AUTHOR-FINAL VERSION OF IN-PRESS MANUSCRIPT AT THE JOURNAL OF EXPERIMENTAL PSYCHOLOGY: LEARNING, MEMORY, AND COGNITION

Spatial sequences, but not verbal sequences, are vulnerable to general interference during retention in working memory

Candice C. Morey

Department of Psychology, University of Edinburgh

Monica D. Miron

Department of Psychology, University of Edinburgh

Author Note

We thank Tuulia Torppa, Josephine Steeghs, Sabine Kästner, and Leire Martin Mendez for assistance with data collection related to this project. Preliminary data influencing the experiments reported here were collected at the Heymans Institute for Psychological Research at the Rijksuniversiteit Groningen. Experiment 1 formed part of Monica Miron's MSc dissertation at the University of Edinburgh. Data, analysis scripts, and experimental software are freely available on our Open Science Framework page (https://osf.io/s6w4b/).
Address correspondence to Candice C. Morey, Department of Psychology, 7 George Square, EH8 9JZ Edinburgh, U.K. C.Morey@ed.ac.uk.

Word count: 6,437
Date submitted: 2 November 2015; Resubmitted 25 January, 2016; Accepted 7 March 2016

Abstract

Among models of working memory, there is not yet a consensus about how to describe functions specific to storing verbal or visual-spatial memories. We presented aural-verbal and visual-spatial lists simultaneously and sometimes cued one type of information after presentation, comparing accuracy in conditions with and without informative retro-cues. This design isolates interference due specifically to maintenance, which appears most clearly in the uncued trials, from interference due to encoding, which occurs in all dual-task trials. When recall accuracy was comparable between tasks, we found that spatial memory was worse in uncued than in retro-cued trials, whereas verbal memory was not. Our findings bolster proposals that maintenance of spatial serial order, like maintenance of visual materials more broadly, relies on general rather than specialized resources, while maintenance of verbal sequences may rely on domain-specific resources. We argue that this asymmetry should be explicitly incorporated into models of working memory.

Keywords: working memory, spatial memory, verbal memory, short-term memory, dual-task interference

Abstract word count: 148

Spatial sequences, but not verbal sequences, are vulnerable to general interference during retention in working memory

Several models of working memory propose distinct mechanisms devoted to the maintenance or rehearsal of strictly verbal or strictly visual information (Baddeley, 2012; Barrouillet & Camos, 2015; Logie, 2011). These proposals are supported by research showing that concurrently performing two verbal tasks or two visual-spatial tasks results in worse performance than concurrently performing one verbal and one visual task (Cocchini, Logie, della Sala, MacPherson, & Baddeley, 2002; Logie, Zucco, & Baddeley, 1990). However, performing two cross-domain tasks concurrently can also lead to substantial costs in performance compared to performing either task alone (e.g., Logie, et al., 1990), and task combinations with higher cognitive load increase dual-task costs regardless of the tasks' materials (Vergauwe, Barrouillet, & Camos, 2009; 2010). This combination of evidence calls for explanations that incorporate both domain-specific and domain-general processes within a working memory system. Various frameworks accomplish this in different ways, some calling for domain-specific maintenance modules (Baddeley, 2012), others adding specialized rehearsal mechanisms for verbal information (Baddeley, 2012; Barrouillet & Camos, 2015), or for both verbal and visual information (Logie, 2011). Other influential hypotheses suppose that short-term maintenance is accomplished without activation in dedicated short-term memory buffers. Short-term memory phenomena may be attributed to long-term memories maintained in a privileged state of heightened activation (e.g., Cowan, 2005) or considered to be determined by similar principles which determine longer-term memory, such as cue-directed retrieval (Nairne, 2002).

Observations of different dual-task costs for within- and cross-domain task combinations have traditionally motivated modular accounts of working memory that posit distinctly specialized storage buffers (Baddeley & Hitch, 1974; Logie, et al., 1990). However, very little research has attempted to isolate the main processes that the domain-specific modules are believed to support, namely storage and rehearsal. There is a large scientific literature investigating effects of doing one task with and without a memory load (e.g., Baddeley & Hitch, 1974; Cocchini, et al., 2002), or of combining storage in one modality with processing in another (e.g., Shah & Miyake, 1996; Vergauwe, et al., 2010; Vergauwe, Dewaele, Langerock, & Barrouillet, 2012), but none of these designs requires verbal and visual-spatial information to be *maintained* simultaneously. Given that modular working memory models differ in precisely which modules are posited, and that some modules are specifically for maintenance, it is surprising that more effort has not been made to measure interference due specifically to concurrent maintenance of verbal and visual-spatial information, rather than to encoding or processing one kind of information while maintaining another. This sort of evidence is needed to validate proposals of domain-specific short-term memory buffers, and whether or not these ideas are supported, to better understand the limits that arise in working memory.

There is ample reason to question whether storage in working memory is less specialized than is widely believed. Regardless of sensory domain or code, serial recall tasks produce similar error functions, with superior performance for the first and final items in a list compared to the middle items (e.g., Cortis, Dent, Kennett, & Ward, 2015; Smyth, Hay, Hitch, & Horton, 2005; Ward, Avons, & Melling, 2005; see also Guérard & Tremblay, 2008 for evidence from order reconstruction). Memory for serial order is more sensitive to interference from other order-memory tasks than from tasks requiring memory for only item identity (Depoorter & Vandierendonck, 2009), regardless of their presentation modalities

(Vandierendonck, 2015), and interleaved verbal and spatial lists cannot be remembered as accurately as lists presented in isolation (Morey & Mall, 2012). Together this evidence implies that verbal and visual-spatial memory depend on common resources, but does not reveal whether those resources are needed primarily for maintenance or for other cognitive processes. Application of domain-general resources during maintenance would be consistent with embedded activation models of working memory (e.g., Cowan, 2005; Oberauer, 2013), while classic modular working memory models (e.g., Baddeley, 2012; Logie, 2011) instead suppose that domain-general resources are needed for manipulating or attending information, but not actually for maintaining it. The most recent version of the Time-Based Resource Sharing model (TBRS; Barrouillet & Camos, 2015) offers a hybrid account positing that domain-general resources are necessary for attending information and maintaining it, but additionally proposes a specialized rehearsal mechanism capable of maintaining only verbal representations.

Other evidence does suggest specialization of function, but likewise cannot distinguish between multi-component working memory frameworks that include specialized verbal and visual resources in addition to domain-general ones (Baddeley, 2012; Logie, 2011) and frameworks that include only verbal and domain-general resources (Barrouillet & Camos, 2015). Saito, Logie, Morita, and Law (2008) compared recall for sequences of kanji varying in phonological and visual similarity. Because articulatory suppression impaired recall of phonologically-similar sequences only, Saito et al. argued that serial memory is subserved by at least two specialized processes (see also Logie, et al., 2000), one of which must be verbal. It is implied that the remaining resource would be exclusively for maintaining visual representations, but positing an additional domain-general resource capable of holding visual representations would accommodate these findings equally well (cf. Phillips & Christie, 1977).

Similarly, though tasks meant to selectively interfere with verbal rehearsal (e.g., articulatory suppression) or spatial rehearsal (e.g., sequential tapping) produce patterns consistent with assumptions of code-based modularity (Guérard & Tremblay, 2008; Meiser & Klauer, 1999; though see Jones, et al., 1995), these double dissociations could be accommodated just as well by the restricted modularity of TBRS (Barrouillet & Camos, 2015) as by more extensive multi-component frameworks. Though much evidence for theorizing exists, none of it provides comparisons that can firmly distinguish between models that posit domain-specific storage buffers and models that do not.

Evidence suited for this purpose may be obtained by combining dual-task memory tests with retro-cue paradigms (Griffin & Nobre, 2003). In retro-cue paradigms, researchers manipulate the content of cues after the stimuli have been presented but before recall is allowed to occur. These designs are ideal for restricting inferences to processes occurring during maintenance. When two stimulus sets have been encoded and then one is cued for testing, participants may focus any maintenance attempts on that set while neglecting the other set. In trials in which no cue is provided, participants must divide whatever resources are available to support maintenance between the two sets because they do not know which will be tested. Suppose that one of the to-be-remembered stimulus sets contains verbal items and one contains visual or spatial items. If there are separate resources for maintaining verbal versus visual materials, then retro-cueing may provide little, if any, benefit to recall, because the cognitive resources for maintaining the stimuli do not conflict with each other. However, if resources for maintaining verbal and visual materials are shared, then focusing on only one of the two stimulus sets should boost performance. Most importantly for hypothesis testing, effects of retro-cues isolate processes that occur after the stimuli have been presented, allowing researchers to hypothesize about processes specific to maintenance.

Retro-cue designs have previously been applied to elements of visual displays (e.g., Griffin & Nobre, 2003; Rerko, Souza, & Oberauer, 2014; Shimi, Nobre, Astle, & Scerif, 2014) and to encoding and maintaining two sets of stimuli (Cowan & Morey, 2007; C. Morey, R. Morey, van der Reijden, & Holweg, 2013). Cowan and Morey (2007) presented two stimulus sets (either two verbal lists, two color arrays, or one of each) sequentially and then retro-cued one set in some trials. Cowan and Morey's data suggested that concurrently maintaining verbal sequences of letters or digits and visual arrays of colored squares required a common resource, but that encoding dissimilar sets incurred little cost (see also Fougnie & Marois, 2009). Morey et al. (2013) similarly presented two stimulus sets sequentially (always one verbal and one visual-spatial), and sometimes retro-cued one set. Morey et al. manipulated the number of to-be-remembered items in each set and showed that maintenance of visual arrays deteriorated as the number of to-be-remembered verbal items increased, while maintenance of verbal lists remained constant regardless of visual memory load. Altogether, these findings demonstrate that resources needed for maintenance of verbal and visual memoranda overlap, but that this conflict is more damaging for visual than verbal memoranda. This suggests a possible asymmetry in resources needed to explain how verbal and visual information are maintained at once.

In previous studies using retro-cues to isolate maintenance of verbal and visual-spatial materials, stimulus sets were presented sequentially. These designs are ideal for learning whether common or distinct resources are needed for maintenance or encoding specifically. By manipulating the order of stimulus presentation, it is possible to discern whether encoding a second stimulus set provokes interference even if the first set is ultimately retro-cued for retention. This design affords estimates of interference costs attributable to encoding and maintenance specifically. But one might argue that previous findings of Cowan and Morey

(2007) and Morey et al. (2013) depended on procedural differences between encoding sequences versus simultaneous arrays. If memory for verbal and spatial sequences were contrasted instead, the procedures for presenting and testing the memoranda may be better equated, allowing for clearer inferences about whether asymmetries in interference during maintenance are really due to differences in the domain of the memoranda, rather than other task differences.

No previous study has employed a retro-cue design to test whether maintenance of two sequences of memoranda from different domains requires common resources. Such evidence could distinguish between the various configurations of multiple-component working memory models (Baddeley, 2012; Barrouillet & Camos, 2015; Logie, 2011). Each of the current modular working memory models posits domain-general resources for controlling attention, thus each modular model predicts that concurrently encoding a verbal and a spatial list would provoke some cost compared to encoding either type alone. But if there are distinct modules specialized for maintaining verbal and visual-spatial information, then no additional cost should be observed in an uncued condition compared to a retro-cued condition for either kind of memoranda, because whatever information has been encoded would be maintained within its designated module, protected from any further interference from the maintenance occurring in the other module. However, if specialized resources are only available for verbal maintenance as Barrouillet and Camos (2015) suggest (consistently with Camos, Lagner, & Barrouillet, 2009, and Camos, Mora, & Oberauer, 2011), then maintenance of verbal lists will benefit from extra protection that is not available for spatial information. Memory for spatial lists will deteriorate more than memory for verbal lists when they must both be maintained in the uncued conditions, compared to when the retro-cue focuses maintenance efforts on the spatial memoranda.

Some evidence already hints that maintenance of spatial sequences suffers more from non-specific interference than maintenance of verbal sequences. For instance, memory for spatial sequences suffers more than verbal ones from interleaved judgments, even if the judgments are not based on spatial reasoning (Bayliss, Jarrold, Gunn, & Baddeley, 2003; Shah & Miyake, 1996; Vergauwe, et al., 2010; Vergauwe, et al., 2012). However, these findings do not allow us to conclude that *maintaining* verbal information differentially interferes with maintaining spatial information. Performing processing tasks should require use of the domain-general resources that each variant of the multi-component working memory model posits. This existing evidence therefore cannot distinguish between the various multi-component working memory models.

The research we report in this article provides evidence better fit for the purpose of distinguishing between the existing modular working memory models. We examined whether serial verbal and visual-spatial information are both vulnerable to interference from maintaining another list, or whether visual information is more vulnerable. We compared serial order reconstruction for verbal and spatial stimuli under three conditions: 1) single-list presentation, where no interference due to encoding or maintaining another list could occur; 2) dual-list presentation with an informative retro-cue; and 3) dual-list presentation without a retro-cue. If these tasks do not conflict during encoding or maintenance, then performance should be equivalent across conditions. This outcome would be consistent with any model that suggests that both encoding and maintenance of information occurs within separate, domain-specific systems (e.g., Shah & Miyake, 1996). If there is conflict during encoding, which may occur due to the division of a shared attention resource like the central executive (Baddeley, 2012), then accuracy will decrease with dual-list compared to single-list presentation. If conflict is limited to encoding and does not occur during maintenance, then no difference between the two dual-list conditions will emerge because the retro-cue can only have an impact after encoding. However, if conflict occurs during both encoding and maintenance, then dual-list recall with retro-cues will exceed recall without retro-cues, consistently with models that do not posit separate storage resources for verbal and spatial memory at all (e.g., Cowan, 2005; Oberauer, 2013). The focused modularity proposed in the most recent TBRS model (Barrouillet & Camos, 2015) further predicts that more interference during maintenance will be observed for spatial than for verbal memoranda. In all cases, any interference occurring during encoding of the stimulus sets is held constant, allowing us to infer that any additional decrease in accuracy in the uncued compared to the retro-cued conditions reflects common resources needed to maintain the two stimulus sets, beyond the contributions any common resource makes to encoding the sets.

We also manipulated the occurrence of irrelevant deviants during list presentation to further observe how attentional capture during encoding might affect performance when attention is divided. This was exploratory, because predicting whether novelty should facilitate or hinder concentration is complicated by a number of other factors (Hughes, Vachon, & Jones, 2007; Parmentier, Elsley, & Ljungberg, 2010; Schomaker & Meeter, 2014). Infrequent deviant stimuli may orient attention to the channel including the deviant (e.g., Cherry, 1953); if so, then an aural deviant during stimulus presentation could distract from visual encoding, and vice versa. Deviant stimuli could also provoke an overall broadening of attention that may facilitate performance (e.g., Schomaker & Meeter, 2014); if this is the case, deviants in either list might improve performance. Further complicating this matter are predications arising from load theory (e.g., Lavie, 2010). Under high visual perceptual load (or visual memory load; Konstantinou, Beal, King, & Lavie, 2014), visual distractors are less likely to be processed than under low perceptual load. With high visual perceptual load, auditory distractors are likewise less likely to be processed (Raveh & Lavie, 2015) than under low perceptual load. However, in our paradigm the deviants occurred to the to-be-remembered stimuli themselves. It is unclear whether this would constitute a mnemonic or perceptual load in Lavie's terms, and thus uncertain how to predict the effects of the distractors under the principles of load theory. Because the disparities of previous literature suggest asymmetries in effects of attentional capture across domains, our sole aim in including this manipulation was to see whether any effects of cross-domain novel deviants could be detected.

Experiment 1

Method

Participants. Twenty-four participants recruited from the University of Edinburgh community took part in the experiment, receiving £7 per hour in compensation (16 females and 8 males, 21-28 years old, *M*= 24.63, *SD*= 2.02).

Materials. All of our materials are publicly available on Open Science Framework (OSF) (https://osf.io/s6w4b/). Stimulus presentation and response collection were controlled using E-Prime (Schneider, Eschmann, & Zuccolotto, 2002).

Spatial positions were marked by 20×20 pixel (0.53 cm) black squares set against a gray background, and were randomly selected without replacement on each trial from 10 pseudo-randomly chosen potential positions, which were scattered within a 13.23×14.29 cm area in the center of the screen, with at least 2.65 cm between the closest squares in the set.

Verbal list items were selected from a set of ten consonants (B, F, J, K, L, N, R, V, Y, Z). Recordings were generated in an artificial female voice and ranged from 430 to 520 ms in duration. Five consonants were randomly selected without replacement per trial.

Irrelevant deviant features were sometimes introduced within a trial. For spatial lists, the placeholder for one spatial position was red rather than the typical black. For consonants, one consonant was spoken in an artificial male voice. When a deviant occurred, the serial position of the deviant was selected randomly.

After stimulus presentation, a cue appeared. The “eye”, “ear”, or “either” retro-cue was aurally presented along with a corresponding visual icon. The recall screen presented either the set of ten possible consonants or positions, from which participants selected a total of 5 items, clicking them in serial order.

Procedure. Participants were individually tested in small, private rooms containing two computer stations separated by a panel. The experimenter obtained written consent and supervised a short practice session, but otherwise remained on the opposite side of the panel from the participant unless summoned.

Participants read scripted instructions for the serial reconstruction task from the computer screen. Instructions emphasized accurate memory for order and explained that the retro-cue sometimes warned them which of two simultaneously presented lists would be tested. Before beginning practice trials, the participant was asked to explain the instructions to the experimenter, so that the experimenter could confirm that the participant understood the task. During the supervised practice session, the participant completed four single-list trials followed by four dual-list trials. Participants received feedback during practice indicating the number of items chosen in the correct order out of their five responses.

An example trial sequence is depicted in Figure 1. Each trial started with a fixation “+” presented in the middle of the screen along with an auditory signal “Ready”. Stimulus presentation began 250 ms later. For both single- and dual-list presentations, each stimulus was presented for 500 ms (or for sounds, for their slightly varying lengths) with blank inter-stimulus intervals (ISIs) of 300 ms. Onsets of auditory and visual stimuli were simultaneous in dual-list presentations. After the presentation of fifth and final item(s), a retro-cue including a visual icon (an eye, a pair of headphones, or a ?) and an aural instruction ("eye", "ear", "either") occurred, followed by a 3000-ms retention interval. Note that in all conditions, participants observed some aural and visual signal after the final stimulus, thereby controlling for potential suffix effects in both modalities (Parmentier, Tremblay, & Jones, 2004). In single-list trials, participants always received the *eye* cue for spatial position lists and the *ear* cue for consonant lists. For dual-list presentations, when participants received the *eye* cue, they knew they would definitely be tested about the spatial list. Likewise, the *ear* cue meant that the consonant list would definitely be tested. The *either* cue provided no information about the eventual test. The response screen showed either a horizontally-oriented list of the 10 possible consonants or a display of the 10 possible positions. Participants responded by attempting to click the stimuli in the order in which they were presented during the trial using the mouse. When the participant clicked an item, it turned green and could not be selected again. There was no time limit for responding; after five valid selections the participant was invited to initiate the next trial.

Participants completed four blocks of 48 experimental trials each, for a total of 192 trials per 90-minute session. Of these, half tested spatial serial memory and half tested verbal serial memory. For each of these, 24 trials were single-list trials, with 36 trials each in the cued and uncued dual-list conditions. Dual-list trials were equally like to contain no irrelevant deviants, an irrelevant visual deviant, or an irrelevant aural deviant. Deviants occurred on half of single-list trials. Participants took a break of at least one minute in between blocks. The experimental session lasted no more than 90 minutes.

Results

We first present analyses of the verbal and spatial reconstruction tasks separately in order to provide rich descriptive information about task performance. We analyzed proportion of correct responses by serial position (1-5), presentation-cue condition (single-list presentation, dual-list with cue, dual-list without cue), and irrelevant deviant condition (none, visual, or aural), as well as between-participant variance. Our main prediction involved the retro-cues: if verbal and spatial memoranda compete for maintenance resources, then accuracy will be lower in the uncued than in the cued dual-list condition. If interference only occurs because of conflict during encoding the stimuli, then accuracy will not be any lower in the uncued than in the cued dual-list conditions, though there would be a general cost for dual- compared to single-list presentations. After presenting full separate analyses for each task, we examine the effect of presentation-cue condition across tasks.

All of our inferences are based on Bayes factor ANOVA (Morey and Rouder, 2015; Rouder, Morey, Speckman, & Province, 2012). In this approach, the full ANOVA model is specified and Bayes factors are computed comparing all the subsidiary models including every combination of the factors against a baseline of only between-participant variance. We conducted all of our modeling using the default settings of the BayesFactor package (version 0.9.12-2). The model with the largest Bayes factor is favored by the data. Unlike *p*-values, Bayes factors are meant to be interpreted continuously: the larger the value, the greater the evidence favoring the model. Interpretation of the magnitude of a Bayes factor depends on context, but for guidance, various scales agree that Bayes factors exceeding 20 constitute strong evidence (e.g., Jeffreys, 1961; Kass & Raftery, 1995). Unlike classical analyses, these analyses allow us to argue in favor of null hypotheses.

Verbal serial reconstruction. Mean proportions correct for verbal serial reconstruction are shown in Figure 2. The best ANOVA model included main effects of serial position, presentation-cue condition, and deviant condition (BF=7.33x10^{60}, ±1.32%). This model exceeded the model excluding the presentation-cue condition by a factor of more than

63,000, indicating that the effect of presentation-cue condition is robust. Reconstruction was best in the single-list condition (*M*=0.85, *SD*=0.17), and did not appear to differ much in the two dual-list conditions (M_{Cued}=0.81, *SD*=0.19; M_{Uncued}=0.79, *SD*=0.21). The effects of the deviants were also quite small, with somewhat lower accuracy with aural (*M*=0.79, *SD*=0.19) and visual (*M*=0.81, *SD*=0.21) deviants compared with no deviants (*M*=0.83, *SD*=0.19). However, the model including the deviant effect was favored over the model excluding it by a factor of more than 1300, indicating a robust, though small, effect. Our analysis yielded no evidence that deviant condition interacted with any other factor, with *BF*s favoring the simpler main effects model by factors of at least 3.80.

Spatial serial reconstruction. The same Bayesian ANOVA was performed on spatial serial reconstruction data (see Figure 3). The best ANOVA model included main effects of serial position, presentation-cue condition, and deviant type ($BF=5.33\times10^{57}$, ±1.34%). Again, evidence for the effect of presentation-cue condition was overwhelming (*BF*> 1 million). Performance was best with single-list presentation (*M*=0.63, *SD*=0.23); accuracy in both dual-list conditions was lower (M_{Cued}=0.55, *SD*=0.24; M_{Uncued}=0.51, *SD*=0.22). Though deviant type was present in the best model, the best model was favored only slightly over the same model with deviant type excluded (*BF*=1.15). If anything, occasional aural deviants presented in the verbal lists seemed to reduce accuracy (*M*=0.51, *SD*=0.24) relative to trials without deviants (*M*=0.58, *SD*=0.23) with intermediate performance with occasional color deviants in the spatial lists (*M*=0.56, *SD*=0.23).

Combined analysis. Our main hypotheses concerned whether the presentation-cue manipulation would affect spatial reconstruction performance differently than verbal reconstruction performance, despite general decreases in dual-list performance observed in both domains. To test this formally, we entered the mean number of items correctly ordered per list into a common Bayesian ANOVA, including probe domain (verbal or spatial) and presentation-cue condition as factors. Mean values are shown in Figure 4. For simplicity, we collapsed across serial positions and deviant conditions because there was no evidence to suggest interactions between these factors and the presentation-cue factor in either separate analysis. We compared two ways of coding presentation-cue in order to test whether differences emerged for one task and not another: we ran ANOVA models with a presentation-cue variable that differentiated between all three conditions (single-list presentation, dual-list with informative cue, dual-list without the cue) and also with a presentation-cue variable specifying only two levels (single-list, and dual-list, where dual-list included all dual-list trials regardless of the informativeness of the cue). If there are no differences in the effect of presentation-cue condition between the verbal and spatial tasks, then the best ANOVA model will not include an interaction of these factors. If the interaction is included, but reflects only a steeper dual-list cost for the spatial task, then the interaction term including the simpler two-level presentation-cue factor will be preferred. If the interaction is due to differences between the two dual-list conditions in the spatial task but not in the verbal task, then the best model will include an interaction between probe domain and the three-level presentation-cue factor.

The best ANOVA model included an interaction between probe domain and the two-level presentation-cue factor ($BF=3.57 \times 10^{65}$, ±2.07%), which differentiated between single- and dual-list conditions, but not between the cued and uncued dual-list conditions. Given the values in Figure 4, this interaction can only mean that dual-task costs were steeper in the spatial than in the verbal serial reconstruction task. However, this model was not drastically favored over a model excluding the interaction (BF = 1.14). The best model also was not favored very much over the more complex interaction model. Compared to the model including an interaction between probe domain and the more complex coding of the presentation-cue factor, the simpler interaction model was preferred only by a factor of less than 5, which is not compelling. These small Bayes Factor comparisons indicate that these data are insufficient to distinguish between these models, providing only tentative evidence for interpretation.

Discussion

These data supported the contention that concurrently encoding and maintaining two serial lists induces a cost, even when constructed from memoranda of differing domains. As noted in our introduction, this outcome is consistent with the idea that some processes are shared between verbal and visual-spatial materials, an assumption which each multi-component working memory model somehow incorporates. Our results were also consistent with the possibility that these dual-task costs are larger for spatial than for verbal memoranda, but inconclusive. Based on these data, we could not decisively distinguish between models that would support differences between cross-domain costs in the verbal and spatial tasks, nor between models differing in the precise effects of retro-cues on verbal and spatial maintenance. In both domains, irrelevant deviants (especially auditory ones) seemed to slightly impair performance in both tasks. Evidence for a small deviant effect was clear in the verbal serial reconstruction task and somewhat more likely than not in the spatial serial reconstruction task.

These results confirm much previous research suggesting that verbal and spatial memories interfere with each other. However, these results do not allow us to confidently go beyond that assertion and firmly limit these effects to *maintenance*, which a clear difference in the retro-cue effects between domains would allow. The discrepancy between performance with the verbal and spatial stimuli hampers our ability to make inferences. Though we detected dual-task effects in both tasks, verbal task performance was quite high, which restricts the potential to observe dual-task costs. If the verbal task was easy enough that participants did not need to devote much effort to maintaining the stimuli, then it could not be expected to exert much effect on concurrent maintenance of the spatial memoranda. Furthermore, interactions between task domain and dual-task condition may reflect effects unique to the different ranges of performance, which clouds interpretation of them. We could more confidently interpret any interactions (or the absence of them) between probe domain and presentation-cue effects if verbal and spatial reconstruction performance were in the same mediocre range. We rectify this in Experiment 2.

Experiment 2

Method

Participants. Twenty-six participants recruited from the University of Edinburgh community took part in the experiment, receiving either £7 per hour or partial course credit (16 females and 10 males, 19-39 years old, *M*= 21.73, *SD*= 3.78).

Apparatus, stimuli, and procedure. Instead of consonants, two-syllable nonwords were used as the verbal memoranda. A set of ten nonwords (*duplip, erdest, fluing, griple, hilder, lerman, mogleg, pliver, spavel, toplin*) was selected from the English Lexicon Project (Balota, et al., 2007). Recordings were generated in an artificial female voice and ranged from 400 to 555 ms in duration. Five nonwords were randomly selected without replacement on each trial. We reduced the number of trials compared to Experiment 1. Participants completed four blocks of 38 experimental trials each for a total of 152 trials. Half of these tested verbal and half spatial serial reconstruction, with 36 single-list trials per task, 20 trials per task each with informative and non-informative cues. Most trials did not include an irrelevant deviant;

irrelevant deviants occurred on 16 of the dual-list trials during the session (equally divided between trials testing verbal and spatial memory, and equally divided between occurrences to the verbal and spatial lists). Deviants occurred to only 4 each of the verbal and spatial single-list trials. All other elements of the procedure and analysis were identical to those described in Experiment 1.

Results

Verbal serial reconstruction. Mean proportions correct for verbal serial reconstruction organized by presentation-cue condition are shown in Figure 5. The best ANOVA model included main effects of serial position and presentation-cue condition (BF=1.13x10^{66}, ± 1.06%). This model exceeded a model including only serial position by a factor of more than 1 million, indicating the robustness of the effect of presentation-cue condition. Reconstruction was best in the single-list condition (M=0.56, SD=0.28), and did not appear to differ in the two dual-list conditions (M_{Cued}=0.45, SD=0.32; M_{Uncued}=0.45, SD=0.30). Exclusion of effects of deviant type (M_{None}=0.49, M_{Aural}=0.48, M_{Visual}=0.45) was favored by a factor of about 80. There was no evidence suggesting interactions between deviant type and any other factor (BFs > 460).

Spatial serial reconstruction. The same Bayesian ANOVA was performed on spatial serial reconstruction data (see Figure 6). The best ANOVA model included main effects of serial position, presentation-cue condition, and deviant type (BF=4.79×10^{71}, ±1.58%). Again, evidence for the effect of presentation-cue condition was overwhelming (BF> 1 million). Performance was best with single-list presentation (M=0.70, SD=0.25). Though accuracy in both dual-list conditions was lower, in contrast to the verbal task, reconstruction performance for the cued lists (M=0.52, SD=0.29) exceeded performance for the uncued lists (M=0.44, SD=0.28). This contention is supported by the outcome of the combined analysis below.

Occasional aural deviants presented in the verbal lists seemed to reduce accuracy (*M*=0.45, *SD*=0.29) relative to trials without deviants (*M*=0.55, *SD*=0.25) whereas occasional color deviants in the spatial lists did not impair performance (*M*=0.57, *SD*=0.32). Evidence for including the main effect of deviants was far weaker than that for effects of presentation-cue condition (*BF*=9.46), but too strong to dismiss.

Combined analysis. As in Experiment 1, we collapsed across serial position and deviant type in order to test our primary hypotheses regarding effects of retro-cueing. The presentation-cue manipulation appeared to affect spatial serial memory differently than verbal serial memory. Specifically, it appears that there was always some cost of presenting dual-lists, but cuing only benefited spatial memory. To test this formally, we entered the mean number of items correctly ordered per list into a common Bayesian ANOVA, including probe domain (verbal or spatial) and presentation-cue condition as factors, running versions that coded presentation-cue condition with three levels (single-list, and dual-list with or without a retro-cue) and with only two levels (single-list versus dual-list, ignoring the cue variable). Mean values are shown in Figure 7.

The best ANOVA model included an interaction between probe domain and the three-level presentation-cue factor, differentiating between dual-list presentations with and without informative cues (*BF*=5.85×10^{20}, ±2.49%). Compared to the model including the same effects but no interaction, the interaction model was favored by a factor of more than 87. Compared to the model including an interaction between probe domain and the simpler, 2-level coding of the presentation-cue factor, the more complex model was preferred by a factor of 18. Given the mean values for each combination of presentation-cue and probe domain level, the only viable explanation for this interaction is that all three levels differ in the spatial but not in the verbal task.

Combined analysis including data from Experiments 1 and 2. We entered data from Experiments 1 and 2 into the same Bayesian ANOVA described above, adding experiment as a grouping factor, in order to learn whether the crucial interaction between task domain and the 3-level presentation-cue factor would emerge when the inconclusive data from Experiment 1 were combined with the more decisive data of Experiment 2. The best model in this analysis indeed included the interaction between the 3-level presentation-cue condition and probe domain (BF=3.61x10^{83}, ±2.76%). This model was favored against the same model substituting the 2-level presentation-cue condition by a factor of about 8.

Discussion

With performance in the verbal and spatial serial reconstruction tasks in a similar range, we observed clear evidence supporting the hypothesis that concurrently maintaining verbal memoranda interferes with maintenance of spatial memoranda, over and above any interference attributable to dividing attention between encoding the verbal and spatial lists simultaneously. Verbal and spatial memoranda both suffered from simultaneous encoding: this is evident from the clear differences between single-list and dual-list conditions in both tasks. However, we observed no evidence for any further decrease in verbal memory in the uncued compared to the retro-cued condition, indicating that no further interference occurred due specifically to maintaining the spatial memoranda.

General Discussion

The results of both of our experiments unambiguously confirm previous evidence that verbal and visual-spatial serial memory tasks interfere with each other (Depoorter & Vandierendonck, 2009; Morey & Mall, 2012; Vandierendonck, 2015). Moreover, our results provide further detail that is needed to advance theory. We observed substantial decreases in serial reconstruction accuracy to both verbal and spatial serial memory under dualpresentation conditions, indicating that dividing attention between encoding two lists is problematic even when stimuli are presented via different senses. This interference was stronger for spatial than for verbal memoranda, but differing levels of performance across tasks limited our ability to draw conclusive conclusions in Experiment 1, where verbal memoranda were consonant lists. Experiment 2 replicated the design of Experiment 1 except that verbal memoranda were non-words, which produced performance in a range comparable to the spatial task performance. Here, spatial but not verbal reconstruction decreased even further without an informative retro-cue. Thus maintenance of spatial, but not verbal, stimuli remains vulnerable to general interference even when interference can only be attributed to concurrent maintenance activities.

While these results imply that there are differences between maintaining verbal and visual-spatial serial memories that models ought to account for, they do not support models that posit distinct, specialized memory stores or rehearsal mechanisms for handling visual materials. We think that the preponderance of evidence showing interference between verbal and visual-spatial memory in a variety of experimental designs shows that the notion of distinct modules for temporarily maintaining visual-spatial information no longer merits serious consideration. However, it is likewise clear that at acquisition aural-verbal and visual-spatial information are routed through different perceptual systems, and that information acquired from different senses is not equivalent. These recurring disparities, particularly the superior resistance of verbal information to general interference, have likely led to exaggerated claims about the domain-specificity of working memory representations. Models should account somehow for asymmetries that appear to protect verbal information from cross-domain interference not by insisting on segregation of verbal and non-verbal storage, but by focusing on the unique benefits verbalization or auditory perception might afford.

Several proposals that attempt to account for this should be directly compared. Possibilities that handle this pattern of data without resorting to verbal and visuo-spatial storage modules include the supposition of domain-specific verbal rehearsal with no specialized non-verbal counterpart (e.g., Barrouillet & Camos, 2015), the idea that some temporal features of aural lists are inherited from auditory perception (e.g., Conway, Pisoni, & Kronenberger, 2009; Jones, Hughes, & Macken, 2006) and propagate into memory, and the possibility of domain-specific contributions to general working memory coming from sensory memory systems (e.g., Cowan, 1988, Saults & Cowan, 2007) which are not necessarily equivalent.

According to the latest version of the TBRS model (Barrouillet & Camos, 2015), our findings occur because verbal lists may be maintained by a specialized rehearsal mechanism but spatial lists must be maintained by a general attention resource. If there are two resources to devote to two stimulus sets, why was interference during maintenance observed at all? It should not be assumed that rehearsal of a verbal sequence can be carried out entirely without general attention. Naveh-Benjamin and Jonides (1984) showed that initiating the rehearsal of a verbal sequence requires attention, and that this process only becomes effortless after some iterations of the sequence. Applied to our procedure, the decrease in spatial reconstruction performance would have occurred due to the need to initiate a verbal rehearsal sequence, not to perpetuate it. Combining TBRS with the maintenance model of Naveh-Benjamin and Jonides suggests that to design a dual-task procedure with minimal interference, one must avoid combining a second task with the effortful phase of initiating a verbal rehearsal sequence.

We also described exploratory manipulations of irrelevant aural and visual deviants on our verbal and spatial serial reconstruction data. The effects of this manipulation differed somewhat in Experiment 1, where verbal reconstruction performance was on average highly accurate, and Experiment 2, where reconstruction performance on both tasks was mediocre. Nonetheless, there were some consistencies. In both experiments and tasks, hearing an aural deviant was more likely to impair memory. For the spatial memory task, this effect appeared in both experiments, though evidence favoring it never exceeded a Bayes factor of 10. This evidence tentatively supports the contention that spatial memory is susceptible to interference from a variety of sources (e.g., Morey & Bieler, 2013; Stevanovski & Jolicoeur, 2007), not only domain-specific ones. Verbal task performance was affected by deviants, mainly aural ones, in Experiment 1 where verbal memory performance was quite high, but not in Experiment 2, where performance was much lower. This pattern is logically consistent with perceptual load theory (Lavie, 2010) extended to memoranda: with a more demanding verbal memory task, irrelevant sounds were more likely to be successfully ignored than with an easier task, in which participants may have had spare capacity to allow for the processing of the deviants. In the spatial task, which was arguably challenging in both experiments, we did not observe clear effects of visual deviants. This is consistent with previous findings that a visual memory load induces attentional blindness to stimulus-driven visual attention capture (e.g., Konstantinou, Beal, King, & Lavie, 2014; Todd, Fougnie, & Marois, 2005).

In conclusion, the differential effects of retro-cues on verbal and spatial memory tasks are incompatible with multi-component models of working memory that posit parallel storage and rehearsal modules for verbal and visual-spatial materials. Instead, comparing alternative frameworks for explaining why verbal memories resist interference better than visual ones should become a priority.

References

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

Barrouillet, P., & Camos, V. (2015). *Working memory: Loss and reconstruction*. Hove, U.K.: Psychology Press.

Bayliss, D. M., Jarrold, C., Gunn, D. M., & Baddeley, A. D. (2003). The complexities of complex span: Explaining individual differences in working memory in children and adults. *Journal of Experimental Psychology: General*, *132*, 71–92.

Camos, V., Lagner, P., & Barrouillet, P. (2009). Two maintenance mechanisms of verbal information in working memory. *Journal of Memory and Language*, *61*, 457–469.

Camos, V., Mora, G., & Oberauer, K. (2011). Adaptive choice between articulatory rehearsal and attentional refreshing in verbal working memory. *Memory & Cognition*, *39*, 231–244.

Cherry, E. (1953). Some experiments on the recognition of speech, with one and with two ears. *Journal of the Acoustical Society of America*, *25*, 975–979.

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.

Conway, C. M., Pisoni, D. B., & Kronenberger, W. G. (2009). The importance of sound for cognitive sequencing abilities: The auditory scaffolding hypothesis. *Current Directions in Psychological Science*, *18*, 275–279.

Cortis, C., Dent, K., Kennett, S., & Ward, G. (2015). First things first: Similar list length and output order effects for verbal and nonverbal stimuli. *Journal of Experimental Psychology-Learning Memory and Cognition*, *41*, 1179–1214.

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

Depoorter, A., & Vandierendonck, A. (2009). Evidence for modality-independent order coding in working memory. *The Quarterly Journal of Experimental Psychology*, *62*, 531–549.

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.

Griffin, I. C., & Nobre, A. C. (2003). Orienting attention to locations in internal representations. *Journal of Cognitive Neuroscience*, *15*, 1176–1194.

Guérard, K., & Tremblay, S. (2008). Revisiting evidence for modularity and functional equivalence across verbal and spatial domains in memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*, 556–569.

Hughes, R. W., Vachon, F., & Jones, D. M. (2007). Disruption of short-term memory by changing and deviant sounds: Support for a duplex-mechanism account of auditory distraction. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*, 1050–1061.

Jeffreys, S. H. (1961). *The Theory of Probability* (3rd ed.). Oxford, U.K.: Oxford University Press.

Jones, D., Farrand, P., Stuart, G., & Morris, N. (1995). Functional equivalence of verbal and spatial information in serial short-term memory. *Journal of Experimental Psychology:*

Learning, Memory, and Cognition, *21*, 1008–1018.

Jones, D. M., Hughes, R. W., & Macken, W. J. (2006). Perceptual organization masquerading as phonological storage: Further support for a perceptual-gestural view of short-term memory. *Journal of Memory and Language*, *54*, 265–281.

Kass, R. E., & Raftery, A. E. (1995). Bayes Factors. *Journal of the American Statistical Association*, *90*(430), 773–795.

Konstantinou, N., Beal, E., King, J.-R., & Lavie, N. (2014). Working memory load and distraction: dissociable effects of visual maintenance and cognitive control. *Attention Perception & Psychophysics*, *76*, 1985–1997.

Lavie, N. (2010). Attention, distraction, and cognitive control under load. *Current Directions in Psychological Science*, *19*, 143–148.

Logie, R. H. (2011). The functional organization and capacity limits of working memory. *Current Directions in Psychological Science*, *20*, 240–245.

Logie, R. H., Della Sala, S., Wynn, V., & Baddeley, A. D. (2000). Visual similarity effects in immediate verbal serial recall. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology*, *53*, 626–646.

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

Meiser, T., & Klauer, K. C. (1999). Working memory and changing-state hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*, 1272–1299.

Morey, C. C., & Bieler, M. (2013). Visual short-term memory always requires general attention. *Psychonomic Bulletin & Review*, *20*, 163–170.

Morey, C. C., & Mall, J. T. (2012). Cross-domain interference costs during concurrent verbal and spatial serial memory tasks are asymmetric. *Quarterly Journal of Experimental*

Psychology, *65*, 1777–1797.

Morey, C., Morey, R., van der Reijden, M., & Holweg, M. (2013). Asymmetric cross-domain interference between two working memory tasks: Implications for models of working memory. *Journal of Memory and Language*, *69*, 324–348.

Morey, R. D. (2008). Confidence intervals from normalized data: A correction to Cousineau (2005). *Tutorial for Quantitative Methods in Psychology*, *4*, 61–64.

Nairne, J. S. (2002). Remembering over the short-term: the case against the standard model. *Annual Review of Psychology*, *53*, 53–81.

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

Oberauer, K. (2013). The focus of attention in working memory-from metaphors to mechanisms. *Frontiers in Human Neuroscience*, *7*, 673.

Parmentier, F. B. R., Elsley, J. V., & Ljungberg, J. K. (2010). Behavioral distraction by auditory novelty is not only about novelty: The role of the distracter's informational value. *Cognition*, *115*, 504–511.

Parmentier, F. B. R., Tremblay, S., & Jones, D. M. (2004). Exploring the suffix effect in serial visuospatial short-term memory. *Psychonomic Bulletin & Review*, *11*, 289–295.

Raveh, D., & Lavie, N. (2015). Load-induced inattentional deafness. *Attention Perception & Psychophysics*, *77*, 483–492.

Rerko, L., Souza, A. S., & Oberauer, K. (2014). Retro-cue benefits in working memory without sustained focal attention. *Memory & Cognition*, *42*, 712–728.

Rouder, J. N., Morey, R. D., Speckman, P. L., & Province, J. M. (2012). Default Bayes factors for ANOVA designs. *Journal of Mathematical Psychology*, *56*, 356–374.

Saito, S., Logie, R. H., Morita, A., & Law, A. (2008). Visual and phonological similarity effects in verbal immediate serial recall: A test with kanji materials. *Journal of Memory and Language*, *59*, 1–17.

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, PA: Psychology Software Tools, Inc.

Schomaker, J., & Meeter, M. (2014). Facilitation of responses by task-irrelevant complex deviant stimuli. *Acta Psychologica*, *148*, 74–80.

Shah, P., & Miyake, A. (1996). The separability of working memory resources for spatial thinking and language processing: An individual differences approach. *Journal of Experimental Psychology: General*, *125*, 4–27.

Shimi, A., Nobre, A. C., Astle, D., & Scerif, G. (2014). Orienting attention within visual short-term memory: development and mechanisms. *Child Development*, *85*, 578–592.

Smyth, M. M., Hay, D. C., Hitch, G. J., & Horton, N. J. (2005). Serial position memory in the visual-spatial domain: Reconstructing sequences of unfamiliar faces. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology*, *58A*, 909–930.

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

Todd, J. J., Fougnie, D., & Marois, R. (2005). Visual short-term memory load suppresses temporo-parietal junction activity and induces inattentional blindness. *Psychological Science*, *16*, 965–972.

Vandierendonck, A. (2015). Modality independence of order coding in working memory:

Evidence from cross-modal order interference at recall. *The Quarterly Journal of Experimental Psychology*, 1–19. http://doi.org/10.1080/17470218.2015.1032987

Vergauwe, E., Barrouillet, P., & Camos, V. (2009). Visual and spatial working memory are not that dissociated after all: A time-based resource-sharing account. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*, 1012–1028.

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

Vergauwe, E., Dewaele, N., Langerock, N., & Barrouillet, P. (2012). Evidence for a central pool of general resources in working memory. *Journal of Cognitive Psychology*, *24*, 359–366.

Ward, G., Avons, S. E., & Melling, L. (2005). Serial position curves in short-term memory: Functional equivalence across modalities. *Memory*, *13*, 308–317.

Figure 1

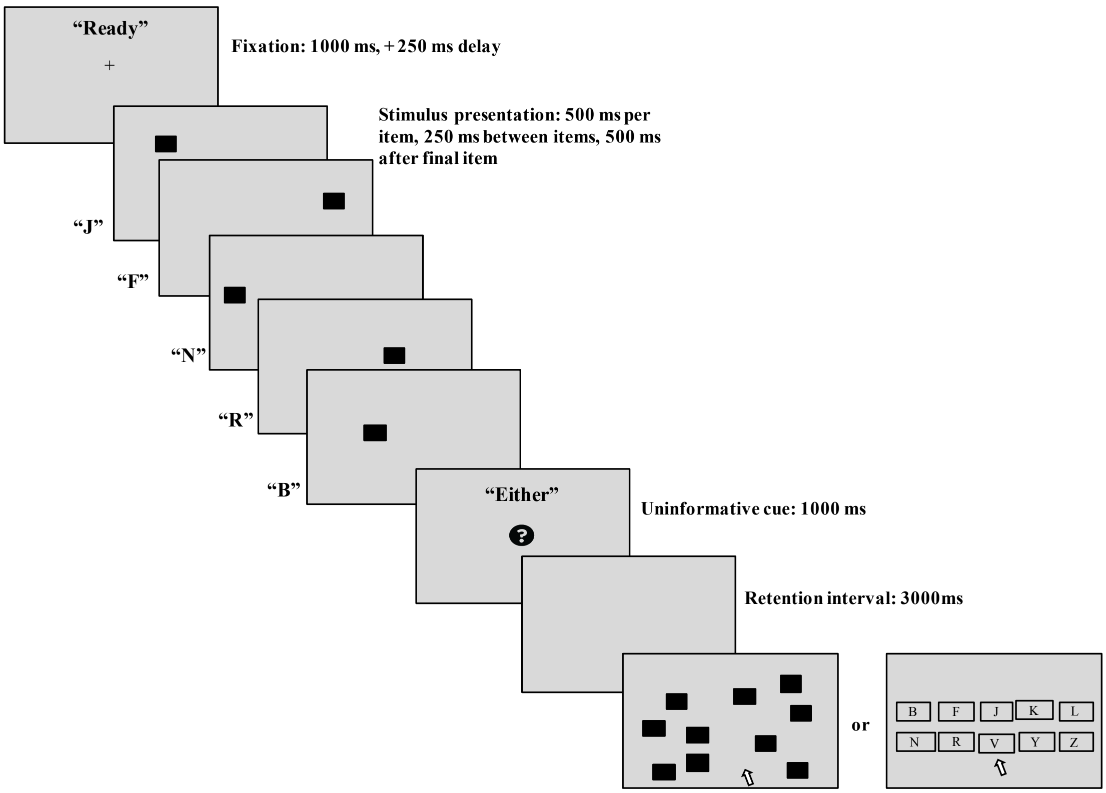

Figure 1. Example of a trial procedure, in the uncued condition. Text in quotes indicates aural presentation. At test, either the spatial reconstruction or the verbal reconstruction prompt would have been presented. In retro-cued conditions, the uninformative cue in the example was replaced by an informative cue indicating which stimulus set would be tested. Note that the images are not depicted to scale.

Figure 2

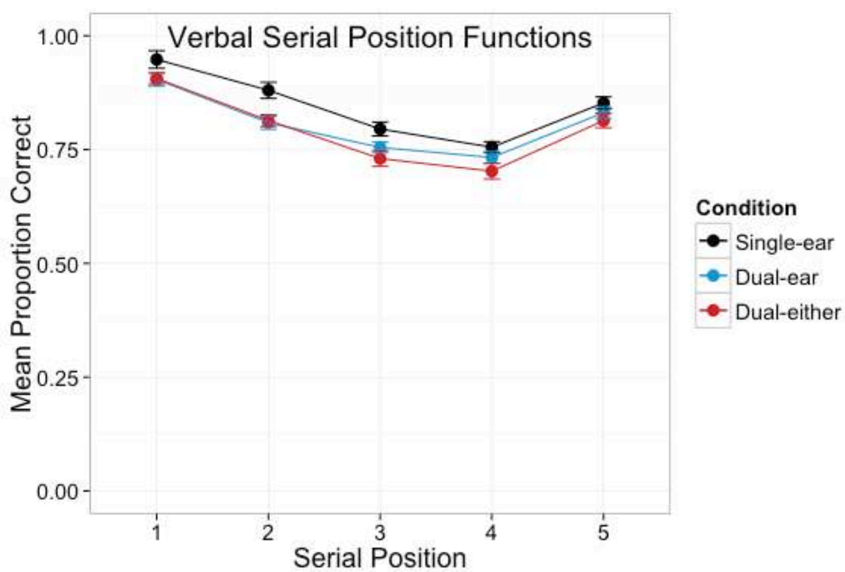

Figure 2. Mean proportions correct on verbal serial reconstruction as a function of serial position, and presentation-cue condition, Experiment 1. Error bars are standard errors of the mean with the Morey-Cousineau (2008) correction applied. *N*=24.

Figure 3

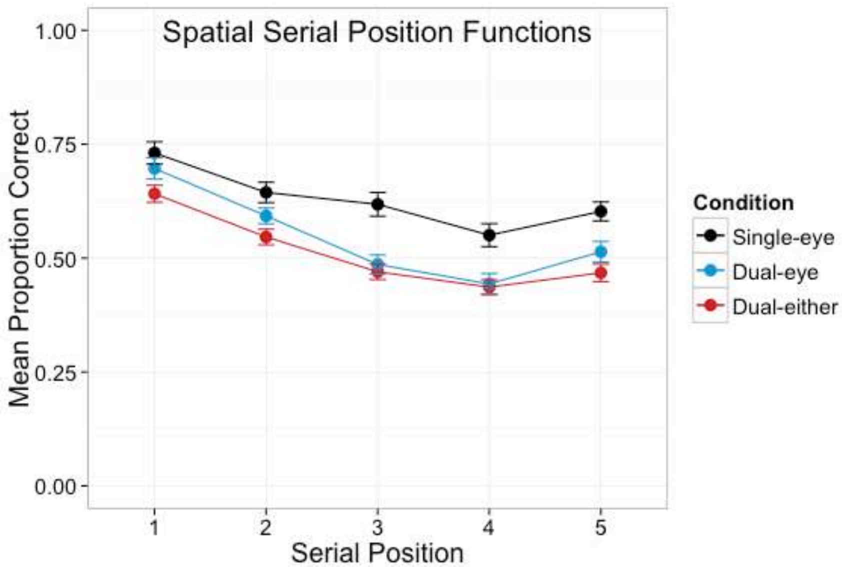

Figure 3. Mean proportions correct on spatial serial reconstruction as a function of serial position and presentation-cue condition, Experiment 1. Error bars are standard errors of the mean with the Morey-Cousineau (2008) correction applied. *N*=24.

Figure 4

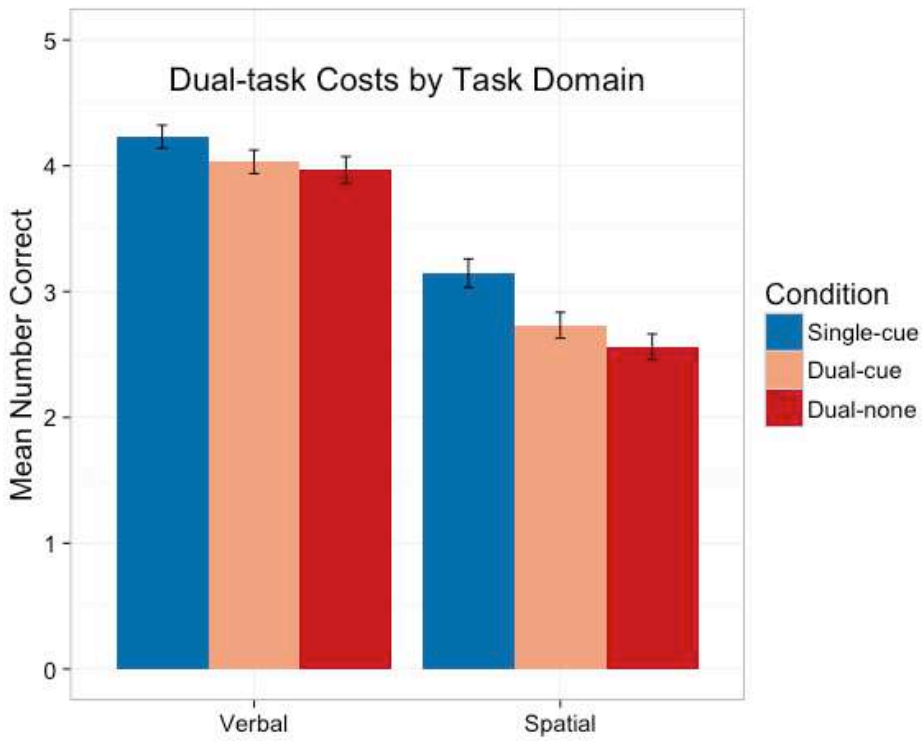

Figure 4. Mean number correct (out of 5 possible per trial) organized by task domain and presentation-cue condition, Experiment 1. Error bars are standard errors of the mean with the Morey-Cousineau (2008) correction applied.

Figure 5

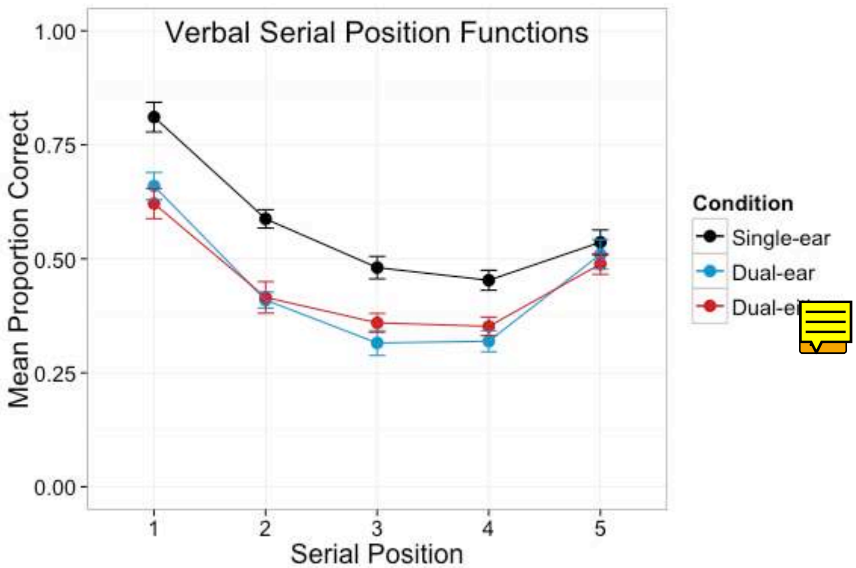

Figure 5. Mean proportions correct on verbal serial reconstruction as a function of serial position and presentation-cue condition, Experiment 2. Error bars are standard errors of the mean with the Morey-Cousineau (2008) correction applied. *N*=26.

Figure 6

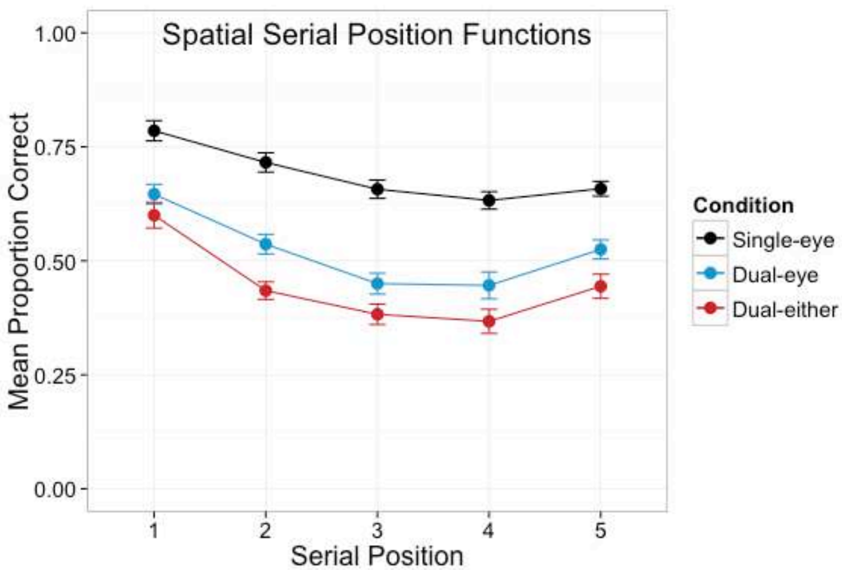

Figure 6. Mean proportions correct on spatial serial reconstruction as a function of serial position and presentation-cue condition, Experiment 2. Error bars are standard errors of the mean with the Morey-Cousineau (2008) correction applied. *N*=26.

Figure 7

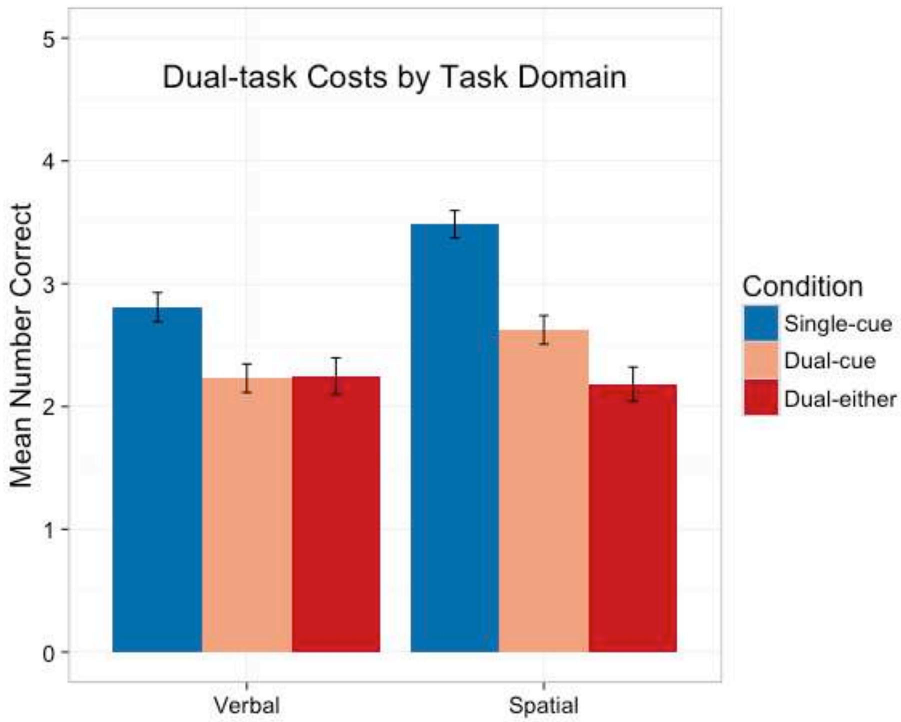

Figure 7. Mean number correct (out of 5 possible per trial) organized by task domain and presentation-cue condition, Experiment 2. Error bars are standard errors of the mean with the Morey-Cousineau (2008) correction applied.