.

Coltheart, M. (1981). The MRC Psycholinguistic Database. *Quarterly Journal of Experimental Psychology*, *33A*, 497–505.

Daneman, M., & Carpenter, P. (1980). Individual differences in working-memory and reading. *Journal of Verbal Learning and Verbal Behavior*, *19*, 450–466.

Guttentag, R. E. (1989). Age-differences in dual-task performance—procedures, assumptions, and results. *Developmental Review*, *9*(2), 146–170.

Isingrini, M., Vazou, F., & Leroy, P. (1995). Dissociation of implicit and explicit memory tests: Effect of age and divided attention on category exemplar generation and cued-recall. *Memory & Cognition*, *23*(4), 462–467.

Kirchner, W. K. (1958). Age differences in short-term retention of rapidly changing information. *Journal of Experimental Psychology*, *55*, 352–358.

Korteling, J. E. (1991). Effects of skill integration and perceptual competition on age-related differences in dual-task performance. *Human Factors*, *33*, 35–44.

Light, L. L., & Prull, M. W. (1995). Aging, divided attention, and repetition priming. *Swiss Journal of Psychology*, *54*(2), 87–101.

Macht, M. L., & Buschke, H. (1983). Age-differences in cognitive effort in recall. *Journals of Gerontology*, *38*(6), 695–700.

McDowd, J. M., & Craik, F. I. M. (1988). Effects of aging and task-difficulty on divided attention performance. *Journal of Experimental Psychology: Human Perception and Performance*, *14*(2), 267–280.

McDowd, J. M., Vercruyssen, M., & Birren, J. E. (1991). Aging, divided attention, and d performance. In D. L. Damos (Ed.), *Multi-task performance* (pp. 387–414). Hove, UK: Psychology Press Ltd.

Newman, M. C., Allen, J. J. B., & Kaszniak, A. W. (2001). Tasks for assessing memory for temporal order versus memory for items in aging. *Aging Neuropsychology and Cognition*, *8*(1), 72–78.

Perfect, T. J., & Maylor, E. A. (2000). Rejecting the dull hypothesis: The relation between method and theory in cognitive aging. In T. J. Perfect & E. A. Maylor (Eds.), *Models of cognitive aging*. Oxford, UK: Oxford University Press.

Perfect, T. J., & Rabbitt, P. M. A. (1993). Age and the divided attention costs of category exemplar generation. *British Journal of Developmental Psychology*, *11*(2), 131–142.

Salthouse, T. A., Fristoe, N. M., Lineweaver, T. T., & Coon, V. E. (1995). Aging of attention: Does the ability to divide decline? *Memory and Cognition*, *23*(1), 59–71.

Somberg, B. L., & Salthouse, T. A. (1982). Divided attention abilities in young and old adults. *Journal of Experimental Psychology: Human Perception and Performance*, *8*(5), 651–663.

Tsang, P. S., & Shaner, T. L. (1998). Age, attention, expertise, and time-sharing performance. *Psychology and Aging*, *13*(2), 323–347.

Tun, P. A., & Wingfield, A. (1993). Is speech special? Perception and recall of spoken language in complex environments. In J. Cerella, J. Rybash, W. Hoyer, & M. L. Commons (Eds.), *Adult information processing: Limits on loss* (pp. 425–457). San Diego, CA: Academic Press.

Tun, P. A., Wingfield, A., Stine, E. A. L., & Mecsas, C. (1992). Rapid speech processing and divided attention—processing rate versus processing resources as an explanation of age effects. *Psychology and Aging*, *7*(4), 546–550.

Watter, S., Geffen, G. M., & Geffen, L . B. (2001). The N-back as a dual-task: P300 morphology under divided attention. *Psychophysiology*, *38*(6), 998–1003.

*Original manuscript received 16 April 2002*
*Accepted revision received 28 January 2003*
*PrEview proof published online 14 July 2003*