



Effects of divided attention at encoding and retrieval: Further data

Fergus I. M. Craik¹ · Eldar Eftekhari¹ · Malcolm A. Binns¹

Published online: 22 June 2018
© Psychonomic Society, Inc. 2018

Abstract

Division of attention (DA) at the time of learning has large detrimental effects on subsequent memory performance, but DA at retrieval has much smaller effects (Baddeley, Lewis, Eldridge, & Thomson, 1984, *Journal of Experimental Psychology: General*, 113, 518–540; Craik, Govoni, Naveh-Benjamin, & Anderson, 1996, *Journal of Experimental Psychology: General*, 125, 159–180). Experiment 1 confirmed the relatively small effects of DA on retrieval and also showed that retrieval operations do consume processing resources. The experiment also found that the effect is not attributable to a trade-off in performance with the concurrent task or to recognition decisions made on the basis of familiarity judgments. Participants made levels-of-processing (LOP) judgments during encoding to check whether deeper semantic judgments were differentially vulnerable to the effects of DA. In fact DA did not interact with LOP. Experiment 2 explored reports that the comparatively slight effect of DA on recognition accuracy is accompanied by a compensatory increase in recognition latency (Baddeley et al., 1984). The experiment replicated findings that neither DA nor differential emphasis between recognition and a concurrent continuous reaction time (CRT) task affected recognition accuracy, but also found evidence for a lawful trade-off in decision latencies between recognition and CRT performance. Further analysis showed that the relationship between response rates on the two tasks was well described by a linear function, and that this function was demonstrated by the majority of individual participants. It is concluded that the small effect of DA on recognition performance is attributable to a trade-off within the recognition task itself; accuracy is maintained by a compensatory increase in decision latency.

Keywords Divided attention · Encoding · Retrieval · Response latency · Response rate

It is generally agreed that effective learning requires attentional resources. One way in which this has been demonstrated empirically is to contrast encoding and retrieval under conditions of full attention with conditions in which participants must perform a secondary task simultaneously with encoding or retrieval operations. Many studies have shown convincingly that such conditions of divided attention (DA) at encoding result in substantial decrements in subsequent recall and recognition performance (Baddeley, Lewis, Eldridge, & Thomson, 1984; Craik, Govoni, Naveh-Benjamin, & Anderson, 1996; Guez & Naveh-Benjamin, 2006; Mulligan, 1998; Murdock, 1965). Interestingly and unexpectedly, DA at retrieval has much smaller negative effects on both recall and recognition (Baddeley et al., 1984; Craik et al., 1996; Kellogg, Cocklin, & Bourne,

1982), and the present article focuses on possible reasons for this discrepancy. In particular, DA combined with recognition testing resulted in no memory decrements in some early studies (Baddeley et al., 1984; Craik et al., 1996), although later studies have shown small recognition decrements (Dodson & Johnson, 1996; Gruppuso, Lindsay, & Kelley, 1997; Hicks & Marsh, 2000; Lozito & Mulligan, 2006).

Why should the effects of DA during retrieval be less than the effects of DA during encoding, especially for recognition memory? Initially Baddeley et al. (1984) concluded that retrieval must be ‘automatic’ and cost-free in terms of attentional resources, although they did find that retrieval latency increased under DA conditions. The notion of automatic retrieval was essentially ruled out by later studies, however, including the experiments reported by Craik et al. (1996). They used a continuous reaction-time (CRT) task as the concurrent secondary task, and found that RTs slowed substantially under DA at retrieval conditions relative to the CRT task performed on its own, leading to the conclusion that retrieval processes do in fact require attentional resources. Craik and colleagues also found that instructions to vary the relative emphasis

✉ Fergus I. M. Craik
fcraik@research.baycrest.org

¹ Rotman Research Institute at Baycrest, 3560 Bathurst St, Toronto, ON M6A 2E1, Canada

between the CRT task and memory retrieval changed performance systematically on the CRT task but had no effect on memory performance. Specifically, recognition performance was unaffected by both divided attention at test and by the emphasis manipulation. This pattern of results led the authors to conclude that retrieval processes are in some sense mandatory and protected, and therefore the required attentional resources (typically greater in recall than in recognition) are necessarily withdrawn from the processes mediating the secondary task.

One straightforward account of these findings is that performance on memory tasks in DA at retrieval conditions simply trade off with performance on the secondary task; better memory performance should mean poorer CRT performance by this view. However, in each of Experiments 2, 3, and 4 in the Craik et al. (1996) series, greater emphasis on the memory task was associated with more slowing of the CRT task (showing that memory retrieval requires resources), but with no further beneficial effects on memory performance. These replicated results thus speak against any simple version of the notion that preserved memory performance during DA at retrieval depends on a trade-off with performance of the secondary task.

Another suggestion relevant to recognition memory is that recognition performance under dual-task conditions is largely mediated by processes associated with familiarity rather than with processes of conscious recollection. That is, target words are endorsed as “old” because participants feel they have experienced them recently although they have no explicit recollection of their reaction at the time of encoding. The claim here is that familiarity and recollection are dissociable aspects of recognition memory and that familiarity requires relatively small amounts of processing resources (Jacoby, 1991; Mandler, 1980). A number of studies by Jacoby and colleagues provide evidence for this view by showing that divided attention during retrieval reduces recollection but has essentially no effect on familiarity (e.g. Jacoby, Kelley, Brown & Jasechko, 1989a; Jacoby, Kelley, & Dywan, 1989b). It therefore seems possible that the sustained levels of recognition memory under DA conditions are attributable to participants making recognition decisions on the basis of feelings of familiarity rather than on recollection. This is one idea tested in the first experiment reported here. We asked participants to make remember/know (R/K) judgments for each item they claimed to recognize, under both full and divided attention conditions. Given that know judgments are given on the basis of familiarity without recollection (e.g. Gardiner & Richardson-Klavehn, 2000; but see also Cohen, Rotello, & MacMillan, 2008; Ingram, Mickes & Wixted, 2012; Wixted & Mickes, 2010, for alternative views), we might expect that correct recognition decisions under DA at retrieval conditions would largely reflect K judgments.

A second point that emerges from Jacoby’s (1991) analysis is that when items are encoded richly and semantically, such

encoding supports higher levels of subsequent memory performance (e.g. Craik & Tulving, 1975). However, assuming that these higher levels of retention are attributable to enhanced recollection (R) which *is* affected by DA at retrieval, items encoded semantically should be particularly vulnerable to dual-task conditions at retrieval. Jacoby (1991) presented evidence in favor of this notion by having participants solve anagrams of some words during the encoding phase (thereby increasing semantic processing of the word), whereas other words were simply read aloud; no mention was made of a subsequent memory test. Words encoded as anagrams were later recognized less well under conditions of DA than under full attention at retrieval (hits minus false alarm scores for full attention = 0.66, for DA at test = 0.49), whereas words read aloud at study were not so affected (full attention = 0.32, DA = 0.34, respectively). This idea has been followed up with mixed results. Jacoby’s findings were replicated by Hicks and Marsh (2000), who also used anagram solution as a means of inducing semantic encoding, but not by Lozito and Mulligan (2006), who compared the effects of rhyme versus semantic processing at study, and found a small but reliable effect of DA during test that was equivalent for the two types of encoding. However they did find an interaction between DA at test and yes/no during study, where yes signifies a positive response to the initial encoding question. DA had a stronger negative effect on yes responses, and this result is in line with Jacoby’s view on the assumption that positive responses are associated with more elaborate forms of encoding (Craik & Tulving, 1975). A second purpose of the first experiment was therefore to obtain further evidence on these points by combining a more traditional levels-of-processing (LOP) manipulation at study with division of attention at both encoding and retrieval.

Experiment 1

The purpose of the first experiment was to obtain further evidence on the reasons for the relatively small memory decrements associated with DA at retrieval. In particular, we wished to explore the possibility that recognition decisions made under conditions of DA at retrieval may be made on the basis of familiarity rather than recollection, and whether items encoded deeply and elaborately are more vulnerable to the effects of DA at retrieval. We also wished to obtain further data on the differential effects of DA at encoding and retrieval, and on the pattern of dual-task performance with regard to the respective effects on memory and the secondary task itself. In overview, participants encoded words in response to three types of orienting questions (letters, rhymes, and semantic associations) and were later given both recall and recognition tests for the words. The encoding and retrieval phases were performed under conditions of full attention or DA at either

encoding or retrieval. The design was therefore 3 (levels of encoding) \times 3 (locus of attention) \times 2 (yes/no items), with all variables being within participants. Learning was intentional (participants were informed of the upcoming memory tests), and the secondary task was detection of three successive odd digits (e.g. 3-9-1 or 5-7-5) in a lengthy string of digits presented auditorily (Craik, 1982; Jacoby, 1991).

Method

Participants The participants were 48 undergraduates (35 female) at the University of Toronto who performed the experiment either for course credit or for a stipend of \$12 CAD per hour. The age range was 17–30 years, plus one outlier of 37 years, and the mean age was 20.0 years. The mean number of years of education was 14.0 years.

Materials and design The words used in the study were common two-syllable concrete nouns presented in white font on a black background in the middle of a computer monitor screen. During the encoding phase, each word was preceded by a question that pertained either to its first letter (e.g. ‘Does the word begin with B?’—BAGEL), to its rhyme characteristics (e.g. ‘Does the word rhyme with Mitten?’—KITTEN) or to its meaning (e.g. ‘Is the word related to Picture?’—ARTIST). Each list contained 18 words, six each of the three question types—Letter, rhyme, and associate—presented randomly. Within each type, four questions required a ‘yes’ answer and two required a ‘no’ answer (e.g. ‘Does the word rhyme with Table?’—CABBAGE). The reason for this imbalance was our greater focus on words associated with a positive response, especially in recognition tests. Each question appeared for 2 s followed by its associated word, also shown for 2 s, during which the participant responded yes or no by pressing one of two keyboard buttons. Following the word’s presentation there was a 1-s pause before the next question appeared. Immediately following the presentation of each list, participants were asked to recall as many of the preceding words as possible, in any order, by writing them on a sheet of paper; they were given 60 s for this task.

There were nine lists in all, each followed by a recall session. Three consecutive lists were encoded and recalled under conditions of full attention, three were encoded under divided attention (DA) and recalled under full attention, and three were encoded under full attention and recalled under divided attention. The secondary task in the DA condition was used previously by Craik (1982) and by Jacoby (1991); it consisted of a long string of single digits presented auditorily at a 1.5-s rate. The participant’s task was to detect any runs of three consecutive odd digits (e.g. 5, 9, 1, or 7, 3, 3) and to respond by saying the third digit out loud (e.g. 1 and 3 in the preceding example). Participants were told that the digit task and the

memory task were equally important and that they should devote as much attention as possible to each task.

Finally, participants completed three recognition tests following presentation and recall of all nine lists. Each test was composed of the 36 previously presented words associated with a ‘yes’ response in the encoding phase—12 words from each of three encoding lists. These 36 old words were mixed randomly with 36 new words drawn from the same source. The 36 old words in each recognition list were taken from three encoding lists presented in the same condition with respect to full or divided attention. That is, one of the recognition lists contained old words encoded and recalled under full attention, another contained old words encoded under DA but recalled under full attention, and the third contained old words encoded under full attention but recalled under DA.

The DA task was also performed during the retrieval phase of selected recall and recognition tasks. In the case of recall, DA at retrieval was performed on three of the encoding lists presented under full attention conditions; and in the case of recognition, DA at retrieval was performed on items drawn from these same three lists. In summary, of the nine encoded lists, three were performed under conditions of full attention at encoding followed by full attention during both recall and recognition tests; three were performed under conditions of full attention at encoding followed by DA during both recall and recognition tests; and three were performed under conditions of DA at encoding followed by full attention during both recall and recognition tests. Each recognition list contained 12 words associated with positive responses to each of the letter, rhyme, and associate encoding questions mixed randomly with 36 new (nonstudied) words. Participants responded to words presented in the recognition phase in one of three ways; if they recognized the word as old and also remembered some details of their reaction to it during the encoding phase, then they pressed a key labeled ‘Remember’; if they thought a word was old but could not recall any contextual reaction to it, then they pressed a key labeled ‘Know’; if they thought a word was new, they simply refrained from responding. In the recognition tests, each word was shown for 4 s, during which time they made their response (if the word was judged old). However, responses of R or K immediately caused the next word to appear.

With regard to randomization and counterbalancing of materials and conditions, we carried out the following manipulations. First, 32 participants saw encoding lists in which the 18 words in each list remained in the same list for all participants, with each word associated with the same encoding question. To create some diversity of materials, the remaining 16 participants were given lists in which the words were completely rerandomized with respect to encoding questions, but with no word repeating the type of question answered by the first 32 participants. For all participants, the words in each encoding list were presented in a different random order. The lists

themselves were organized in three blocks of three lists; each block was allocated randomly to one of the three encoding/retrieval combinations for each participant, and the order of the blocks was also randomized for each participant. Finally, each of the three recognition tests contained the 36 ‘yes’ words from one block, plus 36 new words; the three recognition tests were also presented in a different random order with respect to experimental condition to each participant.

The overall design was therefore entirely within participants, combining three levels of processing (letter, rhyme, and associate) with two answers to encoding questions (yes, no) and with three combinations of full and divided attention (full attention at both study and test; DA at study and full attention during both tests; full attention at study and DA during both recall and recognition tests). For recall, participants recalled both ‘yes’ and ‘no’ words, but in each recognition test only 36 ‘yes’ words were presented (12 each for letter, rhyme, and associate encoding) along with 36 new words.

Procedure Participants were tested individually. They were informed that the experiment concerned memory ability, and that they would be shown sets of common words and would then attempt to recall as many words as they could by writing them down on a sheet of paper in any order. The DA task was then explained to them. They were instructed to say aloud the third digit in a run of three consecutive odd digits, and short runs of digits were then presented as practice. Participants were informed that on some runs the DA task would be performed alone, and on other occasions it would be performed during either an encoding phase or during a recall or recognition test. The DA task was performed alone on two occasions—for 135 s before any memory lists were presented and again for 139 s after all memory tests were completed. In both cases there were 10 targets at unpredictable intervals. The baseline level of performance for each participant was calculated as the mean of performance on these two occasions. Participants were instructed to give equal weight to the DA and memory tasks in the DA conditions. Finally, before each list was presented, participants were informed of the conditions (full or divided attention) under which the study and test phases would be performed.

Results

Secondary task performance Mean probability of correct detection of targets on the three-successive-digit task performed on its own (averaged over two sessions) was 0.96. Performance on the digit task during encoding dropped to 0.57, performance during recall was 0.73, and performance during the recognition test was 0.62. Resource usage costs were therefore 0.39 for encoding, 0.23 for recall, and 0.34 for recognition. That is, retrieval costs were less than encoding costs for both recall and recognition. A one-way ANOVA on

all four conditions of digit task performance yielded a significant effect of condition, $F(3, 141) = 62.5, p < .001, \eta_p^2 = .571$. Subsequent pairwise comparisons showed that all values differed significantly from each other ($p < .01$), except for the difference between encoding and recognition.

Recall performance Given that 32 participants were tested on one combination of words and encoding conditions, and 16 participants on a second combination, we first tested for differences in recall levels for the two subgroups. Overall means for the subgroups of 32 and 16 were 0.23 ($SD = 0.18$) and 0.25 ($SD = 0.17$), respectively. A 2 (groups) \times 3 (attention) \times 3 (LOP) \times 2 (yes/no) ANOVA revealed no main effect of groups, $F(1, 46) = 1.06, p > .05$, and no interactions between group and any other factor (all F s < 1.3). The two subgroups were therefore combined into one group of 48 participants for further analyses.

Recall performance levels as a function of attention condition (full attention, DA at encoding, DA at retrieval), LOP (letter, rhyme, associate) and response type (yes, no) are shown in Fig. 1. For both yes and no responses, recall is best for the full attention condition, poorest for DA at encoding, and intermediate for DA at retrieval. The LOP manipulation shows that letter and rhyme encoding resulted in approximately equivalent performance levels, with higher recall levels for associate encoding. Performance was generally higher for words associated with yes responses than for those with no responses. These effects were confirmed in a three-way ANOVA (attention \times LOP \times yes/no) that showed significant effects of attention, $F(2, 94) = 65.1, p < .001, \eta_p^2 = .581$; of LOP, $F(2, 94) = 7.86, p < .001, \eta_p^2 = .143$; and of yes/no, $F(1, 47) = 36.8, p < .001, \eta_p^2 = .439$. There was one significant interaction, between LOP and yes/no, $F(2, 94) = 10.44, p < .001, \eta_p^2 = .182$. The interaction is attributable to a larger advantage of yes over no responses for associate encoding than for the other two differences (yes – no differences were 0.05 for letter, 0.03 for rhyme, and 0.12 for associate). The interaction between divided attention and yes/no was not significant, $F(2, 94) = 1.42, p > .05$. Separate ANOVAs for yes and no responses showed that for yes responses there were significant effects of divided attention, $F(2, 94) = 80.5, p < .001, \eta_p^2 = .631$, and of LOP, $F(2, 94) = 23.7, p < .001, \eta_p^2 = .335$; the interaction was not significant, $F(4, 188) = 1.67, p > .05$. In the case of no responses, the ANOVA yielded a significant effect of divided attention, $F(2, 94) = 20.6, p < .001, \eta_p^2 = .304$, but no effect of LOP or of the DA \times LOP interaction (F s < 1.0). For both the yes and no analyses, all three levels of the attention variable were significantly different from each other (p values $< .01$ in all six cases).

Recognition performance We first tested the two subgroups of 32 and 16 participants for differences in the overall hits minus false-alarm scores. The overall means for the subgroups of 32

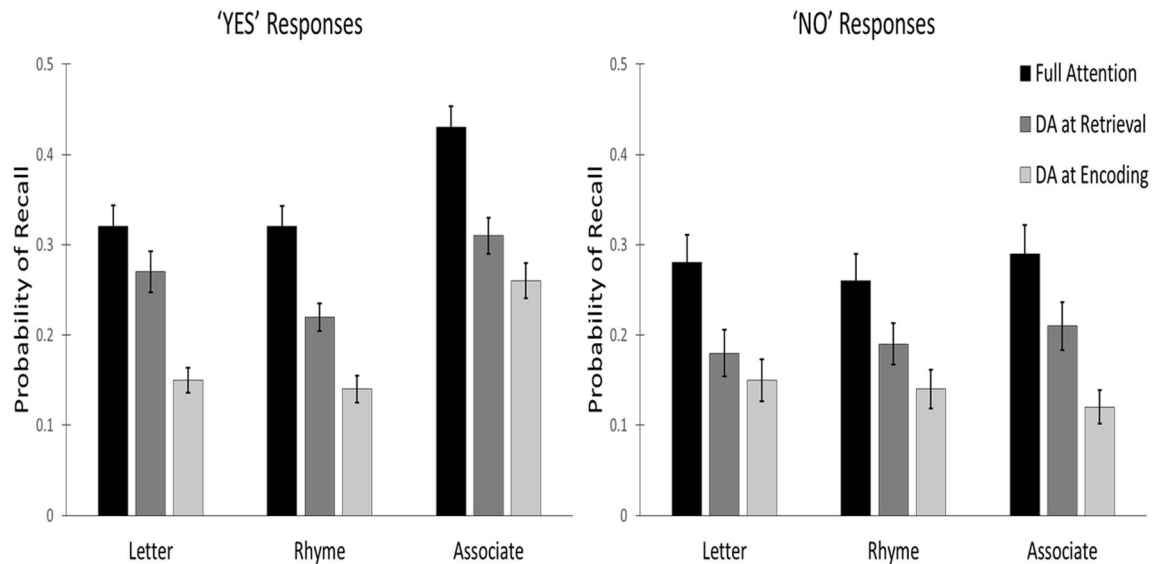


Fig. 1 Probabilities of recall as a function of attention, type of processing (letter, rhyme, associate), and type of response (yes, no). Error bars indicate standard error

and 16 were 0.55 (*SD* = 0.21) and 0.62 (*SD* = 0.19), respectively. A 2 (groups) × 3 (attention) × 3 (LOP) ANOVA revealed no main effect of groups, $F(1, 46) = 3.65, p > .05$, and no interactions between group and either attention or LOP (both F s < 1.2). We therefore again combined the scores into one group of 48 participants for further analyses.

Table 1 shows proportions correct for hits, false alarms (FAs), and for hits minus FAs. The table also shows values for hits and FAs for R and K judgments. The recognition tests involved only ‘yes’ items from the encoding conditions. FA rates are the same for letter, rhyme, and associate encoding conditions within each attention condition as the 36 new words were not assigned to any specific encoding condition.

Table 1 Recognition measures: Proportions of correct responses and false alarms (FA) for overall scores, R, K, and IRK judgments

Condition:	Full attention			DA encoding			DA retrieval		
	L	R	A	L	R	A	L	R	A
Overall hit rate:	0.85	0.85	0.90	0.62	0.66	0.73	0.78	0.78	0.87
Overall FA rate:	0.22	0.22	0.22	0.18	0.18	0.18	0.24	0.24	0.24
Hits – FA:	0.63	0.62	0.68	0.44	0.48	0.55	0.55	0.55	0.63
R Hits:	0.73	0.68	0.83	0.45	0.45	0.58	0.66	0.65	0.78
R FAs:	0.08	0.08	0.08	0.08	0.08	0.08	0.13	0.13	0.13
K Hits:	0.12	0.16	0.08	0.18	0.21	0.15	0.12	0.14	0.09
K FAs:	0.14	0.14	0.14	0.11	0.11	0.11	0.11	0.11	0.11
IRK Hits:	0.44	0.50	0.47	0.33	0.38	0.36	0.35	0.40	0.41
IRK FAs:	0.15	0.15	0.15	0.12	0.12	0.12	0.13	0.13	0.13

Note. DA = divided attention; LOP = levels of processing; L = letter judgment, R = rhyme judgment, A = associate judgment; IRK = independent R/K measure (see text)

An ANOVA on the overall hits minus FA data showed main effects of attention, $F(2, 94) = 20.2, p < .001, \eta_p^2 = .300$, and of LOP, $F(2, 94) = 14.7, p < .001, \eta_p^2 = .238$. The interaction was not significant ($F < 1.0$). In the case of attention, all three conditions were significantly different from each other ($p < .01$). For LOP, associate encoding was superior to both letter and rhyme encoding ($p < .001$), but letter and rhyme encoding did not differ ($p > .05$).

To evaluate R responses, we carried out an ANOVA on the hits minus FA scores shown in Fig. 2. The analysis yielded significant effects of both attention, $F(2, 94) = 38.3, p < .001, \eta_p^2 = .449$, and LOP, $F(2, 94) = 28.1, p < .001, \eta_p^2 = .374$. The interaction was not significant ($F < 1.0$). In the case of attention, all three conditions were significantly different from each other ($p < .01$), with the drop from full attention to DA at encoding (0.25) being substantially greater than the drop from full attention to DA at retrieval (0.09). For LOP, associate encoding was superior to both letter and rhyme encoding ($p < .001$), but letter and rhyme encoding did not differ. The pattern of results for R was thus very similar to that for overall recognition.

In order to compare our results with those from previous studies, we converted each participant’s K scores to measures of familiarity by dividing each K score by its corresponding value of 1 – R (Jacoby, 1998; Jacoby & Hay, 1998). Following Jacoby’s nomenclature, these adjusted K scores were renamed IRK scores (the independent R/K model; Jacoby, 1998). Figure 2 shows mean values of IRK minus their respective false alarm values. One problem in analyzing the resulting IRK scores was that several values of R in some conditions were 1.0 with corresponding K values of zero. It did not seem sensible to apply the formula $K/1 - R$ in such cases, so we transformed the 48 sets of K scores to 24 ‘macroparticipants’ by combining each K value of zero with

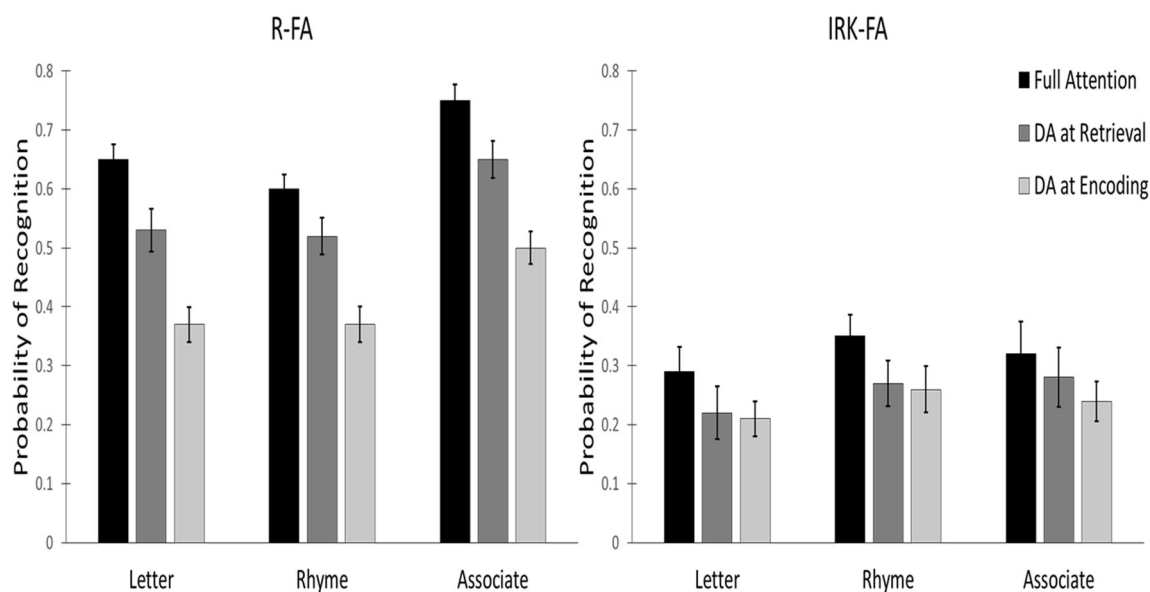


Fig. 2 Probabilities of recognition for R minus false alarms and IRK minus false alarms as a function of attention and type of processing (letter, rhyme, associate). Error bars indicate standard error

a second randomly chosen participant with a K value greater than zero. An ANOVA was then carried out on the resulting sets of IRK-FA values for the 24 new macroparticipants. The ANOVA on these data showed a significant effect of attention, $F(2, 46) = 6.33, p < .01, \eta_p^2 = .216$, but no effect of LOP, $F(2, 46) = 1.70, p > .05$, and no interaction ($F < 1.0$). In the case of attention, full attention was reliably superior to both DA at encoding and DA at retrieval ($p < .05$), which themselves did not differ reliably ($p > .05$).

Discussion

The results of the recall and recognition tests shown in Fig. 1 and Table 1, respectively, confirm a number of previous findings. There were significant effects of DA at both encoding and retrieval for both recall and recognition, with larger effects for DA at encoding in both cases. For recall, the average drop in proportion correct from full attention to DA was 0.16 for DA at encoding and 0.09 for DA at retrieval. For recognition, the comparable values were 0.15 and 0.06, respectively. For both recall and recognition, the mean values for DA at encoding were reliably lower than those for DA at retrieval. The drops in performance from full attention to DA at retrieval, though relatively small, were nonetheless significant. Thus, unlike the results obtained by Baddeley et al. (1984) and by Craik et al. (1996), but in line with the later results of Dodson and Johnson (1996), Gruppuso et al. (1997), Hicks and Marsh (2000), and Lozito and Mulligan (2006), division of attention during retrieval did result in a reliable drop in performance relative to full attention.

One atypical feature of the present results was the equivalence of letter and rhyme encoding manipulations; the ‘levels’

effect was carried entirely by the associate encoding condition. This equivalence does not seem to be attributable to effects of materials, as the same pattern was found in both sets of word–condition combinations—those performed by the groups of 16 and 32 participants. Other possibilities include the point that participants expected a memory test, so they may have processed letter and rhyme words more extensively than the tasks necessitated.

Results from performance on the secondary three-odd-digit task provide evidence that retrieval processes do require processing resources; the drop in performance from performing the task alone was 0.23 for recall and 0.34 for recognition. The greater processing cost for recognition is surprising in light of the opposite result obtained by Craik and McDowd (1987). Initially we thought that this surprising result might be attributable to the fact that the recall period lasted 60 s, which was longer than most participants needed, so they may have shifted attention to the secondary task. This possibility was assessed by counting the numbers of correct targets detected in the first versus second halves of each secondary task period; presumably target detection should be higher in the second half during recall if participants switched attention after exhausting their recall ability. However, target detection rates in the recall task were 0.71 and 0.72 in the first and second halves, respectively ($t < 1.0$); comparable figures for recognition were 0.58 and 0.62, respectively ($t = 1.05, p = .30$). It therefore appears that participants did not switch attention to the secondary task in either case. On the other hand, the recognition task was comparatively effortful in this experiment despite the fact that it was essentially self-paced. Participants took an average of only 1.47 s to answer each positive response, causing the next stimulus word to appear

immediately,¹ so they were fully engaged during the recognition test. By comparison, in the recall task participants recalled an average of only 4.23 words in 60 s under DA conditions, so it seems likely that recognition was atypically more attention demanding than recall in this experiment. However, the possibility that DA during recognition testing resulted in a comparatively slight memory decrement, because participants devoted more processing resources to recognition than they did to encoding, is ruled out by the present data. Although processing costs did not differ reliably between DA at encoding (0.39) and DA during recognition (0.34), the main point is that recognition costs were numerically *less* than encoding costs. The possibility that the asymmetry between encoding and recognition reflects a differential trade-off with the secondary task is thus not supported by the present results.

An ANOVA on the recall data (see Fig. 1) showed significant main effects of level of processing and yes > no responses, and also a significant interaction between yes/no and levels (the yes > no effect was greatest for associate encoding). These results are in line with previous findings (Craik & Tulving, 1975). However, DA did not interact with either yes/no or levels, so the present results do not support the claim that more meaningfully encoded items are more vulnerable to the effects of division of attention (Hicks & Marsh, 2000; Jacoby, 1991). When the yes and no recall data were considered separately, there were significant effects of DA in both cases, but the effect of levels was present only for yes responses. The absence of a levels effect for no responses is not in line with previous results (e.g. Craik & Tulving, 1975), which may again reflect the point that encoding was performed under conditions of intentional learning, leading to equivalent encoding operations between levels when a word was unrelated to the encoding question.

The overall recognition results are shown in Table 1; in this experiment only ‘yes’ response items to the levels questions were tested. An ANOVA on the hits minus false alarm data found significant effects of DA and of levels, but no interaction between these factors, replicating the recall results and also the results of Mulligan and Hirshman (1997). Figure 2 shows the recognition data separately for R and K responses; the values shown are hits minus false alarms in both cases, and the ‘know’ or ‘familiarity’ values, shown in the right-hand panel, are independent remember/know minus false alarms (IRK-FA) values. ANOVAs on these data showed that for R-FA values there were significant effects of DA and levels but

no interaction; for IRK-FA values there was a significant effect of DA but no effect of levels and no interaction. Previous studies have also found that level of processing affects R but not K responses (Gardiner, 1988), and this finding is echoed in studies in which estimates of recollection are separated from estimates of familiarity; level of processing affects the former but not the latter (Jacoby & Dallas, 1981). The significant effect of DA at both encoding and retrieval on IRK responses was unexpected in light of several previous studies, summarized by Kelley and Jacoby (2000), showing no effects of DA on familiarity. The present effect was small, however—an average drop of 0.06 from full attention to the values of DA at encoding and DA at retrieval.

The breakdown of recognition scores into R and IRK fractions is possibly the most novel aspect of the present study. If the relatively small effect of DA on recognition memory is attributable to participants basing their responses on feelings of familiarity rather than on recollection, then it might be expected that, in that condition, values of IRK would be substantial and values of R would decline relative to full attention. Table 1 shows that this is not the case, however; the average hit rate for R responses was 0.70, and the corresponding value for K responses was 0.12. That is, when attention is divided at retrieval, recognition performance is still carried largely by participants' feelings of recollection. Additionally, the relatively slight effect of DA at retrieval was found in recall as well as in the recognition data, and it seems improbable that recall is mediated by familiarity-based representations. It should be noted that several researchers have questioned the validity of the R/K distinction and the notion that recollection and familiarity represent qualitatively different aspects of remembering (e.g. Cohen et al., 2008; Ingram et al., 2012; Wixted & Mickes, 2010). Their alternative suggestion is that R and K simply represent one continuous dimension of memory strength from weak (K) to strong (R). However, Ingram and colleagues concede that under some conditions “performance is (or can be) based on two underlying dimensions” (Ingram et al., 2012, p. 335). These authors also comment (p. 338) that recollection and familiarity may be viewed as separate processes if they are differentially affected by some experimental manipulation, such as levels of processing—as was indeed found in the present experiment.

Jacoby's (1991) finding that deeply encoded items were more vulnerable to the effects of DA was not found generally in the present data. In the overall recognition data (see Table 1), differences between full attention and DA at retrieval were 0.08, 0.07, and 0.05 for letter, rhyme, and associate, respectively. Corresponding differences for the R minus false alarm scores (where an effect is most likely to be found) were 0.11, 0.08, and 0.09, respectively. The effect reported by Jacoby was also found by Hicks and Marsh (2000) but not by Lozito and Mulligan (2006) or by Mulligan and Hirshman (1997). The reasons for these discrepancies are unclear.

¹ The recognition test in the DA at retrieval condition took 192 s on average, in which time participants made decisions on 72 words. Table 1 shows that the mean probability of responding “old” in the DA at retrieval condition was $p = .53$ (average hit rate + false-alarm rate) translating to 38 yes responses and 34 no responses on average. Each target word judged “new” occupied the full 4-s presentation time (since the participant made no response), so distractor items took $34 \times 4 = 136$ s, leaving 56 s for the 38 yes responses, equivalent to 1.47 s per word to decide “old” and whether it should be labeled R or K.

Another feature of the present study was that participants performed the recognition tests on the same words they had previously attempted to recall. It is therefore possible that recognition decisions were based partly on memories of successful recall.² There are few signs of such a biasing effect in the data, however. In the recall data for ‘yes’ responses (left panel in Fig. 1), items encoded in the associate condition were recalled significantly more than items in the other two conditions, yet this benefit to associate items did not carry over to either R-FA values or IRK-FA values (Fig. 2). It therefore does not appear that associate items received a disproportionate boost from the preceding recall test.

Perhaps the most obvious reason for the relatively small effects of DA on retrieval is that participants maintain memory performance by neglecting the concurrent task, but there is essentially no evidence for this outcome (Experiment 1; Craik et al., 1996). However, another possibility is that participants do trade off performance, not between the two tasks, but *within* the memory retrieval task itself (e.g. maintaining accuracy by taking more time to respond under DA conditions). This possibility is supported by the findings of Baddeley et al. (1984), who showed convincingly that although the effects of a demanding concurrent task were negligible on retrieval accuracy, there was a consistent increase in retrieval latency under dual-task conditions. Craik et al. (1996) used a CRT task as the concurrent task in a series of DA experiments, and found that CRT latency increased as participants were instructed to place greater emphasis on the accompanying memory task; recognition latencies were not measured in this series of experiments, however. Combining Craik et al.’s finding of an increase in secondary task latency when emphasis is placed on recognition performance with those of Baddeley and colleagues suggests the possibility that emphasis on memory performance would decrease retrieval latency at the expense of increasing RTs on a concurrent task and vice versa. That is, processing time may trade off between the two tasks, despite the fact that retrieval accuracy is unaffected.

To assess this possibility Experiment 2 combined a recognition memory task, in which both accuracy and retrieval latency were measured, with a CRT secondary task from which latency was also recorded. Participants studied lists of word pairs during the encoding phase; in the test phase, word pairs were again presented, with half being the same pairs as at encoding (intact pairs) and the remainder being studied words but re-paired randomly. The participant’s task was to decide as rapidly as possible whether each test pair was intact or rearranged.

Following the model of Craik et al. (1996), dual-task performance was measured under three conditions of relative emphasis: perform as well as possible on the recognition task while also continuing to perform the CRT task; perform as

well as possible on the CRT task while also performing the recognition task; and pay equal attention to the two tasks. The predicted outcome was that as emphasis shifted from recognition to CRT, recognition latency would increase and CRT latency would decrease, but (in line with previous findings) recognition accuracy would remain constant. An additional exploratory question concerned possible trade-offs in *latency* between the recognition and CRT tasks. It would be interesting to find that faster RTs on one task were associated with equivalent amounts of slowing on the other task—while recognition *accuracy* remained stable despite changes in emphasis.

Experiment 2

Method

Participants The participants were 24 young adults (11 females) who were either fulfilling an undergraduate course requirement or were paid a stipend of \$12 CAD. The age range was 18–28 years, and the mean age was 20.5 years. The mean number of years of education was 14.8 years.

Tasks The recognition memory task comprised eight lists of 12 verbal paired associates, displayed visually on a computer monitor. The words were 192 common two-syllable nouns paired randomly. Two *versions* of the eight lists were constructed; each was a random re-pairing of the total set of words, resulting in 96 word pairs allocated randomly to the eight lists of 12 pairs. Two further sets of 12 pairs were used as practice lists, one with the recognition task performed alone and one in the dual-task setting. Within each list, six pairs were chosen randomly to remain intact in the test list, and the remaining six were randomly re-paired, although each word’s position in an A-B pair was retained. Within each of the two *versions*, A and B *formats* were constructed such that in format B the intact pairs in format A were now re-paired, and the formerly rearranged pairs in format A now remained intact. Thus across the 24 participants, sets of six participants received the same combination of study and test lists; additionally, word pair order within each list was randomized at both study and test for each participant separately. Finally, the order of conditions was counterbalanced across participants as described in the Design section. In the study phase, each word pair was exposed for 2.5 s followed by a plus sign displayed on the screen for 1.5 s. Note that the study phase was always performed under full attention conditions; DA conditions pertained only to the recognition test phase, presented after each study list following emphasis instructions for the relevant DA condition. In the test phase, the word pairs were presented in a different random order from the study phase. Each pair was again presented visually until the participant decided if it

² This point was suggested to us by Larry Jacoby.

was intact or re-paired. Participants conveyed their decision by pressing one of two response keys to indicate either ‘yes’ (= intact) or ‘no’ (= re-paired). This response caused the next word pair to appear immediately.

The DA task in this experiment was a continuous reaction time (CRT) task. Four adjacent boxes, arranged horizontally, were displayed on a computer monitor placed immediately above the monitor showing word pairs for the recognition task. An asterisk appeared in one of the four boxes, and the participant’s task was to press the response key corresponding to the asterisk’s position. The four response keys were also arranged horizontally, underneath the display. A correct response caused the asterisk to jump immediately to one of the other three boxes, cuing the next response. An incorrect key press had no effect on the asterisk’s location, so performance was measured entirely in terms of the total number of correct responses made in that trial. The total time for each trial was also recorded, so a further measure taken for CRT performance was average time per correct response. Two practice trials were given to all participants, who continued the task until they felt comfortable; the four scored baseline CRT trials each lasted 45 s.

Design As described in the Procedure section, the two tasks were first described to participants who were then given an opportunity to practice them, first separately and then together under dual-task conditions. After practice, 12 scored trials were performed. In the trial sequence 1–12, the CRT task alone was presented to all participants as Trials 1, 5, 8, and 12; performance on the CRT task alone provided a baseline measure of performance under full attention to enable comparisons with the task under dual-task conditions. Similarly, all participants were given the recognition task alone on Trials 3 and 10. The DA trials were therefore Trials 2, 4, 6, 7, 9, and 11—two trials for each of the three emphasis conditions—emphasize the recognition test (DA.Rg), ‘50-50’—emphasize both tasks equally (DA.50), and emphasize the CRT task (DA.CRT). Additionally, the DA Trials 7, 9, and 11 always mirrored the order of the three DA conditions given in Trials 2, 4, and 6; for example, if Trials 2, 4, and 6 were represented by the emphasis conditions CRT, Rg, 50, respectively, then Trials 7, 9, and 11 were represented by emphasis conditions 50, Rg, CRT, respectively. The three emphasis conditions have six possible orderings for Trials 2, 4, and 6 across participants (CRT, Rg, 50; CRT, 50, Rg; Rg, 50, CRT; Rg, CRT, 50; 50, CRT, Rg; 50, Rg, CRT), so these orderings (with their mirror orderings for Trials 7, 9, and 11) provided six counterbalanced versions. These six orderings were crossed with the two *versions* described under Tasks and with the two *formats* (A & B) of each recognition list, yielding 24 different presentations. That is, each of the 24 participants performed a unique version of the study. Note that the CRT task was randomized on each trial, so no further counterbalancing was necessary.

Procedure The CRT task was first described, and the participant was then given two self-paced practice trials. Each participant performed the task using his or her dominant hand, and speed of responding was emphasized. The recognition task was then explained and practiced under single-task conditions. Participants were advised that while learning the word pairs, it helps to make a meaningful connection—an image or short story—between the words. Participants were instructed that they should be as accurate as possible, but also to respond as rapidly as possible. They were also shown that the next word pair appeared as soon as they pressed either response key. Participants were then given one further practice trial of 12 word pairs while also performing the CRT task during the recognition test phase. The emphasis conditions were then conveyed by the following instruction: “Before each test trial I will tell you whether the CRT task or the recognition task is the more important—please try to perform the more important task as well as when you do it alone, while doing the other task as well as possible. The third emphasis condition is 50-50—the two tasks are equally important.” The relevant emphasis instruction was given before each DA test trial. Participants then proceeded to perform the 12 scored trials.

Results

Recognition accuracy Table 2 shows the proportions of correct recognition responses as the mean for the two full attention conditions and for all six DA conditions; that is, the two trials under each of the three emphasis conditions are shown separately. We argue that participants may interpret and act on the emphasis instructions somewhat differently, even from trial to trial, and that if there is a trade-off between tasks, then a person who pays more attention to one task on a given trial should have fewer resources to apply to the other task. In this sense, emphasis may be more of a continuous variable than a set of categories. This assumption allowed us to plot more data points to illustrate possible trade-offs between the recognition and CRT tasks. Recognition accuracy performance is also shown for the mean full attention and DA conditions in Fig. 3a. Replicating previous results, the table and figure show that performance was highest under full attention, but that performance under DA conditions is only slightly lower and does not vary systematically as a function of DA emphasis. The means of the two replication conditions for each DA task were 0.77, 0.80, and 0.76 for the emphasis conditions DA.Rg, DA.50, and DA.CRT, respectively. A one-way ANOVA comparing these values to full attention (0.83) was not significant, $F(3, 92) = 1.31, p > .05$. A further one-way ANOVA on the six DA conditions shown in Table 2 was also nonsignificant ($F < 1.0$). The ‘standard’ finding that divided attention at retrieval has either no effect on recognition memory (e.g. Baddeley et al., 1984; Craik et al. 1996), or very slight effects (e.g. Dodson

Table 2 Recognition accuracy, decision latency for Rg and CRT tasks, also numbers of responses in 6 s for Rg and CRT

Condition	Prop. correct Rg	Rg latency	CRT latency	Rg responses	CRT responses
Full attn.	0.83	1211	434	5.07	14.02
DA.Rg1	0.74	1842	2459	3.49	2.98
DA.Rg2	0.80	1629	2458	3.89	3.06
DA.501	0.79	1865	1243	3.39	5.24
DA.502	0.81	2013	1088	3.12	6.10
DA.CRT1	0.76	2891	834	2.40	7.97
DA.CRT2	0.75	2863	724	2.40	8.79

Note. Rg = recognition, CRT = continuous response task, DA = divided attention. Latencies are in ms. Rg and CRT correct responses in 6 s. The numbers 1, 2 (e.g. DA.Rg1) refer to the first and second replications of each DA condition

& Johnson, 1996; Lozito & Mulligan, 2006), was thus replicated in the present experiment.

Decision latencies Table 2 also shows the average decision latencies for (a) correct recognition decisions and (b) individual CRT decisions; mean values for each condition are also shown in Fig. 3b. The data show first that both CRT latencies and recognition decision times were slowed considerably by performing the tasks under DA conditions, confirming earlier results by Baddeley et al. (1984) and Craik et al. (1996). These effects of DA were assessed by separate ANOVAs on the recognition and CRT tasks. A one-way ANOVA on the four means for recognition latency shown in Fig. 3b revealed a main effect of condition, $F(3, 92) = 25.20, p < .001, \eta_p^2 = 0.451$. Subsequent post hoc analyses (Tukey's HSD) found that the full attention condition was only slightly faster than DA.Rg ($p = .044$), showing that participants followed instructions to perform the recognition task under this DA condition

as rapidly as under full attention. The other two DA conditions were associated with latencies that were significantly longer than under full attention. A corresponding ANOVA on the CRT latency data shown in Fig. 3b found an overall effect of condition, $F(3, 92) = 47.66, p < .001, \eta_p^2 = 0.608$. Subsequent post hocs found that CRT latency under full attention was significantly faster than both DA.Rg and DA.50, but was not faster than DA.CRT ($p = .233$), again showing that participants followed instructions to perform the CRT task as well under DA.CRT conditions as under full attention. The latency value for DA.Rg was significantly greater than values for DA.50 and DA.CRT, but the values for DA.50 and DA.CRT did not differ significantly.

Figure 3b shows that whereas recognition latency rises from emphasis on recognition to emphasis on the CRT task, the CRT latencies decline over the same range of emphasis conditions. This trade-off in latencies between the two tasks was assessed by a two-way ANOVA utilizing all six emphasis

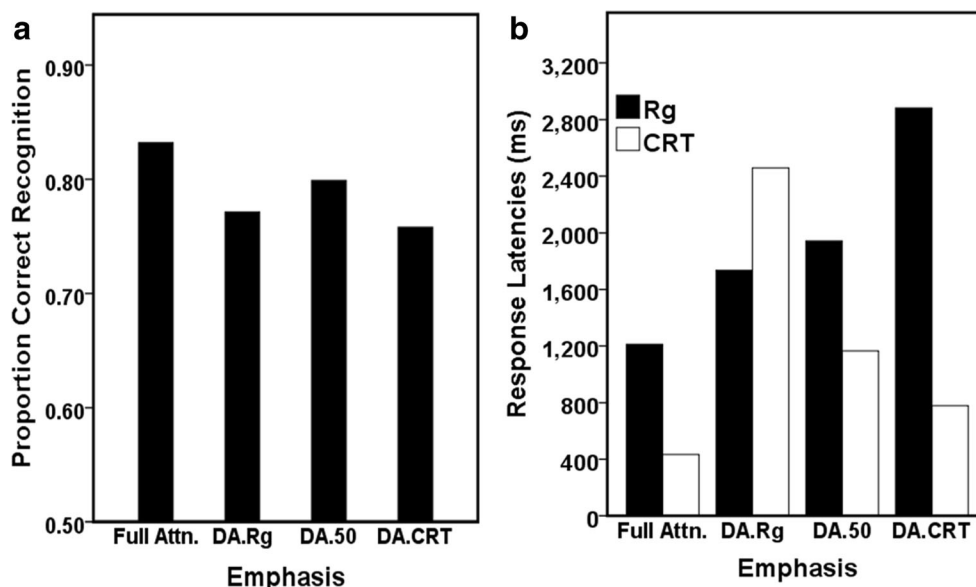


Fig. 3 a Proportions of correct recognition (Rg) responses for full attention and three levels of emphasis b Response latencies for Rg and CRT tasks as a function of attention and emphasis

values shown in Table 2. The analysis revealed a main effect of task, $F(1, 276) = 59.49, p < .001, \eta_p^2 = 0.18$, a main effect of emphasis, $F(5, 276) = 4.71, p < .001, \eta_p^2 = 0.08$, and a reliable interaction between the two variables, $F(5, 276) = 31.06, p < .001, \eta_p^2 = 0.36$. The task effect reflects the fact that recognition latencies were longer than CRT latencies (2,184 ms vs. 1,468 ms, respectively); the emphasis effect shows that overall latencies were shorter in the DA.50 condition (1,553 ms) than in either the DA.Rg condition (2,098 ms) or the DA.CRT condition (1,828 ms). The highly significant interaction confirms the trade-off pattern between the two tasks.

Further one-way ANOVAs were carried out on the data in Table 2 to examine each task in greater detail. CRT latencies were strongly affected by DA emphasis, $F(5, 138) = 25.48, p < .001$, again in line with previous results (Craik et al., 1996). Post hoc comparisons among the six CRT latencies (Tukey's HSD) showed that the comparison between DA.Rg1 and DA.Rg2 was not reliable ($p > .90$), but that both DA.Rg values were significantly different from the other four CRT values ($p < .001$), which did not differ among themselves (all p values $> .18$). More interestingly, recognition latencies were *also* systematically affected by DA emphasis, being shorter under Rg emphasis, and progressively longer under conditions of equal emphasis and CRT emphasis (means of 1,736, 1939 and 2,877 ms, respectively). A one-way ANOVA on the six DA emphasis conditions for this measure showed a significant difference among the six latencies, $F(5, 138) = 11.10, p < .001$. Follow-up post hoc analyses (Tukey's HSD) found that none of the pairs at each level of DA emphasis (e.g. DA.Rg1 and DA.Rg2) differed significantly (all p values $> .90$); similarly, the latencies for the four DA.Rg and DA.50 conditions did not differ (all p values $> .50$); but latencies for the two DA.CRT conditions were significantly longer than those for the other four conditions (all p values $< .01$).

In summary, the decision *latency* data for the recognition and CRT tasks displayed in Table 2 and Fig. 3b show that performance levels on the two tasks *do* trade off against each other under variable conditions of DA, despite the concurrent finding that recognition *accuracy* levels remained constant across the same range of differential DA emphasis conditions. A further point to note is that the rather conservative post hoc tests show that the major differences within the recognition latency data are between the DA.CRT condition and the other DA conditions, which do not differ among themselves. Correspondingly, the major differences within the CRT latency data are between the DA.Rg condition and the other DA conditions, which also do not differ among themselves. That is, the greatest increases in latency for both tasks are associated with the condition in which performance on the other task is emphasized. Emphasis to one task causes delays on the other task, while accuracy levels remain constant.

Response rates As discussed below, the finding of a trade-off between CRT latency and recognition latency essentially solves the problem of why DA at retrieval has such a small effect. Apparently participants show a strong desire to perform as well as possible on recognition accuracy, and simply defer their recognition responses until they feel comfortable with their decision. As processing effort is progressively switched to the CRT task, responses on that task become faster, leaving less attentional capacity for recognition decisions which are therefore slowed. Figure 3 thus provides an answer to the puzzle about DA at retrieval, but we had no preconceived ideas about the form of a possible trade-off function. The simplest function would presumably be a symmetrical linear trade-off in response latencies between the CRT and recognition tasks, but Fig. 3 shows that is not the case. For example, recognition latency increases by an average of 203 ms from DA.Rg emphasis to DA.50 emphasis, whereas average CRT latency decreases by 1,293 ms. Similarly, CRT latency decreases by 387 ms from DA.50 to DA.CRT, whereas recognition latency increases by 938 ms. A plot of the relationship between CRT and Rg latencies exhibited asymptotic properties, so we strategically examined both logarithmic and reciprocal transforms of the time measures; the latter was found to fit the relationship better. The reciprocal of latency gives a measure of the *rates of responding* for the two tasks in a given time in the three emphasis conditions. In order to assess this function, we chose the arbitrary time of 6 s as a unit that was easy to grasp and was meaningful in the context of our two tasks. Table 2 thus shows the average numbers of CRT and recognition responses in 6 s under both full attention (averaged over four and two conditions for CRT and Rg, respectively) and under the six different emphasis conditions; these numbers were calculated by taking the reciprocal of CRT and recognition latencies in seconds, and multiplying by 6.

A within-participant product-moment correlation between the response rates for the six pairs of DA conditions was calculated for each participant. Fisher-transformed coefficients were averaged and then back-transformed; it was found that $r = -.825$, with a corresponding value of $R^2 = .68$. This result suggests a strong linear relationship between the two response rates. A second way to assess the goodness of fit to a linear relationship is given by the proportion of variance explained by a linear combination of the two rates. We calculated this measure using eigenvalues of the covariance matrix for the two rates across the six DA emphasis conditions. For all 24 participants, the median proportion of variance was 0.98 (95% CI for the median: [0.96, 1.00]); the first and third quartiles were 0.97 and 0.99. We also wished to explore the consistency of the observed linear relation, first to assess the generality across participants and second to see if the linear association is modified by general ability. Accordingly, we split the participants into four quartiles of six participants on the basis of their overall recognition accuracy. Medians within

each quartile are presented in Table 3 for the proportions of correct recognition trials (accuracy), proportions of variance explained by a linear association calculated as for the whole group of 24, and linear trade-off costs. For Groups A, B, C, and D, the median proportions of correct recognition responses were 0.94, 0.83, 0.72, and 0.61, respectively (chance responding = 0.50). The functions relating the two response rates are shown for all four quartile groups in Fig. 4. The straight-line functions are the best-fit linear regressions of CRT on Rg using the six emphasis condition means in each quartile. The figure, in conjunction with Table 3, shows that the functions for Groups A, B, and C are strikingly linear, but group D showed clear departures from an optimal linear relationship, possibly because these lowest performing participants did not take the task seriously enough, were not effectively induced to vary emphasis between the two tasks in response to the instructions, or for some other reason were unable to divide attention between the two tasks. The proportions of variance accounted for by linear functions are somewhat different when assessed as correlations using the means of the six emphasis conditions (see Fig. 4) and when taking median values of variance calculated from eigenvalues of the covariance matrix for separate participants (see Table 3). The second method is probably preferable, but for present purposes the point is that a strong linear relationship between the two tasks holds for the majority of the participants.

In order to measure the trade-off relation between the two response rates, the best-fit linear relationship between CRT and recognition response rates was calculated for each participant using the first eigenvector of their covariance matrix. Across all 24 participants, the median trade-off between the two tasks was -3.72 CRT responses for each recognition response (95% CI for the median: $[-11.6, 4.2]$); the first and third quartiles were -4.83 and -2.71 . That is, for an increase of one further correct recognition response, the rate of responding to the CRT task was reduced by 3.72 responses in the same interval.

Discussion

In our opinion the long-standing puzzle of why divided attention during memory retrieval has such a small effect is solved

Table 3 Medians for recognition accuracy (proportions correct), proportions of variance accounted for by a linear association, and linear trade-off costs for quartiles

Quartile	Prop. correct	Prop. variance	Trade-off costs
A	0.94	0.58	-3.72
B	0.83	0.73	-4.48
C	0.72	0.86	-3.61
D	0.61	0.02	-1.28

by the finding that accurate performance holds up at the expense of longer decision times. This general point was already made by Baddeley et al. (1984), but the present experiment adds the finding that the processing times trade off rather precisely between a recognition task and a continuous RT task. The further result that the trade-off between the present tasks is a linear function involving response rates rather than decision latencies was unexpected but well supported by the strength of the variance accounted for by linear functions. The linearity was shown by most participants (see Fig. 4), but clearly broke down for the participants in quartile D. From the present results, it is not possible to say why these poorest performing participants show functions that depart from linearity; speculatively, these participants were either not motivated to take the experiment seriously or were somehow incapable of dividing their attention effectively.

General discussion

To summarize the main results, Experiment 1 confirmed previous findings of a greater effect of DA at encoding than at retrieval; the effects of DA at retrieval were comparatively slight (though significant) for both recall and recognition. The study also confirmed that retrieval under DA conditions does consume processing resources, and that the small effect of DA at retrieval was not simply due to a greater trade-off of resources with the secondary task. In both the recall test and the overall recognition data there were significant effects of both divided attention and levels of processing, but no interaction between these variables. The possibility that DA at retrieval in recognition testing has relatively small effects because recognition is mediated by familiarity rather than recollection was not supported by the present results, given that recognition under that condition was attributable largely to R responses rather than K responses (means of 0.70 and 0.12, respectively). We did not replicate the finding reported by Jacoby (1991) and by Hicks and Marsh (2000) that semantically encoded words were differentially vulnerable to the effects of DA, in that DA did not interact with level of processing for either recall or recognition (see also Lozito & Mulligan, 2006; Mulligan & Hirshman, 1997). Finally, when K values were transformed to IRK values (see Fig. 2, right panel), it was found that levels had no effect on IRK, but that DA did reduce the values by a small but significant amount.

Experiment 2 does substantially more to solve the puzzle of the relatively small effects of DA on retrieval. By measuring decision latencies on both the recognition task and the secondary task, the experiment demonstrated that response rates on the two concurrent tasks trade off against each other while recognition accuracy remains constant. Thus, as noted many years ago by Baddeley et al. (1984), DA at retrieval is associated with an increase in response latency, although accuracy is

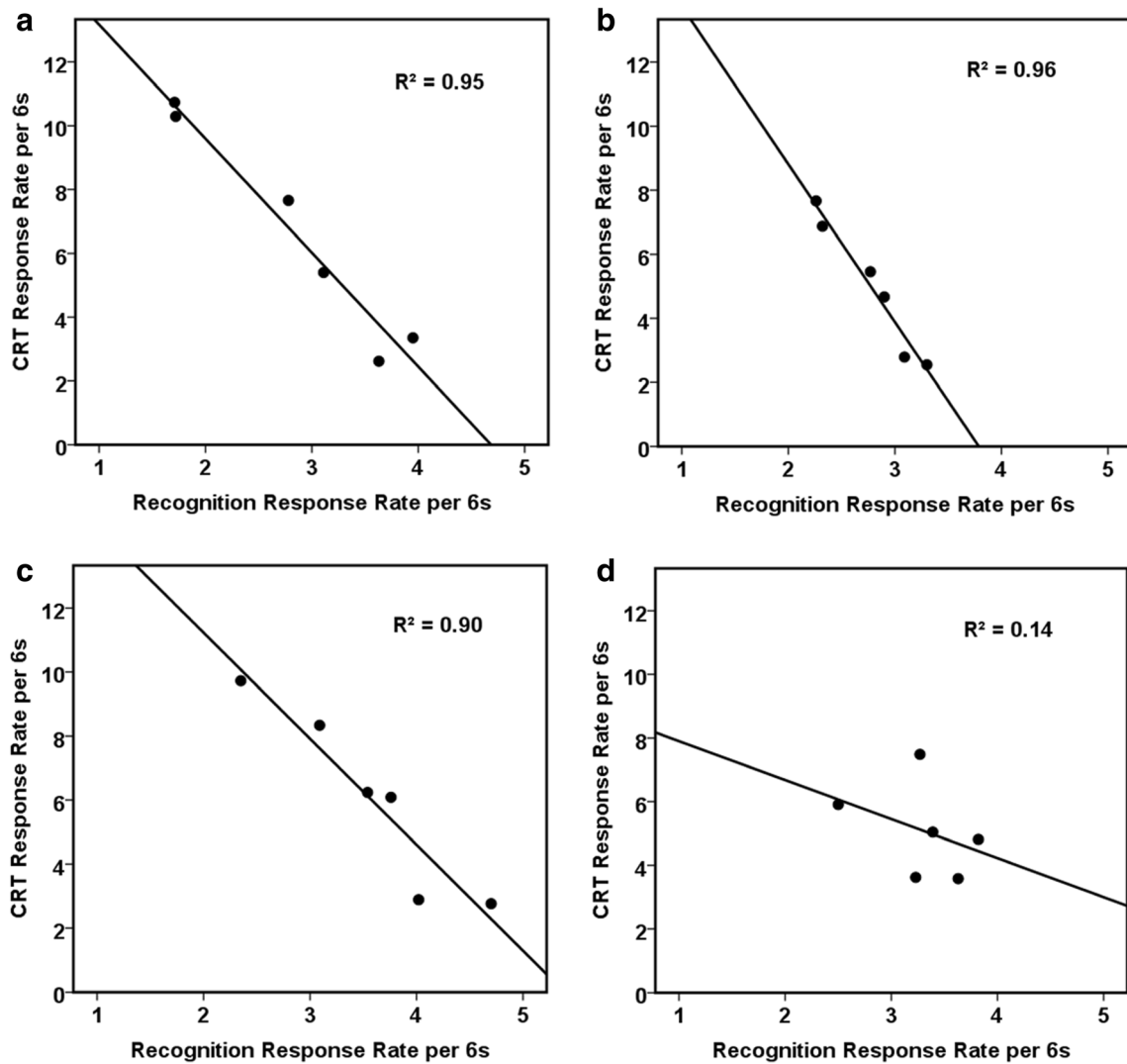


Fig. 4 Response rates for recognition (R_g) and CRT tasks as a function of six emphasis conditions. Data are shown in quartiles in terms of R_g accuracy (see text). The functions shown are best-fit linear regressions of CRT on R_g , using means of six participants within each quartile

unaffected. Rohrer and Pashler (2003) suggested that retrieval processing may be blocked momentarily by concurrent processing of the secondary task to account for their finding that free recall was reduced in both total recall and speed of recall by a demanding concurrent RT task. We prefer an alternative speculation that participants continue to process the recognition stimuli until the system accrues sufficient evidence to satisfy some internal criterion of confidence in their decision. As processing resources are progressively diverted from recognition processing to the concurrent task, the accumulation of evidence is systematically slowed, and it takes longer to satisfy this criterion value. It is possible that participants feel that their ability level will be shown directly by accurate performance on the memory task, whereas the time they take to make the decision is of lesser importance. In this sense perhaps the finding that recognition accuracy is affected only slightly by DA is not so much that recognition is ‘obligatory’ (Craik et al., 1996; see also Rohrer & Pashler, 2003), but that

participants choose to accrue evidence until some criterion of acceptability is reached.

We had no specific expectations regarding the nature of the trade-off function between the recognition and CRT tasks, but found that the relation was well captured by a linear function relating the two response rates. In the present experiment, the trade-off relation was a reduction of 3.72 CRT responses for each additional correct recognition decision in the same time interval. As shown by the overall assessment, a linear function accounted for over 90% of the variance, and this strong relationship also held for the 18 of the 24 participants who scored highest in recognition accuracy. Further work is clearly required to explore this interesting relationship and to see whether the performance of participants who do not fit the linear model can be modified by offering higher rewards, or whether some participants have an inherent inability to divide attention between two concurrent tasks. Although this is a controversial issue (see, e.g. Baddeley et al., 1984), the

finding that response rates do trade off rather exactly between the two tasks appears to be in line with the notion of a general pool of attentional resource (e.g. Kahneman, 1973) that can be drawn on differentially by two or more tasks.

We acknowledge that the differences between DA at encoding and retrieval are not entirely solved at this point. Why exactly is it that DA at encoding decreases subsequent memory so reliably whereas DA at retrieval has such a small effect on recall and recognition? Speculatively, it may be suggested that DA at encoding is likely to reduce the depth and elaboration of encoded stimuli, resulting inevitably in lower memory performance. Additionally, there is good evidence that individuals are poor at judging the effectiveness of their encoding processes and are correspondingly weak at predicting subsequent memory levels (Shaw & Craik, 1989). Thus, participants in a memory experiment may feel that they have encoded information sufficiently, whereas in fact they have not. At retrieval, in contrast, success or lack of it is more obvious, especially in the case of recall; even recognition testing is more ‘public’ than encoding in the sense that participants’ recognition performance can be judged by the experimenter. Thus, plausibly, participants devote more attention and effort to retrieval than to encoding. The main conclusions of the present article, however, are first that division of attention during retrieval has minimal effects on recognition memory accuracy levels because retrieval processing times are lengthened, and second that there is a lawful linear relationship between response rates on the recognition task and response rates on a concurrent continuous reaction-time task.

Acknowledgements This work was supported by a research grant from the Natural Sciences and Engineering Research Council of Canada to Fergus Craik (A8261). Correspondence concerning the article should be addressed to Fergus Craik, Rotman Research Institute at Baycrest, 3560 Bathurst St., Toronto, Ontario, Canada, M6A 2E1. Email: fcraik@research.baycrest.org

References

- Baddeley, A., Lewis, V., Eldridge, M., & Thomson, N. (1984). Attention and retrieval from long-term memory. *Journal of Experimental Psychology: General*, *113*, 518–540. doi:<https://doi.org/10.1037/0096-3445.113.4.518>
- Cohen, A. L., Rotello, C. M., & MacMillan, N. A. (2008). Evaluating models of remember-know judgments: Complexity, mimicry, and discriminability. *Psychonomic Bulletin & Review*, *15*, 906–926. doi:<https://doi.org/10.3758/PBR.15.5.906>
- Craik, F. I. M. (1982). Selective changes in encoding as a function of reduced processing capacity. In F. Klix, J. Hoffmann, & E. Van der Meer (Eds.), *Cognitive research in psychology* (pp. 152–161). Berlin, Germany: DVW.
- Craik, F. I. M., Govoni, R., Naveh-Benjamin, M., & Anderson, N. D. (1996). The effects of divided attention on encoding and retrieval processes in human memory. *Journal of Experimental Psychology: General*, *125*, 159–180. doi:<https://doi.org/10.1037//0096-3445.125.2.159>
- Craik, F. I. M., & McDowd, J. M. (1987). Age differences in recall and recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *13*, 474–479.
- Craik, F. I. M., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General*, *104*, 268–294. doi:<https://doi.org/10.1037/0096-3445.104.3.268>
- Dodson, C. S., & Johnson, M. K. (1996). Some problems with the process-dissociation approach to memory. *Journal of Experimental Psychology: General*, *125*, 181–194. doi:<https://doi.org/10.1037/0096-3445.125.2.181>
- Gardiner, J. M. (1988). Functional aspects of recollective experience. *Memory & Cognition*, *16*, 309–313. doi:<https://doi.org/10.3758/bf03197041>
- Gardiner, J. M., & Richardson-Klavehn, A. (2000). Remembering and knowing. In E. Tulving & F. I. M. Craik (Eds.), *The Oxford handbook of memory* (pp. 229–244). New York, NY: Oxford University Press. doi:<https://doi.org/10.1002/acp.822>
- Gruppuso, V., Lindsay, D. S., & Kelley, C. M. (1997). The process-dissociation procedure and similarity: Defining and estimating recollection and familiarity in recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*, 259–278. doi:<https://doi.org/10.1037/0278-7393.23.2.259>
- Guez, J., & Naveh-Benjamin, M. (2006). Divided attention at encoding and retrieval for once- and thrice-presented items: A micro-level analysis of attentional costs. *European Journal of Cognitive Psychology*, *18*, 874–898. doi:<https://doi.org/10.1080/09541440500485854>
- Hicks, J. L., & Marsh, R. L. (2000). Toward specifying the attentional demands of recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 1483–1498. doi:<https://doi.org/10.1037/0278-7393.26.6.1483>
- Ingram, K. M., Mickes, L., & Wixted, J. T. (2012). Recollection can be weak and familiarity can be strong. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *38*, 325–339. doi:<https://doi.org/10.1037/a0025483>
- Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language*, *30*, 513–541. doi:[https://doi.org/10.1016/0749-596x\(91\)90025-f](https://doi.org/10.1016/0749-596x(91)90025-f)
- Jacoby, L. L., & Dallas, M. (1981). On the relationship between autobiographical memory and perceptual learning. *Journal of Experimental Psychology: General*, *3*, 306–340. doi:<https://doi.org/10.1037/0096-3445.110.3.306>
- Jacoby L. L. (1998) Invariance in automatic influences of memory: Toward a user's guide for the process-dissociation procedure.. *Journal of Experimental Psychology: Learning, Memory, and Cognition* *24* (1):3–26
- Jacoby, L. L., & Hay, J. F. (1998). Age-related deficits in memory: Theory and application. In M. A. Conway, S. E. Gathercole, & C. Cornoldi (Eds.), *Theories of memory* (Vol. 2, pp. 111–134). Hove, UK: Psychology Press.
- Jacoby, L. L., Kelley, C., Brown, J., & Jasechko, J. (1989a). Becoming famous overnight: Limits on the ability to avoid unconscious influences of the past. *Journal of Personality and Social Psychology*, *56*, 326–338. doi:<https://doi.org/10.1037/0022-3514.56.3.326>
- Jacoby, L. L., Kelley, C. M., & Dywan, J. (1989b). Memory attributions. In H. L. Roediger & F. I. M. Craik (Eds.), *Varieties of memory and consciousness: Essays in honour of Endel Tulving* (pp. 391–422). Hillsdale, NJ: Erlbaum.
- Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs, NJ: Prentice Hall.
- Kelley, C. M., & Jacoby, L. L. (2000). Recollection and familiarity: Process-dissociation. In E. Tulving & F. I. M. Craik (Eds.), *The Oxford handbook of memory* (pp. 306–340). New York, NY: Oxford University Press.

- Kellogg, R. T., Cocklin, T., & Bourne, L. E. (1982). Conscious attentional demands of encoding and retrieval from long-term memory. *The American Journal of Psychology*, *95*, 183–198. doi:<https://doi.org/10.2307/1422465>
- Lozito, J. P., & Mulligan, N. W. (2006). Exploring the role of attention during memory retrieval: Effects of semantic encoding and divided attention. *Memory & Cognition*, *34*, 986–998. doi:<https://doi.org/10.3758/bf03193246>
- Mandler, G. (1980). Recognizing: The judgment of previous occurrence. *Psychological Review*, *87*, 252–271. doi:<https://doi.org/10.1037/0033-295x.87.3.252>
- Mulligan, N. W. (1998). The role of attention during encoding in implicit and explicit memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 27–47. doi:<https://doi.org/10.1037/0278-7393.24.1.27>
- Mulligan, N. W., & Hirshman, E. (1997). Measuring the bases of recognition memory: An investigation of the process-dissociation framework. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*, 280–304. doi:<https://doi.org/10.1037/0278-7393.23.2.280>
- Murdock, B. B., Jr. (1965). Effects of a subsidiary task on short-term memory. *British Journal of Psychology*, *56*, 413–419. doi:<https://doi.org/10.1111/j.2044-8295.1965.tb00983.x>
- Rohrer, D., & Pashler, H. E. (2003). Concurrent task effects on memory retrieval. *Psychonomic Bulletin & Review*, *10*, 96–103.
- Shaw, R. J., & Craik, F. I. M. (1989). Age differences in predictions and performance on a cued recall task. *Psychology and Aging*, *4*, 131–135.
- Wixted, J. T., & Mickes, L. (2010). A continuous dual-process model of remember/know judgments. *Psychological Review*, *117*, 1025–1054. doi:<https://doi.org/10.1037/a0020874>