**The role of working memory capacity in the control of recollection**

**Rachael L. Elward\*, Lisa H. Evans, Edward L. Wilding**

Cardiff University Brain Research Imaging Centre (CUBRIC)

School of Psychology, Cardiff University,

Cardiff, CF10 3AT, Wales, UK.

**Running Head:** Control of Recollection

**\* Corresponding Author:** Rachael Elward (Rachael.elward@utdallas.edu), Centre for Vital Longevity, University of Texas at Dallas, 1600 Viceroy Drive, Suite 800, Dallas, TX 75080. Tel: +1 469 341 7738; Fax: +1 972 883 3250.

**Acknowledgements:** This research was supported by the Wales Institute of Cognitive Neuroscience (WICN) and the UK Biotechnology and Biological Sciences Research Council (BBSRC). **Abstract:** The links between control over recollection and working memory capacity (WMC) were investigated using event-related potentials (ERPs) and behavioural assays. Electrophysiological evidence for a relationship between greater control over recollection and higher scores on a measure of WMC were obtained. In addition, people with high WMC who first completed a task requiring cognitive control showed no electrophysiological evidence for control over recollection on a subsequent task. This outcome suggests a causal link between control over recollection and the availability of WMC, in so far as the consequence of completing the first task was a time-limited reduction in WMC that impacted on completion of the subsequent task. All participants also completed a final recall task, on which they were asked to remember the stimuli they had encountered during the task in which ERPs were acquired. Only those participants who showed electrophysiological evidence for the exertion of control over recollection showed differences in the likelihood of recall between stimuli over which control either had or had not been exerted. In combination, the findings provide insights into the conditions under which control over recollection occurs, and make a strong argument for including individual difference measures of resource availability when assessing how and when people exert control over what they remember.

**Key Words:** Recollection, Episodic Memory, Cognitive Control, Event-related potentials (ERPs), Inhibition. **1. Introduction:** Recollection is a retrieval process associated with the recovery of detailed qualitative information about studied events. The process is widely assumed to be under some degree of voluntary control {e.g. /Burgess, 1996 #101;Johnson, 1992 #423;Schacter, 1998 #828;Yonelinas, 2002 #1064}. In this paper, we describe findings in a study where event-related potential (ERP) data were acquired, alongside behavioural assessments, to understand: (i) how control over recollection is exerted when people must adjudicate between the contexts in which events were encountered, and (ii) pre-requisites for exerting control over recollection.

The study described here was motivated by findings in a series of ERP studies in which systematic changes in the magnitude of an electrophysiological index of recollection – the left-parietal ERP old/new effect – have been used to infer that some form of control over recollection has taken place. ERP old/new effects are measured by comparing neural activities that are elicited in response to old (previously studied) and new (unstudied) stimuli that attract correct old/new judgments during a retrieval task {Rugg, 1994 #776}. The left-parietal ERP old/new effect comprises a greater relative positivity for old than for new stimuli, which is largest at left-posterior/parietal scalp locations between approximately 500 and 800 msec post-stimulus {Wilding, 2003 #1028}. The evidence linking this effect to the process of recollection is substantial, and is not re-reviewed here {see / Donaldson, 2003 #176;Friedman, 2000 #242;Rugg, 2007 #1436}. The key finding to emphasise is that the left-parietal effect is sensitive to either the amount or quality of contextual information that is recovered from memory {Vilberg, 2006 #1382;Vilberg, 2009a #1552;Vilberg, 2009b #1553;Wilding, 2000 #1025}. Consequently, changes in the magnitude of the effect have been argued to index the extent to which recollection has occurred {Parkin, 1995 #1221}. Moreover, in some circumstances, changes in the magnitude of this ERP effect across critical experimental conditions have been used to make inferences about when control has been exerted over the extent to which recollection occurred {Evans, 2010 #1572;Herron, 2003 #1089;Wilding, 2006 #1369}.

The critical ERP data for this latter inference were acquired in exclusion tasks, where participants are exposed to an equal number of study items in two separate contexts {Jacoby, 1991 #397}. Participants make a binary response on an ensuing memory test, during which they are exposed to study items from both contexts, as well as new (unstudied) items. The test requirement is to use one response option for items from one of the two study contexts (hereafter *targets*), and the other response option for new items, as well as those from the other study context (hereafter *non-targets*).

Comparisons between the left-parietal ERP old/new effects elicited by targets and non-targets have formed the basis for inferences about control over recollection. In a number of studies, it has been shown that the likelihood of recollecting information about targets influences the magnitude of the effect for non-targets: the old/new effect for targets is larger than that for non-targets when the likelihood of recollection is relatively high, but more similar to the effect for non-targets as the likelihood of recollecting information about targets decreases {Dzulkifli, 2006 #1140;Dzulkifli, 2005 #203;Herron, 2003 #1089}. This finding has been interpreted as evidence for the selective control of recollection, in so far as the ERP data indicates that some kinds of recollected content are being prioritised over other kinds {Eimer, 1996 #1291;Fraser, 2007 #1458}.

One explanation offered for this pattern of findings is that using the presence or absence of recollected content about targets to make the binary judgment in an exclusion task (an A/not A strategy) is a good approach when recollection of target content is likely, hence the larger parietal old/new effects for targets than for non-targets under those circumstances. The utility of this strategy, however, diminishes as the likelihood of recollecting target content diminishes, thus, the explanation for the circumstances under which target and non-target old/new effects are comparable is that this reflects a strategy of relying on an assessment of recollected information about non-targets as well as targets when it is beneficial to do so {Herron, 2003 #1089}.

An important development of this account has been offered recently, which is that another determinant of when control over recollection can be exerted is the availability of sufficient cognitive resources. Elward and Wilding {, 2010 #1590} demonstrated that people with increased working memory capacity (WMC) showed greater evidence of selective recollection. That is, they had larger target than non-target ERP old/new effects than people with lower working memory capacity. This outcome did not vary with levels of response accuracy in the exclusion tasks they used.

Elward and Wilding {, 2010 #1590} interpreted WMC as a measure of the availability of cognitive resources, and suggested that, as the likelihood of recollecting information in a task decreases, the demands upon cognitive resources increase. For example, retrieval itself may be more resource demanding if the quality of recovered information is not high, and this in turn might place additional load on processes involved in monitoring the outputs of retrieval. Elward and Wilding {, 2010 #1590} proposed that selective control over recollection would be implemented only when there was sufficient cognitive resource available to do so. Hence individual differences in WMC, and not solely the likelihood of recollecting task-relevant information, determine when control over recollection can be exerted.

In so far as exerting control over recollection is resource-demanding, the findings of Elward and Wilding {, 2010 #1590} might not be regarded as surprising, but there are two related points that are worth making. First, for high WMC participants, response accuracy did not predict when recollection of target content would be prioritised over recollection of non-target content. This raises the possibility that the availability of resource results in a processing style that is not always optimal. Second, WMC is an individual difference variable that is rarely controlled for in functional imaging studies of memory, but the ERP findings comprise marked changes in neural activity despite little evidence of behaviour change. In so far as the findings reported by Elward and Wilding {, 2010 #1590} are not peculiar to the exclusion task, the outcome they obtained raises the possibility that group averaged functional imaging data might not reflect accurately the activity (hence the processes engaged) for all participants contributing to the group average. Moreover, when participants are selected without recourse to WMC scores, it may be that differing WMC profiles across cohorts are the primary driver for differences between measures of neural activity, rather than other task characteristics to which the divergences might be (and most commonly are) assigned.

These possibilities motivated in part the study described here, in a paradigm that is immune to some of the objections that might be raised on the basis of Elward and Wilding’s initial findings and interpretations. Elward and Wilding {, 2010 #1590} employed a version of the exclusion task in which items were encountered in one context at study {Jennings, 1997 #1138}. These were re-presented at test and designated as targets. A proportion of new test items were repeated, and participants were asked to treat the repetitions as non-targets. One potential problem with this design is that the targets and non-targets differ with respect to the time between first and second presentations. This is not typically the case in studies depending upon memory for contextual details (with the exception of studies of temporal context), and it may be that participants relied upon factors such as the relative strengths of memories to make decisions in the test phase. If this was the case, it undermines claims that the data obtained reflect prioritisation of recollection of some contents over others. A related challenge for Elward and Wilding’s account stems from the fact that the version of the exclusion task they employed provides no means for estimating how memorable non-targets actually were. Consequently, the outcomes they obtained might be explained by linking WMC to differential encoding of targets and non-targets, rather than differential exertion of control over what contents were recollected.

This consideration motivated one component of the experiment described below, in which participants completed the more widely used version of the exclusion task where items were first encountered in one of two study contexts {Jacoby, 1991 #397}. In separate test blocks targets were designated as items coming from one or the other study context. This design permits an assessment of how memorable targets and non-targets are. Evidence for selective control over recollection in this task only in those participants with higher WMC scores would suggest that the findings of Elward and Wilding {, 2010 #1590} are not peculiar to one kind of exclusion procedure or a consequence of targets and non-targets not being equally memorable.

In a second component of the study, we investigated the possibility of a causal link between resource availability and ERP evidence for selective recollection. Half of the participants completed a Stroop task for 6.5 minutes immediately before the exclusion task {MacLeod, 1991 #1574;Stroop, 1935 #1573}. Completing the Stroop task has been shown to influence negatively performance on subsequent tasks that are assumed to require cognitive control, such as negative priming {Baumeister, 1998 #1591;Baumeister, 2000 #1592;Muraven, 2000 #1593;Muraven, 1998 #1594}. Importantly, these performance decrements are not evident when demanding and complex tasks that have little or no cognitive control requirements (such as algebra) are completed following Stroop {Baumeister, 2000 #1592}. This outcome has motivated the view that resources important for cognitive control are temporarily unavailable after having been taxed heavily, hence the subsequent performance decrements. If this view is correct, and if changes in the magnitude of the left-parietal ERP old/new effect arise because of the exertion of control over recollection, then depletion of the resource necessary for control will attenuate differences between the magnitudes of left-parietal ERP old/new effects for targets and non-targets in high WMC participants. This outcome would be evidence for a causal link between WMC and the ability to exert cognitive control over recollection.

Finally, one further manipulation was included in the experiment. After completing the exclusion task in which ERPs were acquired, all participants completed a free recall task where they were asked to write down all of the words they could remember encountering in the exclusion task. If changes in the left-parietal old/new effects in the exclusion task are in fact indexing some kind of selective prioritisation of recollected content, this should influence the subsequent memorability of the relevant content. Specifically, if the ERP evidence for selective recollection is a consequence of differential prioritisation of targets and non-targets, then larger target than non-target left-parietal ERP old/new effects should be associated with superior recall of targets.

**2. Methods**

**2.1. Participants:** 57 were recruited through the Cardiff University Experiment Management System and were paid £20 or awarded course credits for their participation. All spoke English as a first language, were right-handed, did not have a diagnosis of dyslexia, were not taking any psychoactive medication, were between 18 and 30 years of age and gave informed consent before taking part. Data from 9 were rejected due to: i) poor response accuracy on the exclusion task (discrimination between targets and non-targets < 0.1; 4 participants), excessive EOG artefact (3) and researcher error (2).

**2.2. Overview of Procedure:** Each participant completed five phases in a fixed order. All phases other than the second were the same for all participants. In the first phase, participants completed an automated operation span (ospan) task, which is a widely employed measure of working memory capacity {Turner, 1989 #1554}. In the second phase, participants completed either a Stroop or a control task. Assignment to Stroop/control was pseudo-random: 24 participants completed each task. In the third phase, all participants completed an exclusion task and ERPs were acquired during the test (retrieval) blocks. The fourth phase was the free recall task, and the final phase comprised completion of a battery of assessments of the integrity of frontal cortical function. This final phase was not described in the Introduction. It was introduced to assess the specificity with which ospan scores are linked with changes in other experimental variables. Given the links between frontal cortical integrity and higher cognitive functions, one possibility is that changes attributed to differences in working memory capacity are parasitic upon frontal cortical function

**2.2.1. Phase 1:** Ospan. Participants completed an automated ospan task {Unsworth, 2005 #1595} (for task, see http://psychology.gatech.edu/renglelab). On each trial, participants first saw a mathematical equation (e.g. [3x2] + 4 = ?). They were instructed to make a key press when they had calculated the solution, at which point a possible solution was displayed along with true/false response options. After the true/false decision, a single letter appeared on the screen. Participants were informed that letters should be remembered for subsequent recall. The number of sequences of equations and letters that were presented before recall (the set size) varied between 3 and 7. A total of three sequences of each set size were presented. The percentage of correct maths solutions was displayed in the upper right-hand corner of the screen and participants were instructed to try and keep this value at 85% or above. A total of 75 maths problems were presented. The ospan score was calculated as the sum of the number of items recalled in perfectly recalled sets. For example, recalling all three sequences with 3 letters, and two with four letters, results in a score of 17 (3+3+3+4+4).

**2.2.2. Phase 2:** Stroop/control task. The Stroop task group completed a Stroop colour naming task. Participants were given 5 A4 cards with 160 colour names printed in capitals on each (an equal number of the words red, green, blue and yellow). The words were arranged in 5 columns, and each colour word was printed 8 times in each column. Two of each 8 presentations were in one of 4 colours (red, green, blue and yellow). This meant that word colour and word meaning were incongruent for 75% of the printed words. Words in each column were ordered pseudo-randomly: no more than 4 words in the same colour were presented consecutively and no columns were identical. Participants were asked to name the colour of the ink that each word was printed in, reading one column at a time, and to ignore word meaning. They were asked to read as many words as possible in the allocated time and to prioritise accuracy over speed. Participants knew that the researcher would record how many colours were named correctly and incorrectly. In the control group, the cards matched those for the Stroop task group with the exception that all words were printed in black ink. Participants were asked to read the words aloud. Participants in both groups were asked to perform the tasks continuously for 6.5 mins, starting again with the first sheet if all 5 had been completed. This time period is the same as that employed in other studies where this between-groups manipulation has been employed before a memory retrieval task {Neshat-Doost, 2008 #1596}.

**2.2.3. Phase 3:** Exclusion Task.Stimuli were three hundred and sixty concrete nouns taken from the Kucera and Francis {, 1967 #484} corpus and presented in white letters on a black background on a computer monitor placed approximately 1 metre from participants. Words had a frequency range of 1-15/million and a letter range of 4-9. Maximum horizontal and vertical visual angles were 4.6º and 0.6 º, respectively. Six groups of 60 words were selected at random for a full experiment list. Each experiment list comprised two study-test cycles. Each study phase comprised two word groups (120 words). These were repeated at test together with a third word group to give 180 test words per cycle. No words were repeated across cycles. Word groups were rotated fully across experiment lists, resulting in the formation of 6 complete lists.

**2.2.3.1. Study:** Participants were fitted with an electrode cap (see below). The researcher read aloud the exclusion task instructions and participants were also given written descriptions. In each study phase, there were two encoding tasks. Cues preceding each word signalled which task to complete; ‘FUNCTION?’ for the function task, ‘DRAW?’ for the drawing task. In the function task, participants were asked to think of a function for the object denoted by each word and make a binary easy/difficult response indicating whether it was easy or difficult to think of an appropriate function for the object. In the drawing task, they were asked how easy it would be to draw the object denoted by each word, again making an easy/difficult response. Cues remained on the screen for 1000 msec, followed by a blank screen for 500 msec. Order of encoding task cues was pseudo-randomised; no more than three consecutive words were preceded by the same cue. Each study word was presented for 300 msec before the screen was blanked. Participants initiated the next trial by pressing a key on a response pad, and the trial started 2000 msec after this response.

**2.2.3.2. Test:** Each trial began with a fixation asterisk which remained on the screen for 500 msec, followed by a 500 msec blanked screen and then the test word for 300 msec. The screen was then blanked until the participant responded. The next trial began 1500 msec later. Participants were instructed to respond using the index finger of one hand to words from one of the two encoding tasks (*targets*), and the index finger of the other hand to new test words as well as those from the other task (*non-targets*). Target designation (function/drawing) changed across study-test cycles. The hands used for responses were balanced across participants, and 50% of participants completed the function task designation first. There was a short break between each study-test cycle and between study and test phases. The researcher reiterated test phase instructions to each participant immediately before each test phase began, at which point target designation (function/drawing) was also indicated. Participants were not informed of the specific requirements of the retrieval tasks at the outset of the experiment. All participants completed one block in which words from the function task were designated as targets and one block where words from the draw task were designated as targets.

**2.2.4. Phase 4:** Free Recall. This task was completed immediately after the exclusion task and before removal of the EEG cap. Participants were given a blank sheet of paper. They were asked to write down all of the words they could remember from any phase of the exclusion task.

**2.2.5. Phase 5:** Prefrontal cortical function tests. These were completed after the EEG cap had been removed. They were: two versions of the Tower of Hanoi (ToH) task (3 and 4 disk versions), the Wisconsin Card Sorting Task {WCST; /Berg, 1948 #1575}, and an assessment of abstract reasoning. Participants in the control condition completed a Stroop task at this stage. Full details of these tasks and their administration are available from the corresponding author on request.

**2.3. Electroencephalogram (EEG) recording:** This was recorded from 32 locations based on the International 10-20 system {Jasper, 1958 #407} including midline (Fz, Cz, Pz, Oz) and left/right hemisphere sites (FP1/FP2, F7/F8, F5/F6, F3/F4, F1/F2, T7/T8, C5/C6, C3/C4, C1/C2, P7/P8, P5/P6, P3/P4, P1/P2, O1/O2). Additional electrodes were placed on the mastoid processes. Electro-ocular activity (EOG) was recorded from above and below the left eye (vertical EOG) and from the outer canthi (horizontal EOG). EEG (range DC-419 Hz; sampling rate 2048 Hz) was acquired referenced to linked electrodes located midway between POz and PO3/PO4, and re-referenced off-line to the average signal at the mastoids. Trials containing large EOG artefact were rejected, as were trials containing A/D saturation or baseline drift exceeding ±80μV. Other EOG blink artefacts were corrected using a linear regression estimate {Semlitsch, 1986 #841}. The data were band-pass filtered off-line (0.03 - 60 Hz), epoched and down-sampled to 167 Hz. Epochs were 1536 msec in length including a 102 msec baseline relative to which all mean amplitudes were computed. A 7-point (22Hz) binomially weighted smoothing filter was applied to the averaged ERPs before analysis.

**3. Results:** Performance on the exclusion task is described first. This is followed by reports of analyses of the ERP data, which is in turn followed by a report of performance on the recall test.

**3.1. Behaviour Measures 1:** Response accuracies and reaction times for each category of stimulus on the exclusion task are presented in Table 1, separated by group (Stroop/control). These data are collapsed across the blocks in which words associated with the function or drawing task were designated as targets, because there were no reliable performance differences with target designation. Two discrimination scores were calculated for each participant {Snodgrass, 1988 #869}. These were: (i) the differences between probabilities of target responses to targets and non-targets, and (ii) the differences between probabilities of target responses to targets and to new test words. Discrimination was reliably above chance for both groups of participants on both measures (minimum t(23) > 10.0, p < .01). These discrimination measures were entered into 2\*2 mixed ANOVA along with the factor of group. There was a main effect of measure only (F(1,44) = 403.80, p < .001), reflecting superior target/new than target/non-target discrimination. In this and all subsequent ANOVAs, the Greenhouse-Geisser correction for non-sphericity was applied when necessary {Greenhouse, 1959 #302}.

PLEASE INSERT TABLE 1

A 3\*2 ANOVA was conducted on the reaction times (RTs) to correct responses for each stimulus category, including group as a between-participant factor. There was a significant main effect of stimulus category only (F(2,88) = 130.00, p < .001) and follow up t-tests (collapsed across task) revealed only that RTs were reliably shorter for new test words than for targets and non-targets (smallest t(47) > 11.00, p < .001).

**3.2. Electrophysiological Measures:** The first analyses were conducted to assess whether the ERPs elicited in this task included a left-parietal ERP old/new effect. Figure 1 shows the scalp distributions of the old/new effects for targets (left-hand side) and for non-targets in the 500-800 msec time period, which is the epoch in which parietal old/new effects are commonly observed. These maps were computed from difference scores obtained by subtracting mean amplitudes associated with correct rejections from those associated with targets and with non-targets, respectively. The maps were generated from data collapsed across group (Stroop/control).

A greater relative positivity for targets and for non-targets in comparison to correct rejections is evident at left-posterior scalp sites. The figure shows that the scalp distributions of the old/new effects for targets and non-targets are very similar, and a direct comparison of the scalp distributions of these two effects using re-scaled data computed over all electrode locations revealed no reliable differences between the two**1**. In a second analysis conducted to assess the correspondences between the target and non-target old/new effects, the amplitudes of the two effects at left- and right-parietal sites P5 and P6 were subjected to a 2\*3 ANOVA with factors of hemisphere and response category (target, non-target, new). As per the analysis of scalp distributions, this analysis was conducted on data collapsed across group. A reliable interaction between hemisphere and category (F(2,94) = 13.55, p < 0.001) was followed up by separate analyses of the target and non-target old/new effects. These revealed that the amplitudes associated with targets and non-targets were reliably larger than those associated with new items at P5 and P6 (minimum t(47) = 2.3, p < 0.05), and that both old/new effects were larger at P5 than at P6 (targets: t(47) = 4.6, p < 0.001; non-targets t(47) = 3.4, p < 0.01). The outcomes of these initial analyses confirm the presence of reliable (and reliably left-lateralised) ERP old/new effects for targets and for non-targets.

**Focused analyses at left parietal scalp sites:**

The remaining analyses were conducted to address directly the hypotheses set out in the Introduction. The first of these was conducted to replicate Elward and Wilding’s (2010) finding of a positive relationship between ospan score and the difference measure obtained by subtracting left-parietal old/new effects for non-targets from targets. These analyses were conducted on data for all control group participants from the P5 electrode, where the old/new effects for targets and for non-targets have been shown to be largest. These data were entered into a regression model with Working Memory Score (as measured by ospan) as the predictor variable and the amplitude difference between targets and non-targets attracting correct judgments as the dependent variable. In keeping with the earlier finding {Elward, 2010 #1590} there was a significant relationship between these two R2 = .20, F (1, 22) = 5.38, p<.05 (see Figure 2).

The next set of analyses was conducted to determine whether the Stroop manipulation influenced control over recollection. To accomplish this, the control and Stroop groups were subjected to a median split based on the WMC scores. Appendix 1 shows the behavioural data on the exclusion task split into the high and low WMC groups. An ANOVA identical to the one reported above for the exclusion data but including high/low WMC revealed no reliable effects involving this factor. Table 2 shows descriptive statistics and scores on the ospan and Stroop tasks, as well as the measures of frontal function, for the high and low WMC participants. In 2\*2 ANOVAs conducted on each of these measures (factors of WMC (high/low) and group (Stroop/control)) the only reliable outcome (see the first row) arose because WMC was reliably greater in high than in low WMC participant groups. The table emphasizes the marked similarities between the WMC scores for the high and low WMC groups when separated for Stroop and for control participants.

Figure3 shows the ERP old/new effects at the P5 location separated by task and group. While positive-going ERP old/new effects for targets are evident in each case, old/new effects for non-targets appear markedly smaller than those for non-targets only for high WMC participants in the control group. The ERP data submitted to analysis were the magnitudes of the target and the non-target old/new effects in the 500-800 msec epoch at electrode P5. These were obtained by subtracting the mean amplitudes associated with correct rejections from those associated with the target and the non-target categories. These difference scores were analysed in a 2\*2\*2 ANOVA with factors of category (target-new/nontarget-new), WMC (high/low) and group (Stroop/control). In keeping with the impression given by the figure, the ANOVA revealed a reliable three-way interaction (F(1,46) = 5.9, p < .05). Follow up t-tests revealed that the target old/new effect was reliably larger than the non-target effect only for high WMC control group participants (t(11) = 2.9, p < .05; in the other three contrasts, largest t(11) = 0.7).

**3.3. Free Recall Measure:** The analyses of these data were guided by the findings in the analyses of the ERP data. If the magnitude of left-parietal old/new effects indicates the extent to which recollection has occurred, and if recollection during the exclusion task has an impact on the subsequent memorability of words that were encountered, then only high WMC participants in the control group should recall more targets than non-targets. Figure 4 shows the numbers of targets, non-targets and new test words recalled during the exclusion task, separated according to WMC and task group. This prediction was supported by the presence a reliable difference between recall of targets and non-targets for this group only (t(11) = 5.9, p < .001; for the remaining contrasts, largest t(11) = 1.3). Two separate one-way ANOVAs for numbers of targets and non-targets recalled across groups revealed no reliable differences (both F < 1).

**4. Discussion**

There are three principal outcomes in this experiment. First, the findings provide an important replication of those reported by Elward and Wilding (2010) in a paradigm which has some advantages over the one they employed. Second, the findings provide new evidence for a causal link between the availability of cognitive resources and the exertion of cognitive control over recollection. Third, they indicate that control over recollection has consequences for how memorable the material that was subject to cognitive control will be subsequently. These three outcomes are discussed in turn.

The data points that are directly comparable to those reported previously by Elward and Wilding (2010) are those from the control condition. Elward and Wilding (2010) showed that the degree to which left-parietal ERP old/new effects elicited by non-targets are attenuated relative to the effects elicited by targets is correlated positively with WMC. This outcome was replicated here, and this is important because of the differences between the designs of the two experiments. Elward and Wilding (2010) had no direct measure of how memorable non-targets were (see Introduction). As a result the differences they observed might have arisen because targets and non-targets were not equally memorable, rather than because participants prioritised recollection of some contents over others. In this experiment target/non-target designation varied across blocks, and performance across this manipulation did not vary. The findings in this experiment are therefore not vulnerable to this alternative interpretation, hence the correlation between WMC and the differences between the magnitudes of the target and non-target parietal old/new effects is a strong argument for the view that the ability to exert control over recollection is related to the resources available to do so.

This claim is bolstered by the absence of ERP evidence for prioritisation of recollection of target content in participants who completed a Stroop task prior to the exclusion task, irrespective of whether WMC scores were high or low. Completion of this task has been shown to impact negatively on performance on some kinds of subsequent tasks, with the explanation offered being that completing the Stroop task results in a time-limited depletion of the cognitive resources that can assist performance on the subsequent task {e.g. /Baumeister, 2000 #1592}. The result of this manipulation in this experiment was to attenuate differences between the amplitudes of target and non-target old/new effects for high WMC participants, rendering them electrophysiologically indistinguishable from low WMC participants. The outcomes therefore suggest a causal link between the availability of cognitive resource and the exertion of cognitive control over recollection.

The third key outcome in this experiment was differences in performance on the free-recall task all participants completed after the exclusion task. Only high WMC control participants recalled reliably more targets than non-targets, and these participants were the only ones for whom there was electrophysiological evidence for prioritisation of targets over non-targets. This recall pattern provides a complementary source of evidence to support the claim that differences between the magnitudes of target and non-target left-parietal ERP old/new effects indicate that the recollection of the former occurred to a greater degree (or for a higher proportion of stimuli) than for the latter. By this account, the extent to which targets and non-targets were associated with processes tied to recollection on the exclusion task influenced the likelihood that they would be recalled subsequently.

One question that the recall data prompts is whether the differences between target and non-target recall in the high WMC controls come about because of ways in which targets or non-targets were processed. Unfortunately, the behavioural data on the recall task do not shed light on this issue, because across groups the likelihoods of recalling targets did not differ reliably, and neither did the likelihoods of recalling non-targets. Figure 3 shows that the principal divergence between the ERPs across groups comprises a markedly larger target old/new effect: the non-target old/new effects are comparable across groups. While this might be considered evidence in support of the view that the principal difference across groups is in the processing of targets. Elward and Wilding (2010) observed that inferring from this outcome that control over recollection arose because of selective processing of target content depended on the assumption that WMC had no overall influence on the magnitude of left-parietal ERP old/new effects, arguing that, if it did (for example, if WMC was correlated with the magnitude of the left-parietal ERP old/new effect), then drawing inferences based on the similar magnitudes of the non-target ERP old/new effects would not be appropriate.

It may be the case that the relatively small group sizes in this experiment preclude extraction of differences in recall performance across groups that might help to understand the mechanisms that underlie control over recollection. Irrespective of the accuracy of either of the foregoing accounts, however, the ERP data clearly indicate that prioritisation of some recollected contents over others occurred only for the high WMC control group, and that this had consequences for the relative memorability of targets and non-targets subsequently. In addition, the view that these outcomes are a consequence of resource availability, rather than being parasitic upon a general deficit in one or more higher cognitive functions, is supported by a failure to find a difference between the four groups and several measures of fronto-cortical integrity (see Appendix 2).

These findings also prompt several further considerations. In this experiment, individual differences in WMC resulted in marked changes in neural signatures of memory processes but no discernible changes in response accuracy on the task where the neural measures were obtained {for a related ERP finding, in which reaction times but not accuracy predicted changes in ERP old/new effects, see /Wilding, 2004 #1029}. This outcome might be regarded as a cautionary note in functional imaging studies where the task requirements include adjudicating between two or more study contexts: if variations in WMC enable different retrieval strategies, then aggregate data in which WMC is not accounted for may not generalise. A related observation is that discrepancies across functional imaging experiments, such as failures to replicate, might be explained by different distributions in the WMC scores of the participant population. Likewise, in other tasks where cognitive control over retrieval is assumed to be exerted, such as the Think/No-Think paradigm {Anderson, 2001 #1577} and the retrieval-induced forgetting paradigm {Anderson, 1995 #1338}, variations in WMC might influence performance, and null results might arise because of inadvertent sampling biases (because WMC is not controlled for) across otherwise similar studies.

These practical observations aside, the fact remains that the ERP data obtained in the exclusion task (in combination with the subsequent recall data) have been interpreted here as evidence for differences in the exertion of cognitive control, but neither response accuracy nor response times predicted the key ERP differences. One possibility is that differences in performance at the levels of difficulty documented here are hard to detect, and that they might be revealed on an exclusion task that was more demanding. A related possibility is that the specific set of task demands on the exclusion task, irrespective of difficulty as indicated by response accuracies or reaction times, are insufficient to elicit substantive differences in behaviour. Of relevance here is the finding that recall of autobiographical details is correlated with WMC, and completion of a Stroop task prior to autobiographical recall reduces the number of details that are recovered subsequently {Neshat-Doost, 2008 #1596}. This outcome suggests that there are circumstances under which the accuracy of responding on retrieval tasks is linked with the availability of cognitive resources as indexed by WMC.

What mechanisms might be responsible for differential prioritisation of content associated with targets and with non-targets on the exclusion task? Dzulkifli and Wilding {, 2005 #203} discussed three possibilities, motivated by earlier considerations about how selective retrieval could come about {Anderson, 1994 #25;Bjork, 1989 #75}. Two of these can be conceived of as ‘front-end’ cognitive control mechanisms {Halamish, 2012 #1597} of which the most discussed is cue-specification (or cue-bias). This process ensures that internal representations of retrieval cues are modified such that they are more likely to interact with some kinds of memory representations than others. Hence larger old/new effects for targets than non-targets might be a consequence of greater cue-specificity for high WMC participants.

A second possibility is target bias, where processes are engaged (possibly tonically) that influence the relative accessibility of different kinds of representations. It is possible that larger old/new effects for targets than for non-targets result from either a reduction in the accessibility of representations for non-targets, greater accessibility for target representations, or some combination of the two.

Anderson and Bjork {, 1994 #25} discussed a third ‘back-end’ cognitive control process, whereby selective retrieval is a consequence of attending to only task-relevant products of retrieval. Each of these classes of process can explain the data reported here, and in each case it is also possible to link failures to engage these processes to WMC. Each trial of the exclusion task places demands upon retrieval, decision-making, response preparation, selection and execution processes. It may be that WMC estimates are an indicator of the extent to which, in addition to meeting these demands, participants are also able to adopt a content-specific task completion strategy. It is also important to note that this description is not intended to imply that adoption of a content-specific strategy is all-or-none. It may be the case, for example, that sufficient WMC enables an individual to apply effective cue-specification processes consistently throughout an experiment session. Individuals with lower levels of resource might engage effective cue-specification processes either sporadically, or consistently but at a lower level of specificity. In the absence of the ability to assess relevant neural markers on a trial-by-trial basis it is not straightforward to adjudicate between these competing accounts. Similar arguments can be made for target-bias and attention-based accounts of selective retrieval, and of course there is no reason to assume that these putative classes of selective retrieval process are mutually exclusive.

A somewhat different explanation for the ERP data obtained in this experiment stems from investigations of the role of the inferior parietal lobe in memory and attention. O’Connor , Han and Dobbins {, 2010 #1598} have linked activity in this region to expectancy violation. They based this proposal on findings in a functional magnetic resonance imaging (fMRI) study in which participants completed a recognition memory task and were given information at the start of each trial about whether the immediately following stimulus was likely to be studied or unstudied. Valid information resulted in greater activity in the inferior parietal lobe than invalid information, and critically this was the case for studied as well as unstudied stimuli.

The left inferior parietal cortex has many of the same functional properties as the left-parietal ERP old/new effect, and it has been argued that it is the likely neural generator of the effect that is seen at the scalp {e.g. /Vilberg, 2008 #1492}. Consequently, one possibility is that the results reported here are linked to violations of stimulus expectancies that differ according to WMC. One way in which this might occur follows from the proposal that high WMC participants are more likely to adopt an A/not A strategy for the task than are lower WMC participants. In this task there was an equal number of targets, non-targets and new test words, so reducing the decision to a binary category judgment (A/not A or target versus ‘other’) results in a smaller proportion of targets than ‘others’. If this in turn results in the expectancy that words other than targets are the ones most likely to be encountered, then this provides a means for the larger old/new effect for targets than for non-targets to be interpreted as an expectancy violation. By this view, low WMC participants are less likely to reduce the task to a binary category judgment hence expectancy violations are less likely to occur. This account still maintains that the changes in the ERPs reported here reflect the implementation of cognitive control operations, but it does not require that control influences the contents that are recovered from memory.

One immediate starting point for investigating this possibility further is to use the pre-cuing task reported by O’Connor and colleagues {, 2010 #1598} with ERPs rather than fMRI as the neural correlate. There are also good reasons to consider employing the same kinds of task adopted in this experiment along with fMRI measures. If the ERP effects reported here arise because of upstream differences between the amount of content recovered for targets and non-targets, the target/non-target distinction in the parietal lobe might be honoured in other brain regions, such as the hippocampus**2**. This would not invalidate an expectancy violation account, but it would indicate that the different strategies that lead to changed expectancies have a direct influence on what contents are in fact recovered from memory.

**4.1 Concluding Remarks:** It is uncontroversial to assume that some kind of control over what we recollect, and when we recollect, can be exerted. How this comes about remains poorly understood, but the links and outcomes documented here emphasise the utility of employing real-time measures of neural activity, alongside behavioural assays, to understand control over recollection. Important goals for subsequent studies are assessments of whether ERPs point to the use of different retrieval strategies on tasks other than the exclusion procedure, and careful consideration of how electrophysiological and behavioural measures can be combined to establish the mechanisms by which control over recollection is exerted.

**References:**

**Footnotes**

1. Prior to the analysis of scalp distributions, the difference scores (target-new, non-target-new) were rescaled using the max-min method, to avoid confounding changes in the shapes of distributions with the magnitudes of the effects of interest {McCarthy, 1985 #548;Urbach, 2002 #956;Wilding, 2006 #1014}.
2. Isolating hippocampal activity via neural recordings at the scalp is not straightforward.

**Table 1.** Probabilities of correct responses and reaction times for each class of test word on the exclusion task separated according to group (Stroop/control). Standard deviations are in parentheses.

 Control Stroop

Accuracy

Target 0.74 (0.11) 0.74 (0.13)

 Non-target 0.82 (0.12) 0.80 (0.10)

New 0.97 (0.04) 0.93 (0.08)

Reaction Times

Target 898 (237) 972 (272)

 Non-target 948 (258) 1011 (263)

 New 695 (228) 710 (237)

**Figure Legends**

**Figure 1.** Scalp distributions of the target and non-target ERP old/new effects averaged over the 500 and 800 msec post-stimulus time period. The maps were interpolated from difference scores obtained by subtracting mean amplitudes associated with correct rejections from those associated with targets and non-targets attracting correct test judgments. The values below the maps denote the voltage minima and maxima within each map, and can be interpreted relative to the colour bar on the right-hand side of the figure.

**Figure 2.** Mean amplitude of the differences between targets and non-targets at electrode P5 plotted against working memory capacity (WMC) scores for control participants.

**Figure 3.** Grand average ERPs from the P5 left-parietal electrode location associated with correct judgments to targets, non-targets and new test words. The data (N = 12/group) are separated according to WMC (High/Low) and task group (control/Stroop).

**Figure 4.** Mean numbers of words recalled by participants in each group and separated according to the status of the words in the exclusion task (target, non-target, new). Standard deviations are in parentheses.

 Target – New Non-target - New

Target hits minus non-target hits (mv)

Working Memory Score

600ms

0ms

10µv

+

High WMC

Low WMC

Control

Stroop

**Appendix 1.** Probabilities of correct responses and reaction times for each class of test word on the exclusion task separated according to WMC and group (Stroop/control). Standard deviations are in parentheses.

 Control Stroop

 High WMC Low WMC High WMC Low WMC

Accuracy

Target 0.71 (0.13) 0.77 (0.08) 0.76 (0.08) 0.72 (0.17)

Non-target 0.82 (0.14) 0.83 (0.10) 0.80 (0.11) 0.80 (0.09)

New 0.97 (0.02) 0.96 (0.07) 0.93 (0.08) 0.93 (0.09)

Reaction Times

Target 892 (207) 903 (268) 1019 (231) 925 (314)

Non-target 959 (252) 937 (263) 1065 (225) 956 (301)

New 744 (274) 657 (183) 779 (266) 652 (208)

**Appendix 2.** Participant performance on experiment tasks and assessments of frontal function. Ospan: scores on the operation span task. ToH4 and ToH5: Tower of Hanoi task with 4 and 5 discs, respectively, scores are calculated as (minimum number of moves required / total moves played) x 100. WCST: the percentage of perservative errors on the Wisconsin Card Sorting Task. Reasoning: the percentage of correct responses on the abstract reasoning test. Standard deviations are in parentheses. F-values (right-hand columns) show the outcomes of 2\*2 ANOVAS for each measure with factors of WMC and task. \* = p < .001.

Control Stroop F-Values

 High WMC Low WMC High WMC Low WMC WMC Task WMC\*Task

Ospan 53 (7.9) 28 (8.8) 52 (8.3) 27 (4.6) 129.7\* 0.1 1.0

Stroop score 530 (122) 461 (151) 464 (82) 497 (146) 0.2 0.2 1.9

Stroop errors 4.4 (2.4) 6.3 (3.8) 5.5 (3.0) 6.4 (7.7) 1.1 0.2 0.1

ToH4 81% (20%) 67% (20%) 66% (24%) 58% (26%) 2.5 2.8 0.2

ToH5 72% (28%) 64% (21%) 72% (29%) 68% (20%) 0.8 0.5 0.9

WCST 10% (5%) 9% (4%) 8% (3%) 9% (3%) 0.4 0.9 0.5

Reasoning 50% (19%) 51% (15%) 47% (15%) 49% (16%) 0.5 0.2 0.1