Flexible attention allocation to visual and auditory working memory tasks: manipulating reward induces a trade-off

Candice Coker Morey · Nelson Cowan ·
Richard D. Morey · Jeffery N. Rouder

Abstract Prominent roles for general attention resources are posited in many models of working memory, but the manner in which these can be allocated differs between models or is not sufficiently specified. We varied the payoffs for correct responses in two temporally-overlapping recognition tasks, a visual array comparison task and a tone sequence comparison task. In the critical conditions, an increase in reward for one task corresponded to a decrease in reward for the concurrent task, but memory load remained constant. Our results show patterns of interference consistent with a trade-off between the tasks, suggesting that a shared resource can be flexibly divided, rather than only fully allotted to either of the tasks. Our findings support a role for a domain-general resource in models of working memory, and furthermore suggest that this resource is flexibly divisible.

Keywords working memory · attention · change detection · rewards

Introduction

Working memory is a complex system that holds information while it is temporarily available to be processed further and manipulated (Baddeley, 2007; Baddeley & Hitch, 1974; Cowan, 1988, 2005; Miller, Galanter, & Pribram, 1960; Miyake & Shah, 1999). A key issue is whether information in working memory is held entirely in separate modules that do not affect one another (e.g., Baddeley, 1986; Baddeley & Logie, 1999), or whether it is held at least partly in a common faculty in which information from various sources share limited resources (e.g., Baddeley, 2001; Cowan, 2001, 2005). In order to determine which view is correct, one must explore the nature of interference between concurrent working memory loads imposed in different domains, such as the visual-spatial and acoustic-verbal domains.

Such interference is often observed, but there are still controversies concerning the nature of that interference. One issue has been whether cross-domain interference can be obtained at all. At least some experiments have shown such interference (Morey & Cowan, 2004, 2005; Saults & Cowan, 2007; Stevanovski & Jolicoeur, 2007; Vergauwe, Barrouillet, & Camos, 2010), though others have shown little or no such interference (Allport, Antonis, & Reynolds, 1972; Cocchini, Logie, Della Sala, MacPherson, & Baddeley, 2002; Logie, Zucco, & Baddeley, 1990). We will reexamine this issue in a manner complementary to previous studies.

A second issue is whether the individual can choose the proportions of working memory resources to allocate to the two tasks, or whether that proportion is immutable. If attention is involved in working memory storage and maintenance, the allocation of that resource should be at least partly voluntary. That is true regardless of the particular model of working memory. If one accepts a modular conception of working memory (e.g., Baddeley, 1986), attention governing the central executive could regulate how many stimuli held in each domain-specific buffer are rehearsed or refreshed; at least the initiation of a rehearsal cycle seems to require some attention

C. C. Morey (✉) · R. D. Morey
Experimental Psychology, University of Groningen,
Grote Kruisstraat 2/1,
9712 TS Groningen, The Netherlands
e-mail: c.c.morey@rug.nl

N. Cowan · J. N. Rouder
University of Missouri,
Columbia, MO, USA

(Naveh-Benjamin & Jonides, 1984). If one accepts a less modular conception of working memory (e.g., Cowan, 1988, 2005), attention might determine how many stimuli from each domain are represented in a common, central store.

In contrast to both of these approaches, though, some varieties of a modular approach seem to exclude a role of domain-general attention (e.g., Cocchini et al., 2002; Wickens, 1984, 2002). According to such approaches, the amount of interference between working memory tasks might depend solely on the amount of overlap between the stimuli to be remembered in the two tasks; the primary source of interference is the competition between similar stimuli for access to their appropriate storage module. First, we aim to confirm that some cross-domain interference indeed occurs, using motivational incentives as an operationalization of volitional attention. Second, assuming some degree of interference is observed, we shall have new information to restrict the plausible features of any shared resource, whether that resource is considered a process manager (like the central executive) or a shared store (like the focus of attention or episodic buffer). Currently, the descriptions of these resources are so vague that it is not necessarily clear how they might be shared between two competing tasks. Is a general working memory resource shared only in the sense that auditory-verbal or visual-spatial tasks may each use it, but without sharing it concurrently? Alternatively, perhaps a shared resource may be arbitrarily divided and deployed to assist in the maintenance of two kinds of memoranda. If so, how fine can this division be? The results we report provide new limitations to be applied to existing models of domain-general resources in working memory.

The present study

Theoretically, there are at least two methods to examine cross-domain interference between two working memory tasks. In the first method, the difficulty of at least one of the tasks is manipulated and evidence for an effect on performance of the other task is examined. This is the method used by all the dual-task studies noted above, and has certainly proved a fruitful strategy for research. Here, we instead use a second method for examining dual-task performance that is well regarded but not often used: to manipulate the relative payoffs or attention allocation instructions of the two tasks instead of task difficulty (Gopher & Donchin, 1986; Navon & Gopher, 1979; Sperling & Dosher, 1986; for more recent applications, see Alvarez, Horowitz, Arsenio, DiMase, & Wolfe, 2005; Craik, Govoni, Naveh-Benjamin, & Anderson, 1996). The logic of these two approaches is similar; as either the difficulty of, or incentive to, one task increases, perfor-

mance on the other task should decrease if a shared resource is needed to carry out both tasks at once. In the trade-offs approach, if a common attentional resource is needed in both tasks, then emphasizing one task should result in an improvement in that task at the expense of performance on the other task.

Evidence from difficulty manipulations has been mixed, with some researchers observing cross-domain dual-task costs but many observing few or no costs. However, the pattern of performance observed in dual-task studies manipulating difficulty can be difficult to interpret clearly. Where cross-domain dual tasks costs have been observed, increasing difficulty of Task A in one domain results in decreased performance on both Task A and concurrent Task B in some other domain. In these cases, the cost to Task B must reflect some cross-domain resource sharing, but the cost to Task A might reflect either domain-specific resource limitations, domain-general limitations, or some combination of these. Expected trade-offs in a scenario in which reward is manipulated instead of difficulty are more straightforward in some respects. First, because the paradigm encourages selective resource-sharing, one expects performance on one task to improve at the expense of the other, rather than performance on both to decrease. Second, any change to either task with changing reward level must be attributed to whatever resource they share. The extent to which trade-offs are observed under this circumstance might reflect the extent to which individuals can determine which task to prioritize, or to selectively attend.

Manipulating payoffs is therefore a method whose success depends on the availability of a reasonable estimate of storage capacity. Such estimates have been advocated by Pashler (1988), as well as by Cowan (2001) and Cowan et al. (2005), and variations of this estimate have been used successfully in a profusion of studies of working memory based on the visual comparison procedure of Phillips (1974), reintroduced by Luck and Vogel (1997; e.g., Alvarez & Cavanaugh, 2004; Gold, Wilk, McMahon, Buchanan, & Luck, 2003; Rouder et al., 2008; Todd & Marois, 2004, 2005; Xu & Chun, 2006; also with a sequential visual memory task, Kumar & Jiang, 2005). One benefit of a capacity metric is that it leads to a clear expectation for the form of a trade-off function, whereas in many situations, this function must be determined empirically (Alvarez et al., 2005). With a capacity measure for each task, the trade-off should be linear if the shared resource has a constant capacity; when the payoffs change, an increase in X units in capacity for Task 1 should always produce a decrease in $C \times X$ units for capacity in Task 2. This expectation follows from the simple physical metaphor in which a maintaining an item takes up a certain proportion of the shared resources.

In certain circumstances, though, $C \neq 1$. This is the case if the units differ between tasks or the tasks differ in the amount of domain-specific mnemonic capacity. Importantly, in our design, estimates of storage capacity are meant as dependent variables, calculated to provide comparable measures between two different tasks, and it need not be assumed that $C = 1$. Certainly, a trade-off between tasks, which in this case would manifest itself as a reduction in capacity estimates of the less rewarded task and a corresponding increase in capacity estimates of the more rewarded task, could arise due to dependence on any kind of shared general resource, whether that resource is thought to function as a memory store (e.g., Baddeley 2001; Cowan's focus of attention, 2005) or as mnemonic processing (e.g., Baddeley's central executive, 1986). In either case, estimates of storage capacity would change with the proportion of reward given for each task, albeit for different reasons. Observing any trade-off would further confirm sharing of some domain-general resource, as in the working memory models of Baddeley and Cowan, in opposition to the notion that mental resources are predominantly separate on several dimensions (Wickens, 2002). Furthermore, the characteristics of any observed trade-off might improve the specificity with which a domain-general resource can be described, which is essential for making further progress in understanding any relationship between attention and memory.

We used a manipulation of financial payoffs in three dual-task working memory experiments. In different conditions of each experiment, different allocation instructions were used while the number of memoranda in each task was held constant. The concurrent tasks, tone sequence comparison and visual array comparison, were the same across all conditions within each experiment, and challenging levels of difficulty were chosen so that the tasks were unlikely to be accomplished using only automatically-activated memory buffers. In designing these tasks, we endeavored to create two tasks that were as equivalent as possible except with respect to the stimulus domain of the memoranda, so that capacity could be estimated in the same manner for both tasks. The critical difference between conditions in each experiment was the level of reward assigned to correct responses in each task. In most conditions, the total potential reward was fixed, and what distinguished the conditions was how the reward was divided between the two tasks. An exception was two conditions in Experiment 2 that were included to ensure that attention allocation, rather than total effort expended, explained performance differences between conditions.

By using multiple payoff conditions we can also determine how fine-grained the attention allocation process can be. For example, it might be that participants have only two possible attentional states: attend to a task or ignore it. Alternatively, participants might be able to split attention between the two tasks. If participants can allocate some proportion of attention to both tasks, how flexible can this allocation be? Note that this flexibility might occur in terms of either the splitting of attention on an individual trial, or in terms of some proportion of trials with attention to each task; we cannot distinguish between those possibilities, as indeed no other prior study has been able to do.

We report three similar experiments, all contributing to a comprehensive Bayesian analysis. In these studies, we included various combinations of reward levels but, overall, we wanted to assess whether reward levels affect performance, and if so, how flexible this effect might be. Using traditional inference techniques, this question might be addressed with ANOVAs, in which we test for any effect of the reward variable, and subsequent post-hoc tests, by which we test (assuming a main effect of the reward variable) how finely a shared resource may be divided by comparing each level of reward. A simple main effect of the reward variable would indicate that some shared resource can be split between the concurrent tasks. Differences between precise levels of reward in the expected direction (i.e., no reward < low reward < high reward < full reward) could be taken as evidence that fine, rather than all-or-none, allocation of this resource is possible. However, an enormous sample might be required to detect significant differences between each level of reward and, moreover, a flexible division of resources does not depend on observing significant mean differences between each reward level. Rather, the joint orderings of reward level conditions in each task might be evaluated. With a Bayesian analysis, the question of interest can be addressed more powerfully and more directly. Specifically, we constructed a nested series of comparisons of theoretically possible joint orderings, supposing different levels of divisibility in a shared attention resource. For each comparison, we calculated a Bayes factor, and compared these Bayes factors to evaluate the evidence for each theoretical level of divisible attention. Because we observed similar effects of reward across all three experiments, it is best to combine the results of all three of our studies into one comprehensive analysis, which is possible and advantageous when using hierarchical estimation techniques. We therefore present the method for all experiments together so that, in turn, we can consider the results of all experiments together.

General method

Because we wanted to compare cross-modal performance on comparable visual-spatial and auditory tasks, we combined a

tone-sequence comparison task (as in Cowan et al., 2005) with a visual array comparison task (cf. Luck & Vogel, 1997; Phillips, 1974). Both tasks were two-alternative forced-choice recognition tasks. In both tasks, stimuli were selected randomly on each trial from comparably-sized samples. The tone stimuli might not be continuously rehearsed with the passive phonological loop as a verbal list might be, which is important inasmuch as we aimed to measure dual-task interference in a central resource. To further ensure that such rehearsal was not effective for either type of stimulus, participants engaged in articulatory suppression by repeating the word "the" throughout each of the three experiments, a standard precaution taken in many visual change detection experiments.

In our procedure, a sample visual array and a sample tone sequence were presented one after another, in either order. Then, a test array and a test tone sequence were presented, with a response required for each task. Participants indicated whether the two arrays were identical or differed in the color of one square and whether the two tone sequences were identical or differed in the pitch of one tone. This procedure is illustrated in Fig. 1, for one of the two orders of the stimuli that were used. The order of presentation of the sets of stimuli to be remembered always matched the order of the tests.

In order to make the visual array and tone sequence recognition tasks as equivalent as possible, we re-presented

an entire array and sequence at test. The task was to indicate whether one item had changed. This whole-array procedure is commonly employed (cf. Cowan et al., 2005; Luck & Vogel, 1997; Saults & Cowan, 2007; Vogel, McCullough, & Machizawa, 2005), although it is also common to employ probes with only one item present or with only one item marked as possibly having changed (for a comparison of whole- and partial-report tests, see Wheeler & Treisman, 2002). We chose to use whole-report probes to allow us to probe memory for the assignment of items to spatial or temporal positions; this was necessary because probing the simple presence versus absence of a single item leads to poor results for tone sequences, which rely on relational information. We therefore used the whole-report method for both tasks, to ensure that the tasks were as similar as possible except for the stimulus domain of the memoranda. For this variation of change detection task, Pashler (1988) capacity estimate is the appropriate metric, because it was constructed for the procedure in which the entire array is reproduced at test with a possible change in one item, whereas Cowan (2001) estimate was constructed for a single-item probe procedure. Both estimates combine hits and false alarms in a principled manner. The difference in assumptions between Pashler's estimate and Cowan's is reflected after dropping redundant terms by dividing Cowan's estimate by the correct rejection rate.
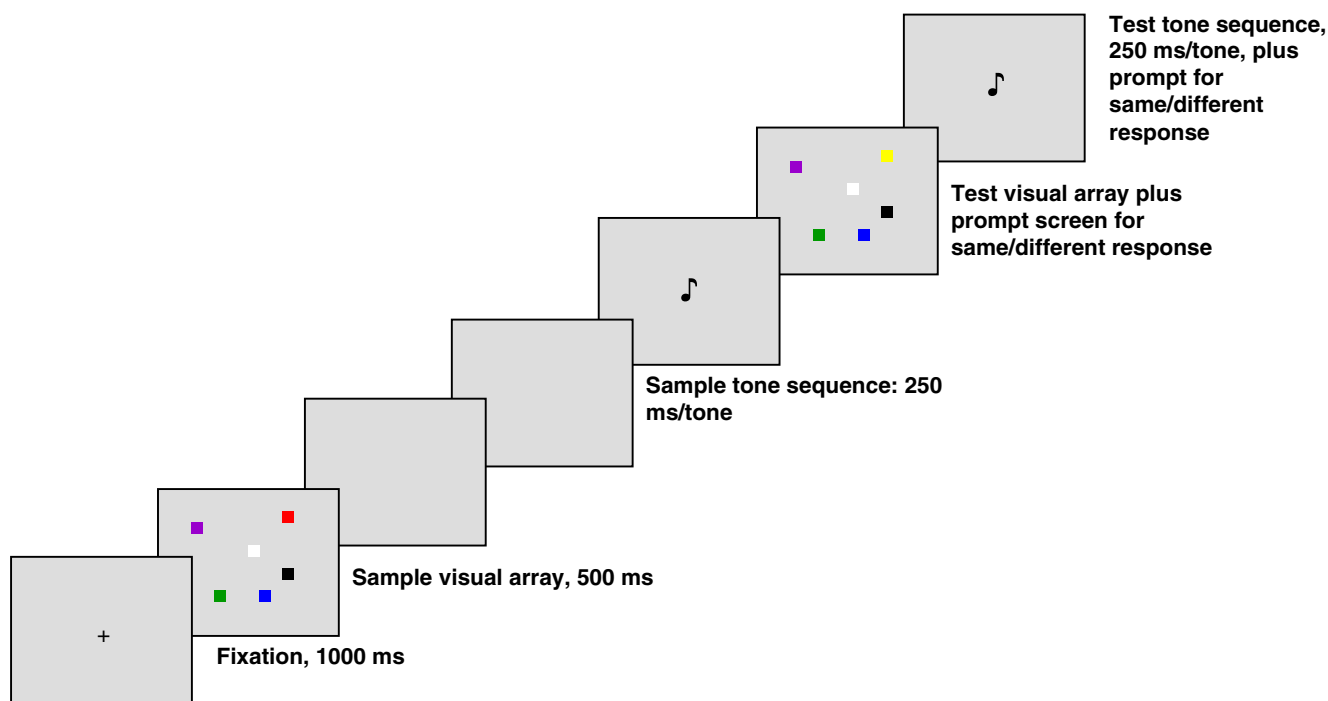

Fig. 1 A graphic depiction of the basic procedure for the visual-array-first presentation order in all experiments. An alternative order in which the tone sequence was presented before the visual array at both study and test was also used. The correct visual response in this case is "different". Feedback was given for both tasks after response to the second task was registered

Experiment 1

Participants

Participants, recruited from the University of Missouri introductory psychology pool, received partial course credit for taking part in the study and a monetary reward of up to $10 for correct responses. Participants earned an average of $8.24 (range: $6.61–9.87). Three were excluded due to chance accuracy in at least one condition for each task and three others were excluded due to computer failure errors, leaving a final $n = 32$ (17 men, 15 women). Each was randomly assigned to one of the two stimulus orders shown in Fig. 1.

Apparatus and stimuli

The experiment was controlled with E-Prime (Schneider, Eschmann, & Zuccolotto, 2002). We chose to use arrays of 8 squares and sequences of 6 tones in an effort to equate task accuracies on the basis of pilot results. Visual stimuli were displayed on a 17-inch (43.18 cm) monitor set to a resolution of 1024 × 768 pixels. Each visual array included eight 20 × 20-pixel squares scattered at randomly chosen locations at least 2° apart within a 270 × 201-pixel area in the center of the screen. The centermost point was excluded as a possible location. Assuming a viewing distance of 50 cm, each square occupied approximately 2° of visual angle. Each square color was selected randomly with replacement from a set of seven easily discriminable colors including red, blue, green, yellow, black, white, and violet, which appeared against a neutral gray background. The first array (*sample*) and the second array (*test*) were either identical or differed in the color of only one square. Tone stimuli were presented via headphones at approximately 78 dB. Each tone sequence contained six tones played at a rate of four per second. Tones were drawn randomly without replacement from a set of nine pitches (87, 174, 266, 348, 529, 696, 788, 880, and 972 Hz). On trials in which one tone changed, the change could be to a pitch that was new to the sequence or a repeat of a tone that occurred at another position in the sequence. Since these pitches did not comprise a typical Western musical scale and the order was randomly determined, the sequences did not sound melodic.

Procedure

After completing eight supervised practice trials, each participant performed five randomly-ordered blocks of experimental trials. In each 40-trial block, values were assigned to accuracy on the visual array and tone discriminations such that the total value of correct responses on any single trial equaled 1,000 points. The allocation of

points to each task was always explicit. Correct responses to each array were worth 1,000, 750, 500, 250 or 0 points each in different blocks, with the reward for the tones set to 1,000 minus the reward for the arrays. Pilot testing of this procedure indicated that participants were more motivated when each trial was worth a large number of points, vague with respect to the actual amount of money each response was worth, rather than a description of the small amount of real money each correct response was worth (rather like the reward structure for playing arcade games). Participants knew that they could earn money by making accurate responses, and were specifically instructed that points corresponded to money and that by accumulating as many points as they could, they would earn the most pay.

Participants viewed two visual arrays and were asked to judge whether the arrays were the same or different. If the arrays differed, only the color of one square changed. Likewise, participants heard two tone sequences and were asked to judge whether the sequences were the same or different. If the sequences differed, they differed only in the frequency of one tone. Half of all the test arrays and sequences differed from the sample stimuli. The visual and tone stimuli were independent; change in one type of stimulus did not increase or decrease the likelihood of change in the other stimulus. Participants were instructed to repeat the word "the" softly (2 repetitions/s) during each trial in order to suppress verbal rehearsal of the stimuli, and an experimenter monitored their suppression throughout the session. Articulatory suppression began when the fixation cross appeared, and continued at least until the first response prompt appeared; participants were not prevented from suppressing throughout the trial if they preferred, as many did. No participant included in the analysis needed to be reminded to suppress articulation more than once after the practice session ended.

The trial events presented in Fig. 1 were similar for all subsequent experiments (which differed slightly in timing, assignment of orders to individuals, and levels of reward conditions). The critical segments of a trial included presentation of a sample stimulus in Task 1 (visual array or tone sequence), sample stimulus in Task 2, test stimulus and query display in Task 1 (requiring a same–different response), and test stimulus and query display in Task 2 (requiring another same–different response). At the end of each trial, feedback was provided for both tasks.

Because a response occurred during either the array or tone inter-stimulus interval depending on the task order, it was not possible to perfectly equate the duration of each task, but it was possible to make them similar. For the visual array comparison task, the onsets of the sample and test arrays were separated by 2,100 ms when the array task began first, and 2,100 ms plus response time to the test tone sequence when the array task began second. For the tone

sequence comparison task, the onsets of the first tone in each sequence were separated by 3,100 ms when the tone task began first and 3,700 ms plus response time to the test array when the tone task began second. (In Experiment 3, when task order was manipulated within-participants, these timings were made as similar as possible.)

Experiment 2

This experiment differed from the first in the inclusion of control trials intended to test whether it could be the absolute, rather than relative, reward assigned to each task that mattered.

Participants

Nine men and 23 women participated in Experiment 2, for partial course credit and up to $10, depending on performance. Monetary reward ranged from $7.11 to $9.41, with a mean reward of $8.27. The data of two participants were removed from the analysis due to chance performance on at least one task, leaving a final $n = 30$.

Apparatus and stimuli

All equipment used in Experiment 2 was the same as that used in Experiment 1. Visual array and tone sequence stimuli were constructed as in Experiment 1.

Procedure

Experiment 2 included two conditions in which reward for correct completion of each task was equal instead of relative. In the low reward condition, participants earned 250 points for each correct response to either the visual-array or the tone-sequence task. In the high reward condition, participants earned 750 points for each correct response to either task. We also included four relative reward conditions, in which each task was worth 0, 250, 750, or 1,000 points, with the total possible reward on each trial equal to 1,000 points. Each combination of reward levels was presented in randomly-ordered blocks of 40 trials each. All other aspects of Experiment 2, including the between-subject order manipulation, and selection and duration of stimuli remained as they were in Experiment 1.

Experiment 3

This experiment differed from the others in that the order of visual and auditory memoranda within a trial was varied within participants.

Participants

After excluding 2 subjects for chance performance in at least one of the tasks, Experiment 3 included 16 men and 8 women, a total $n = 24$. Reward for a 1-h experimental session ranged from $6.54 to $9.21, with a mean of $7.88.

Apparatus and Stimuli

Equipment and tasks used in Experiment 3 were the same as that used in Experiments 1 and 2.

Procedure

Participants completed one set of trials in which the sample array was presented first, followed by the sample tone sequence, test visual array, and test tone sequence, and one set in which the sample tone sequence occurred first, followed by the sample visual array, test tone sequence, and test visual array. Order of these blocks was counterbalanced.

Within each block, relative reward conditions were presented in randomly-ordered blocks. The sum of potential rewards for each trial always equaled 1,000 points, and each task could be worth 0, 500, or 1,000 points, making three reward conditions in each task-order block, or six blocks total per experimental session. Participants first completed an unpaid practice session of eight trials. After the first three blocks of trials, participants were required to take a break lasting at least 1 min. After this break, participants completed eight more unpaid practice trials before continuing with the final three blocks of the session. Each experimental session included 192 paid trials, amounting to 32 trials per reward condition block.

Regardless of task order, the interval between the offset of the first task's stimuli and the appearance of the test stimuli for the first task was always 3,250 ms. This interval for the second task varied depending on participant's response to the first stimulus but for each case measured 3,250 ms plus response time to the other task's test stimulus.

Results

Throughout this paper, inferential results will be reported in terms of capacity estimates rather than proportions correct. These estimates allow a concrete understanding of how many visual items cost how many auditory items in performance as the attention allocation changes, as explained in the introduction.

We calculated capacity estimates using Pashler (1988) formula, which is more appropriate for use with whole-report

probe designs than that of Cowan (2001; also Cowan et al., 2005). If a change occurred, only one item changed, but no cue was given to limit the decision to one particular square in the test array or one tone in the test sequence. That is the situation motivating Pashler's estimate. The relevant formula is

$$k = S(h - f)/(1 - f) \qquad (1)$$

where k is the number of items loaded into working memory, S is the set size in the modality tested, h is hits, the proportion of changes correctly detected, and f is false alarms, the proportion of non-changing displays incorrectly judged to have changed. The formula is based on the assumption that a change in the stimuli can be detected on k/S of the trials in which there is a change and, if no change is detected, the participant nevertheless guesses "change" on some proportion g of the remaining trials. This proportion g is the same for change and no-change trials and thus drops out of the final formula. In our studies, because the number of items in each display is larger than the set size at which k typically reaches asymptote (see Cowan, 2001), we take k to reflect the limit in capacity. Calculating capacity estimates allowed a direct comparison between tasks, even though more to-be-remembered items were presented in the visual task than the auditory task.

Mean hit and correct rejection rates (from which false alarm rates were derived) for each task in each experiment can be found in Tables 1, 2 and 3. These rates were used to calculate estimates of memory capacity, also given in these Tables. Because the reward manipulation produced similar results across experiments, conditions were combined across experiments in statistical analyses.[1]

Our analyses focus on two questions: first, are the effects of reward consistent with a flexible division of resources? To answer this question, we considered data from all three experiments combined. In the combined analysis, we included all reward conditions with a total reward of 1000 points divided between the modalities (0/1,000, 250/750, etc). Because we were primarily interested in discovering whether any effect of reward was ordered with reward size, we used a Bayes factor approach (Jeffries, 1961; Kass &

Raftery, 1995). Using Bayes factors, it is possible to directly test which orderings the data support. Following the Bayesian analysis, we used ANOVAs and post hoc comparisons to compare condition means. This analysis is likely to be more familiar to the typical reader and, as we will show, the results of these two analyses converge upon the same conclusion.

Second, can these effects be attributed to the relative amount of reward assigned to each task, as is generally assumed, rather than the absolute amount of reward assigned to each task? To test this, we considered the equal high and low reward conditions (250/250 and 750/750) from Experiment 2.

Effects of relative reward amount on attention allocation

Bayesian analysis In order to address whether performance ordered with reward size, we fit a Bayesian hierarchical working memory model (Morey, 2010) based on Pashler (1988) assumptions to the data from both the auditory and visual task. Figure 2 (upper panel) shows the posterior mean capacities in the auditory task plotted against the posterior mean capacities in the visual task. The intervals are posterior standard deviations on the differences from the no reward condition for each task. In addition to estimates of the effects of reward, the model also provides estimates of the posterior probability of the true orderings of the reward effects. Because there is always uncertainty in sampling, the observed order of the reward conditions in our data may not be the true orderings. Our hypotheses regard true orderings of the conditions, considered jointly for the visual task and the auditory task. The posterior probability of each joint ordering may be used to construct tests of specific hypotheses regarding the allocation of resources.

We constructed hypotheses by considering three groups of joint orderings that might reasonably describe the deployment of attention resources. These nested cases are depicted in Fig. 3. The most basic theoretically plausible case supposes that attention allocation is inflexible, such that attention might be deployed only in an all-or-none manner. Any ordering in which the 0-point reward corresponds with lower estimates of capacity than all the rewarded conditions is consistent with this possibility, which is represented by A (Fig. 3, white region). A stronger claim is that attention can be flexibly divided between two stimulus sets. In that case, the payoff conditions in which there is some reward allocated to each task should produce capacity levels for a given task that are all higher than is found when there is no reward for that task, and lower than is found when there is reward exclusively for that task. Any ordering in which the 0-point reward corresponds with the lowest capacity estimate and the 1,000-point reward

---

[1] We manipulated the order of the presentation of the tasks between participants in Experiments 1 and 2, and within-participants in Experiment 3, but have collapsed across it in all reported analyses. Although we always observed significant effects of task order on visual capacity estimates (such that visual capacity estimates were higher when the visual task was presented and tested first), auditory capacity estimates did not vary with order, and there were no significant interactions involving both order and reward variables ($ps > .13$), whether order was manipulated within or between participants. Because the order manipulation does not address either of our main questions, and apparently does not impact our inferences regarding task rewards, we have chosen not to include it in our report of the results.

Table 1 Hit rates, correct rejection rates, and estimated capacity for the visual array and tone sequence tasks as a function of reward condition, Experiment 1

Stimuli	Measure	Task Reward in Points per Trial				
		0	250	500	750	1,000
Visual arrays						
	Hit rate	.56 (.17)	.61 (.20)	.59 (.16)	.62 (.20)	.67 (.18)
	CR rate	.69 (.18)	.77 (.12)	.83 (.12)	.81 (.13)	.84 (.12)
	Capacity	2.58 (2.36)	3.92 (1.96)	3.92 (1.90)	3.98 (2.40)	4.84 (1.74)
Tone sequences						
	Hit rate	.52 (.20)	.50 (.17)	.58 (.16)	.56 (.17)	.54 (.15)
	CR rate	.70 (.20)	.75 (.15)	.78 (.13)	.78 (.14)	.78 (.13)
	Capacity	1.59 (2.08)	2.05 (1.17)	2.82 (1.00)	2.59 (1.21)	2.37 (1.28)

Mean capacity estimates for both tasks were computed with Pashler's (1988) formula. $n = 32$. Standard deviations are in parentheses.

corresponds with the highest estimate is consistent with this hypothesis, represented by B (Fig. 3, dark grey region). Finally, we also considered C (Fig. 3, black region), which included only the perfect joint ordering of capacity estimates with reward amount in both tasks. This joint ordering represents the strongest case for a flexible allocation of resources.

To construct our Bayes factor tests, we first considered all three hypotheses equally likely a priori; thus, each hypothesis has even odds against each other hypothesis. A Bayes factor analysis proceeds by determining the amount by which the data change the odds of each hypothesis relative to each other hypothesis. The Bayes factor for the broadly flexible allocation hypothesis (B exclusive of C)

versus the inflexible allocation hypothesis (A exclusive of B and C) was 19, meaning that the data favored the flexible allocation hypothesis by 19 to 1, considered strong evidence against the inflexible allocation hypothesis (Jeffries, 1961). The inflexible allocation hypothesis was rejected even more decisively against the strong flexible allocation hypothesis (that is, C versus A); this Bayes factor was 53, which is considered very strong evidence. Finally, the data substantially favored the strong flexible allocation hypothesis over the weaker flexible allocation hypothesis (C versus B), by a factor of 6 to 1. Overall, the data provide substantial evidence in favor of the flexible resource allocation hypothesis. More details about these analyses and estimation procedures are given in the Appendix.

Table 2 Hit rates, correct rejection rates, and estimated capacity for the visual array and tone sequence tasks as a function of reward condition, Experiment 2

Stimuli	Measure	Task Reward in Points per Trial			
		0	250	750	1,000
Relative Reward Conditions					
Visual arrays					
	Hit rate	.55 (.19)	.63 (.18)	.63 (.20)	.61 (.20)
	CR rate	.73 (.18)	.80 (.15)	.82 (.14)	.87 (.08)
	Capacity	2.93 (1.94)	4.03 (2.10)	4.19 (2.37)	4.41 (1.78)
Tone sequences					
	Hit rate	.52 (.16)	.52 (.15)	.55 (.19)	.62 (.15)
	CR rate	.68 (.21)	.75 (.15)	.78 (.12)	.76 (.13)
	Capacity	1.52 (1.41)	2.00 (1.68)	2.54 (1.40)	2.91 (1.24)
		Low Reward (250)		High Reward (750)	
Absolute Reward Conditions					
Visual arrays					
	Hit rate	.60 (.15)		.63 (.19)	
	CR rate	.80 (.10)		.81 (.09)	
	Capacity	3.93 (1.56)		4.30 (1.95)	
Tone sequences					
	Hit rate	.56 (.17)		.55 (.20)	
	CR rate	.75 (.12)		.76 (.14)	
	Capacity	2.36 (1.40)		2.44 (1.51)	

In the absolute reward conditions, participants received the same number of points for correct responses on each task. Mean capacity estimates were computed with Pashler's formula (1988). $n = 30$. Standard deviations are in parentheses

Table 3 Hit rates, correct rejection rates, and estimated capacity for the visual array and tone sequence tasks as a function of reward condition, Experiment 3

Stimuli	Measure	Task Reward in Points per Trial		
		0	500	1,000
Visual Arrays				
	Hit rate	.60 (.13)	.66 (.10)	.69 (.14)
	CR rate	.68 (.17)	.75 (.10)	.80 (.08)
	Capacity	3.15 (1.40)	4.32 (1.19)	4.87 (1.33)
Tone Sequences				
	Hit rate	.52 (.13)	.57 (.10)	.58 (.12)
	CR rate	.69 (.13)	.73 (.09)	.73 (.10)
	Capacity	1.81 (1.07)	2.42 (0.80)	2.53 (1.05)

Mean capacity estimates were computed with Pashler (1988) formula for estimating capacity. $n = 24$. Standard deviations are in parentheses

Traditional analysis Two ANOVAs were carried out on the conditions with a total of 1,000 points per trial, one on capacity estimates from Experiments 1 and 2 which included task (auditory or visual) and reward for that task (0, 250, 750, or 1,000 points) as within-participants factors and another with the same factors on capacity estimates from Experiments 1 and 3, with three levels of reward (0, 500, or 1,000 points). In both cases, significant main effects of task type and reward were found and, in the analysis of Experiments 1 and 3, a significant interaction between task and reward was observed. Table 4 gives the details of each of these tests. Effects of reward were always consistent with the assumption of a trade-off; as reward to one task increased, capacity for that task increased.

Post-hoc Newman−Keuls comparisons from both ANOVAs suggested that trade-offs occurred and were to some extent flexible. For the visual task, one analysis showed that performance under all three allocations differed from one another (i.e., 0 < 500 < 1,000 points; 0 points: mean (M) = 2.83, SEM = .27, 500 points: M = 4.09, SEM = .22, 1000 points: M = 4.85, SEM = .21) and the other analysis similarly showed that both extremes (0 points: M = 2.75, SEM = .27, 1,000 points: M = 4.63, SEM = .22) differed from each other and from the two intermediate allocations (250 points: M = 3.98, SEM = .26, 750 points:　M = 4.08, SEM = .30), which did not differ, *p* = .69 (i.e., 0 < 250 = 750 < 1,000 points). These analyses provide evidence for at least three states of attention allocation: zero, divided, and full. For the auditory task, one analysis showed that a zero allocation produced poorer performance than a non-zero allocation (i.e., 0 < 500 = 1,000 points; 0 points: M = 1.68, SEM = .23, 500 points: M = 2.65, SEM = .12, 1,000 points: M = 2.44, SEM = .16) ; the other analysis showed that zero allocation (M = 1.56, SEM = .23)

produced poorer performance than 750 (M = 2.57, SEM = .16) or 1,000 allocation (M = 2.63, SEM = .16), and also that a minor allocation of 250 points (M = 2.03, SEM = .18) produced poorer performance than a larger allocation (i.e., 0 = 250 < 750 = 1,000 points). Thus, in the auditory task as in the visual, at least three states of attention allocation exist, though with different cutoff points for the two modalities. Both the ANOVAs and the Bayes factor analyses indicate that the data are highly consistent with the assumption of a trade-off determined by relative reward levels for accuracy in each task, supporting at least some degree of flexibility in the allocation of a shared attention resource.

It is worth noting that the ANOVAs are limited by the inability to include all conditions in a common analysis. Both the Bayesian and the traditional analyses provide strong support for at least weak flexibility with three reward states (Fig. 3b) and the Bayesian approach considering all
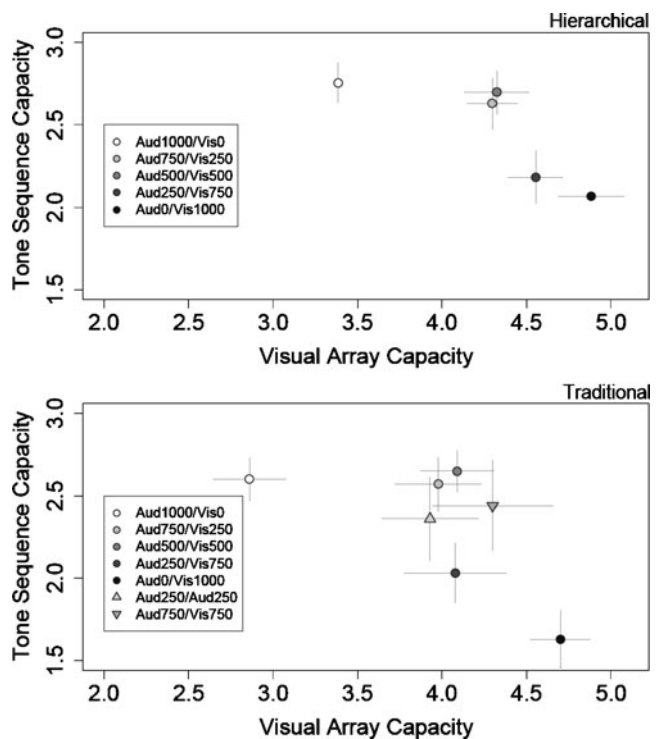

Fig. 2 Across experiments, visual array capacity estimates by tone sequence capacity estimates. Visual arrays included eight items and tone sequences included six items. Each *data point* represents concurrent reward conditions. Capacity estimates in the *upper panel* were calculated using hierarchical Bayesian techniques, with *error bars* representing posterior standard deviations on the differences from the no reward condition. In the *lower panel*, capacity estimates were calculated with Pashler's formula, collapsing across participants, with *error bars* representing standard errors of the mean. For the relative reward conditions (*circles*), *shades* correspond to reward level with the lightest representing the highest auditory reward and the darkest the highest visual reward. The absolute reward conditions from Experiment 2 are represented by the *light, upward-pointing triangle* (low reward) and the *dark, downward-pointing triangle* (high reward) in the *lower panel*
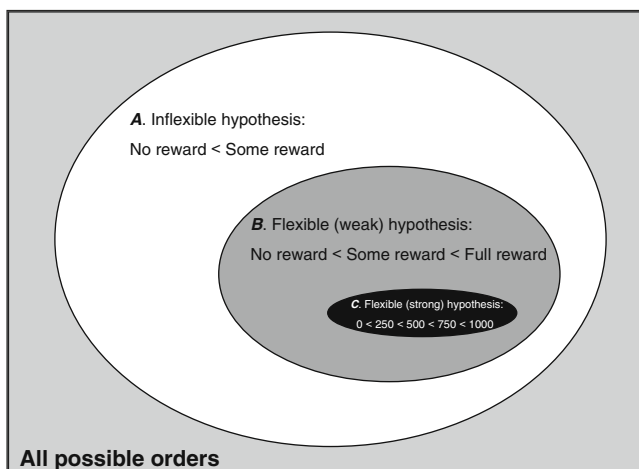
Fig. 3 Nested groups of orderings designated for Bayes factor analysis. The group **a** included joint orderings in which capacity in the 0-point reward condition was lower than for any other reward amount. **b** included joint orderings in which the 0-point reward resulted in the lowest estimates and the 1,000-point reward resulted in the highest estimates, and **c** included only the perfect joint ordering, in which capacity estimates always corresponded to reward value. Hypothesis tests were always between these groups, exclusive of each other

of the data together suggests that there may indeed be stronger, more finely graded flexibility (Fig. 3c).

Even though a large trade-off is apparent between these conditions, it does not seem to be the case that participants ignored the 0-reward task entirely. Although performance on the unrewarded task was low, it was consistently above chance. This, of course, could occur on the basis of memory that is automatically rather than effortfully encoded or if participants were motivated partially by factors other than the monetary reward.

Effects of absolute reward

Are the relative reward assignments causing participants to allocate their attention to one task at the expense of the other, as dual-task logic assumes, or are participants simply trying harder in the higher-reward task because of the absolute amount of reward in that task? Possibly, the amount of resources used is not constant, but is instead greater during blocks that include high-reward trials (e.g., 1,000 points for either task) than during blocks with mediocre rewards for both tasks. This question can be examined by comparing conditions in which the reward for both tasks is the same, both low or both high. If performance levels are higher with more overall reward, it suggests that part of the reward effect in other conditions may not be due to the allocation of attention to one task versus the other after all, but rather to the absolute reward for each task. To our knowledge, this control has never before been considered in an analysis of the effect of

reward on attention allocation. In Experiment 2, two absolute reward control conditions were included to test whether participants allocate attention differently when both tasks are worth the same low (250 points) or high (750 points) amount. To answer this question, we again used a Bayesian approach (we also report an ANOVA).

The hypothesis test of interest is whether there is evidence that the two equal reward conditions (250 / 250 and 750 / 750) yield different performance. To test this hypothesis, we took a nested models approach and compared the fit of a model in which an effect of reward level (250 / 250 or 750 / 750) is allowed, versus one in which this effect is constrained to be 0. For this comparison, it is most convenient to use the deviance information criterion (DIC; Spiegelhalter, Best, Carlin, & van der Linde, 2002). Although the Bayes factors for the discrete hypotheses tested in answering our first question were straightforward to compute, Bayes factors for point-null hypotheses are difficult to compute for complex models. The DIC is an appropriate alternative for model testing. DIC is a penalized-likelihood criterion similar to AIC (Akaike, 1974) and BIC (Schwartz, 1978); models with lower DIC values are preferred. The model with no reward effects yielded a lower DIC (17,949.3) than the model including reward effects (17,953.5), which favors the null hypothesis that absolute reward value does not greatly affect capacity estimates.

Likewise, a two-way ANOVA with task and absolute reward condition (low or high) as factors revealed only a main effect of task domain [$F(1, 29) = 21.75$, $MSe = 4.08$, $\eta^2_p = .43$]; absolute reward amount [$F(1, 29) = 0.79$, $p = 0.38$, $\eta^2_p = .03$] and the task by reward interaction [$F(1, 29) = 0.38$, $p = .54$, $\eta^2_p = .01$] were both non-significant. Though another two-way ANOVA including all of the reward conditions from Experiment 2 revealed a significant main effect of reward, post-hoc Newman–Keuls tests again uncovered no significant differences

Table 4 Task by reward ANOVAs

Experiments 1 and 2 (Rewards: 0, 250, 750, 1,000 points)				
	df	F	MSe	η^2_p
Task	1, 61	57.40*	5.98	.49
Reward	3, 183	22.75*	2.45	.27
Task × Reward	3, 183	2.04	2.43	.03
Experiments 1 and 3 (Rewards: 0, 500, 1,000 points)				
	df	F	MSe	η^2_p
Task	1, 55	87.37*	2.67	.61
Reward	2, 110	37.39*	1.90	.41
Task × Reward	2, 110	7.99*	1.876	.13

Experiment 1 $n = 32$, Experiment 2 $n = 30$, Experiment 3 $n = 24$

*$p < .05$

between the high and low absolute reward conditions, or between the high and low absolute reward conditions and the intermediate relative reward conditions in either task (ps .33 – .87). These analyses both suggest that, for both the visual array and tone sequence tasks, the value of the reward alone had no effect on capacity estimates. It therefore seems reasonable to conclude that the differences observed in the relative reward conditions were due to resource trade-offs and not merely to increased motivation during blocks with higher reward values.

General discussion

Using payoffs to manipulate volitional attention allocation and a novel Bayesian ordering analysis for hypothesis testing, our research contributes to ongoing discussions of the nature of attention allocation to working memory tasks. First, we have verified two assumptions about volition that are commonly made with respect to working memory performance, but have not to our knowledge been directly measured: (1) that a trade-off between visual-spatial and auditory-temporal memories occurs with manipulation of rewards as well as with manipulation of task difficulty, and (2) that the outcome of a payoff experiment is not only the result of changes in overall motivation between conditions. Our results were consistent with the assumption that some constant, limited resource is divided between the two tasks, and furthermore suggest that such a resource can be divided flexibly, not only shared in an all-or-none manner. This knowledge should lead to better specification of how the sharing of resources in working memory should be described. The vagueness with which these resources have been described in the past limits any researcher's ability to clearly falsify hypotheses about resource-sharing in working memory. We think that our findings, if applied to theories, will help to address this problem.

The studies reported above largely replicate previous findings by Morey and Cowan (2004, 2005) manipulating reward for task performance instead of task difficulty, but provide three important clarifications to their previous work. First, cross-domain interference cannot be attributed to any unknown, obligatory priority assignment to one task or stimulus type over the other; if this were the case, then we might not have observed an effect of reward consistent with the assumption of a trade-off. Second, the cross-domain interference Morey and Cowan documented was not due only to interference between verbal and visual-spatial stimuli, since the present studies show interference between tone and visual-spatial stimuli. Finally, our evidence is consistent with the proposition that a shared resource can be flexibly divided between two stimulus sets. This resource could be used to store information directly as

Cowan (2001, 2005) and Oberauer (2002) suggest, possibly in the form of the episodic buffer of Baddeley (2001), or it could reflect the contribution of something like the central executive proposed by Baddeley (1986, 2007). Our data do not allow us to conclude that one of these constructs is superior to the others, but instead establish new limits and possibilities about what theorists may reasonably claim any shared resource does.

The psychological literature includes some impressive cases of successful multi-tasking, such as auditory shadowing while playing the piano or engaging in a visual memory task (Allport et al., 1972) and reading while taking dictation (Hirst, Spelke, Reaves, Caharack, & Neisser, 1980). Although we have observed a trade-off between two concurrent tasks, our results also suggest that shared resources can be divided, at least to some degree. These extreme cases of excellent multi-tasking might be explained as well by a flexibly-divisible shared resource as by separate, independent resources (for a detailed analysis, see Cowan, 1995). Note also that this impressive multi-tasking was not necessarily cross-modal or cross-domain, so positing separate resources for stimuli from different domains does not necessarily explain these instances sufficiently.

Rather than suppose that there are multiple attention resources distinguished by reliance on a particular sensory domain (Wickens, 2002), it is possible to suppose that some other difference between the two tasks enabled their simultaneous completion. In the study of Hirst et al., even though the tasks both involved verbal stimuli, it is plausible that dictation may rely more on an automatic phonological store while the semantic processing needed for reading comprehension relied primarily upon a central attention resource. This explanation would be consistent with the process-based view of interference described by Marsh, Hughes, and Jones (2009), who showed that meaningful irrelevant speech does not equally interfere with all verbal tasks, but instead selectively interferes with a categorization task more than with a serial recall task. Furthermore, in these cases, participants were experts in at least one of the task's domains (e.g., piano-playing, dictation-taking), which might have enabled them to perform that task somewhat automatically, with little reliance on a shared attention resource. All things considered, a shared attention resource that can be flexibly divided seems at least as plausible as separate attention resources, and is a more parsimonious proposition.

Even so, it seems unlikely that all resources in working memory are shared between visual-spatial and auditory-verbal materials. We believe that the strongest versions of single, cross-domain models are probably inadequate, as are the strongest versions of multiple, domain-specific resource models. If memory were accomplished solely by a domain-

general store, one might expect to observe a one-for-one, item-to-item trade-off between items from concurrent stimulus sets, regardless of their domain. Here, we do not observe an item-to-item trade-off between visual-spatial and tone items, although there is clearly a cost for both tasks; the cost of one visual item was worth about 0.5 acoustic items. One possible explanation for this is that we did not include masks to eliminate lingering sensory memories; Saults and Cowan (2007) showed in a similar procedure that this was necessary to observe item-to-item trade-offs. Therefore, in our studies, reliance on a shared resource might have been somewhat reduced compared with the studies of Saults and Cowan. Auditory-verbal stimuli benefit more from automatically-activated echoic memory more than visual stimuli benefit from iconic memory (Darwin, Turvey, & Crowder, 1972), and the results of Saults and Cowan appear to confirm this. This factor might generally force a greater reliance on a central resource in visual, but not the auditory, tasks. This proposition is consistent with our data, in that the variance in capacity estimates with reward amounts was greater for the visual than the auditory task, as might be expected if participants can rely more on auditory than visual sensory memories to make a decision.

Gopher and Donchin (1986) suggested three possible interpretations for data that fit neither a strong separate resources nor a strong central resource model: (1) that one or both tasks are at ceiling, or are data-limited in Norman and Bobrow's (1975) terms; (2) there is a marginal divided attention cost, i.e., with less total usable resource in a divided-attention situation than in a full-attention situation; or (3) there is partial overlapping demand for a common resource. We argue that, in this case, the third explanation is most plausible. Performance levels on both the visual array and the tone sequence tasks were safely below ceiling level. Although we cannot truly test whether a there was a cost for dividing attention (because we have no data from pure single-task conditions), performance in the 1,000-point reward conditions was comparable to single-task performance observed in other studies (Cowan et al., 2005). Moreover, a divided-attention cost alone cannot explain the existence of three states of attention allocation in both modalities. The reason for supposing a partial overlapping demand for a common resource is that the central store might be supplemented by sensory memory, phonological and visuo-spatial memory (Baddeley, 1986), and/or activated elements of long-term memory (Cowan, 1988).

Others have suggested that concurrent tasks may draw upon one attentional resource that can be rapidly switched between tasks (e.g., Alvarez et al., 2005). Our results cannot confirm or falsify an attention-switching account; it is a potential direction for future studies of cross-domain resource sharing in working memory, but a challenging one,

inasmuch as it is difficult to distinguish empirically between resource switching and resource sharing when trade-offs are observed. Thus, it is possible that participants in our studies allocated attention on an all-or-none basis on each trial, and that our observation of flexible division actually reflects a different mixture of all-or-none states in each combination of reward conditions. This unresolved question is analogous to the question of whether capacity limits themselves occur because of concurrent attention to the items (Cowan, 2001) or one-at-a-time processing of a limited number of items repeatedly on a rapid time scale (Lisman & Idiart, 1995). Just as capacity limits are of interest without resolution of this tough issue, it is theoretically important that we observe a flexible trade-off between tasks, regardless of whether that trade-off occurs within individual trials or only in the balance between trials.

The relation between working memory storage and a more general attention resource used in other phases of processing in working memory also warrants further study. Fougnie and Marois (2009) found that, when two visual working memory tasks are presented together, the encoding and responding processes are unaffected, except when the storage capacity limit interferes with further encoding; also, Woodman and Vogel (2005) showed that, although consolidation of new items can occur while other items are being stored, the number of items that can be consolidated is limited by the amount of working memory capacity remaining. Both Fougnie and Marois and Woodman and Vogel ultimately argue that general attention processes, which are typically found to be involved in encoding and retrieval, are not used for visual working memory maintenance. Theoretically, one could extend that suggestion to the present study, but only if a working memory store that is *shared between modalities* is separate from attention. This suggestion is consistent with Baddeley's division of the central executive, a processing resource, from the episodic buffer, a domain-general store. However, it is not currently known whether this possible dissociation of general attention processes from storage in visual working memory tasks also holds for situations in which both visual and auditory or verbal representations must be maintained.

While our results are consistent with the division of a shared store or some shared attentional resource, we believe that at least it might be possible to use attention to boost fragile memories, in order to prevent over-writing (Cowan & Morey, 2006) or to delay the complete loss of complex representations in short-term memory (or "sudden death" as Zhang and Luck 2009 call this particular forgetting phenomenon). This role for attention could theoretically occur as refreshing, thought to be a non-automatic executive function (Raye, Johnson, Mitchell, Greene, &

Johnson, 2007). One difficulty with the suggestion that storage must be domain-specific and does not depend on attention is that it does not explain why the maintenance of items in a visual array is impaired by performance of a tone identification task that does not depend upon working memory storage (Stevanovski & Jolicoeur, 2007). One possibly reconciling speculation is that storage in working memory does not necessarily depend on attention, but nonetheless benefits from the application of it. This topic is certainly one that requires further consideration.

The present studies extend and clarify previous observations of cross-domain interference (Morey & Cowan, 2004, 2005; Stevanovski & Jolicoeur, 2007). Previous research clearly suggested the necessity of including some shared resource, perhaps even a shared memory store (Cowan & Morey, 2007; Saults & Cowan, 2007) in models of working memory. Our findings limit the nature of that shared resource, suggesting that it can be flexibly allocated between two stimulus sets.

Appendix

Bayesian hierarchical analysis details

In this appendix, we outline the details of the Bayesian hierarchical analysis used to argue for the flexible allocation of resources. A detailed description of the hierarchical model may be found in Morey (2010), along with simulations showing that it outperforms traditional analyses. Model details and GUI software (WoMMBAT: Working Memory Modeling using Bayesian Analysis Techniques) used to obtain model estimates can be freely obtained from drsmorey.org/research/rdmorey/.

In order to obtain Pashler (1988) capacity estimates in a hierarchical model, linear models are placed on two parameters: working memory capacity (κ) and the logistic transform of guessing probability (G). The logit function is used to transform the guessing parameters from (0,1) to the set of real numbers, in a manner similar logistic regression. Let $i = 1,...,86$ index participants, $j = 1,2$ index the order in which the auditory and visual tasks were performed, and $k = 1,...,5$ index reward conditions. Then, the capacity and guessing parameters for the i th participant, for the j th order, in the k th reward condition is

$$\kappa_{ijk} = \mu^{(\kappa)} + \eta_{1i} + \eta_{2j} + \eta_{3k}$$
$$\mathrm{logit}(G_{ijk}) = \mu^{(g)} + \gamma_{ijk}$$

where $\mu^{(\kappa)}$ and $\mu^{(g)}$ are the grand mean capacity and guessing parameters, η_1, η_2, η_3 are the random effects of participant, order, and reward condition, respectively, and γ is a random effect on guessing. The model above was fit to both the auditory and visual change detection data, yielding two sets of effect estimates. Details on priors and the specific methodology used to obtain estimates may be found in Morey (2010). One important detail is that the prior distribution on each of the effect parameters is the same; there is no a priori bias toward any particular ordering of the reward effects.

The main question we want to address regards the ordering of the reward effects, indexed by the five η_3 parameters. The output of the hierarchical analysis is a chain of vectors, each one of which is a sample from the joint posterior distribution of the parameters (see Rouder & Lu, 2005, for a tutorial for psychologists). To determine the posterior probability of any particular ordering of reward condition effects, we observe the order of the η_3 parameters on each iteration of the chain. The proportion of times each ordering appears in the chain is an estimate of the posterior probability of that ordering. The posterior probability of a joint ordering is obtained by multiplying together a posterior probability of an ordering in the auditory task and an ordering in the visual task. Summing the posterior probabilities of joint orderings in a group of orderings yields the posterior probability of that group of orderings. In order to compute the Bayes factor of one hypothesis to another, the ratio of posterior probabilities of the two hypotheses is divided by the prior probabilities. The Bayes factor thus represents the proportional change in the odds of one hypothesis to another due to observing the data.

References

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control, 19*, 716–723.

Allport, D. A., Antonis, B., & Reynolds, P. (1972). On the division of attention: A disproof of the single channel hypothesis. *The Quarterly Journal of Experimental Psychology, 24*, 225–235.

Alvarez, G. A., & Cavanaugh, P. (2004). The capacity of visual short-term memory is set both by visual information load and by number of objects. *Psychological Science, 15*, 106–111.

Alvarez, G. A., Horowitz, T. S., Arsenio, H. C., DiMase, J. S., & Wolfe, J. M. (2005). Do multielement visual tracking and visual search draw continuously on the same visual attention resources? *Journal of Experimental Psychology: Human Perception and Performance, 31*, 643–667.

Baddeley, A. D. (1986). *Working memory*. Oxford: Oxford University Press.

Baddeley, A. (2001). The magic number and the episodic buffer. *The Behavioral and Brain Sciences, 24*, 117–118.

Baddeley, A. D. (2007). *Working memory, thought, and action*. Oxford: Oxford University Press.

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. Bower (Ed.), *Recent advances in learning and motivation volume VIII*. New York: Academic Press.

Baddeley, A.D., & Logie, R.H. (1999). Working memory: The mulitple-component model. In A. Miyake & P. Shah (Eds.), *Models of Working Memory*. Cambridge, UK: Cambridge University Press.

Cocchini, G., Logie, R. H., Sala, S. D., MacPherson, S. E., & Baddeley, A. D. (2002). Concurrent performance of two memory tasks: Evidence for domain-specific working memory systems. *Memory & Cognition, 30*, 1086–1095.

Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information processing system. *Psychological Bulletin, 104*, 163–191.

Cowan, N. (1995). *Attention and memory: An integrated framework. Oxford psychology series 26*. New York: Oxford University Press.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *The Behavioral and Brain Sciences, 24*, 87–185.

Cowan, N. (2005). *Working memory capacity*. New York: Psychology Press.

Cowan, N., Elliott, E. M., Saults, J. S., Morey, C. C., Mattox, S., Hismjatullina, A., & Conway, A. R. A. (2005). On the capacity of attention: Its estimation and its role in working memory and cognitive aptitudes. *Cognitive Psychology, 51*, 42–100.

Cowan, N., & Morey, C. C. (2006). Visual working memory depends on attentional filtering. *Trends in Cognitive Sciences, 10*, 139–141.

Cowan, N., & Morey, C. C. (2007). How can dual-task working memory retention limits be investigated? *Psychological Science, 18*, 686–688.

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.

Darwin, C. J., Turvey, M. T., & Crowder, R. C. (1972). An auditory analogue of the Sperling partial report procedure: Evidence for brief auditory storage. *Cognitive Psychology, 3*, 255–267.

Fougnie, D., & Marois, R. (2009). Dual-task interference in visual working memory: A limitation in storage capacity but not in encoding or retrieval. *Attention, Perception, & Psychophysics, 71*, 1831–1841.

Gold, J. M., Wilk, C. M., McMahon, R. P., Buchanan, R. W., & Luck, S. J. (2003). Working memory for visual features and conjunctions in schizophrenia. *Journal of Abnormal Psychology, 112*, 61–71.

Gopher, D., & Donchin, E. (1986). Workload: An examination of the concept. In K. Boff & L. Kaufman (Eds.), *Handbook of Human Perception and Performance*. New York: Wiley.

Hirst, W., Spelke, E., Reaves, C., Caharack, G., & Neisser, U. (1980). Dividing attention without alternation or automaticity. *Journal of Experimental Psychology: General, 109*, 98–117.

Jefferies, H. (1961). *The theory of probability*. Oxford: Oxford University Press.

Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association, 90*, 773–795.

Kumar, A., & Jiang, Y. (2005). Visual short-term memory for sequential arrays. *Memory & Cognition, 33*, 488–498.

Lisman, J. E., & Idiart, M. A. P. (1995). Storage of 7±2 short-term memories in oscillatory subcycles. *Science, 267*, 1512–1515.

Logie, R. H., Zucco, G. M., & Baddeley, A. D. (1990). Interference with visual short-term memory. *Acta Psychologica, 75*, 55–74.

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature, 390*, 279–281.

Marsh, J. E., Hughes, R. W., & Jones, D. M. (2009). Interference by process, not content, determines semantic auditory distraction. *Cognition, 110*, 23–38.

Miller, G. A., Galanter, E., & Pribram, K. H. (1960). *Plans and the structure of behavior*. New York: Holt, Rinehart and Winston.

Miyake, A., & Shah, P. (Eds.). (1999). *Models of working memory: Mechanisms of active maintenance and executive control*. Cambridge: Cambridge University Press.

Morey, R. D. (2010). Hierarchical Bayesian models for estimating working memory capacity. *Journal of Mathematical Psychology.* doi:10.1016/j.jmp.2010.08.008

Morey, C. C., & Cowan, N. (2004). When visual and verbal memories compete: Evidence of cross-domain limits in working memory. *Psychonomic Bulletin & Review, 11*, 296–301.

Morey, C. C., & Cowan, N. (2005). When do visual and verbal memories conflict? The importance of working-memory load and retrieval. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 31*, 703–713.

Naveh-Benjamin, M., & Jonides, J. (1984). Maintenance rehearsal: A two-component analysis. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 10*, 369–385.

Navon, D., & Gopher, D. (1979). On the economy of the human-processing system. *Psychological Review, 86*, 214–255.

Norman, D. A., & Bobrow, D. G. (1975). On data-limited and resource-limited processes. *Cognitive Psychology, 7*, 44–64.

Oberauer, K. (2002). Access to information in working memory: Exploring the focus of attention. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 28*, 411–421.

Pashler, H. (1988). Familiarity and visual change detection. *Perception & Psychophysics, 44*, 369–378.

Phillips, W. A. (1974). On the distinction between sensory storage and short-term visual memory. *Perception & Psychophysics, 16*, 283–290.

Raye, C. L., Johnson, M. K., Mitchell, K. J., Greene, E. J., & Johnson, M. R. (2007). Refreshing: A minimal executive function. *Cortex, 43*, 135–145.

Rouder, J. N., & Lu, J. (2005). An introduction to Bayesian hierarchical models with an application in the theory of signal detection. *Psychonomic Bulletin & Review, 12*, 573–604.

Rouder, J. N., Morey, R. D., Cowan, N., Zwilling, C. E., Morey, C. C., & Pratte, M. S. (2008). An assessment of fixed-capacity models of visual working memory. *Proceedings of the National Academy of Sciences of the United States of America, 105*, 5975–5979.

Saults, J. S., & Cowan, N. (2007). A central capacity limit to the simultaneous storage of visual and auditory arrays in working memory. *Journal of Experimental Psychology: General, 136*, 663–684.

Schneider, W., Eschmann, A., & Zuccolotto, A. (2002). *E-Prime user's guide*. Pittsburgh: Psychology Software Tools.

Schwartz, G. (1978). Estimating the dimension of a model. *Annals of Statistics, 6*, 461–464.

Sperling, G., & Dosher, B. (1986). Strategy and optimization in human information processing. In K. Boff & L. Kaufman (Eds.), *Handbook of Perception and Performance*. New York: Wiley.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & van der Linde, A. (2002). Bayesian measures of model complexity and fit (with discussion). *Journal of the Royal Statistical Society, Series B (Statistical Methodology), 64*, 583–639.

Stevanovski, B., & Jolicœur, P. (2007). Visual short-term memory: Central capacity limitations in short-term consolidation. *Visual Cognition, 15*, 532–563.

Todd, J. J., & Marois, R. (2004). Capacity limit of visual short-term memory in human posterior parietal cortex. *Nature, 428*, 751–754.

Todd, J. J., & Marois, R. (2005). Posterior parietal cortex activity predicts individual differences in visual short-term memory capacity. *Cognitive, Affective & Behavioral Neuroscience, 5*, 144–155.

Vergauwe, E., Barrouillet, P., & Camos, V. (2010). Do mental processes share a domain-general resource? *Psychological Science, 21*, 384–390.

Vogel, E. K., McCullough, A. W., & Machizawa, M. G. (2005). Neural measures reveal individual differences in controlling access to working memory. *Nature, 438*, 500–503.

Wheeler, M. E., & Treisman, A. M. (2002). Binding in short-term visual memory. *Journal of Experimental Psychology: General, 131*, 48–64.

Wickens, C. D. (1984). Processing resources in attention. In R. Parasuraman & D. R. Davies (Eds.), *Varieties of attention*. New York: Academic Press.

Wickens, C. D. (2002). Multiple resources and performance prediction. *Theoretical Issues in Ergonomics Science, 3*, 159–177.

Woodman, G. F., & Vogel, E. K. (2005). Fractionating working memory: Consolidation and maintenance are independent processes. *Psychological Science, 16*, 106–113.

Xu, Y., & Chun, M. M. (2006). Dissociable neural mechanisms supporting visual short-term memory for objects. *Nature, 440*, 91–95.

Zhang, W., & Luck, S. J. (2009). Sudden death and gradual decay in visual working memory. *Psychological Science, 20*, 423–428.