

Visual working memory is disrupted by covert verbal retrieval

TIMOTHY J. RICKER AND NELSON COWAN
University of Missouri, Columbia, Missouri

AND

CANDICE C. MOREY
University of Groningen, Groningen, The Netherlands

If working memory (WM) depends on a central resource—as is posited in some theories, but not in others—it should be possible to observe interference between tasks that share few features with each other. We investigated whether interference with WM for visual arrays would occur, even if the interfering task required neither visual processing nor overt responding. In an auditory–verbal interfering task, a response was to be made if a word was recognized as having come from one of two prelearned lists, but not from the other list. As compared with nonretrieval control conditions, even covert verbal memory retrieval disrupted the storage of visual items held in WM. A second experiment ruled out verbal recoding of the visual arrays as the basis of interference. The results indicate that visual WM and verbal long-term retrieval share a central resource (e.g., attention).

Many theorists (e.g., Cocchini, Logie, Della Sala, MacPherson, & Baddeley, 2002; Friedman & Miyake, 2000) assume that working memory (WM) is partitioned into several code-specific resources (e.g., Baddeley, 1986). The belief in separate resources for different types of information is based largely on evidence that verbal WM tasks conflict more with other verbal tasks than they do with visual tasks, and vice versa. An important open question, though, is whether a central component of WM accepts information regardless of its original coding. Evidence for that possibility includes conflicts between stimulus sets in WM that are visual–spatial versus auditory–verbal (Morey & Cowan, 2004, 2005) or visual–spatial versus nonverbal–acoustic (Stevanovski & Jolicœur, 2007), and there is psychometric evidence of domain-general storage as well (Kane et al., 2004).

We address one key question about these demonstrations of central WM storage: an ambiguity in the source of dual-task interference. Some theories of information processing distinguish between the attention used for central processing, such as perception, WM storage, and memory retrieval, versus the bottleneck encountered in planning and executing overt responses (Pashler, 1993). These have been difficult to dissociate because, in many situations, overt responding is necessary to ensure that central processes are used. For example, Morey and Cowan (2005) found that there was a much larger effect of a verbal memory load on visual array performance when the verbal load was recited aloud. Researchers (e.g., Barrouillet, Bernardin, Portrat, Vergauwe, & Camos, 2007) often depend on

overt responses to distracting stimuli to enhance and assess distraction. Previous studies indicate an attention cost of verbal retrieval (for a review, see Baddeley, 1986), but its effect on visual WM has not been shown.

To confirm that a central capacity limit exists in WM, it is necessary to test the effect of a central process in the absence of overt responding. To determine whether intertask interference would still occur, we used a visual array memory task (Luck & Vogel, 1997) along with an acoustic–verbal secondary task that sometimes required memory retrieval but no overt response. In Experiment 1, the design included variation in the amount of retrieval required by the secondary task and in the need to respond overtly in that task.

EXPERIMENT 1

Instead of using visual stimuli in easily labeled categories, we reduced labeling by using unconventional characters (cf. Alvarez & Cavanagh, 2004). Figure 1 illustrates a typical trial. An array of one or three unconventional, briefly shown characters was to be remembered. (One-character arrays were presented only to demonstrate that the characters could be encoded and retained.) The array was followed by a mask to ensure that what was remembered was not simply sensory in nature (cf. Sauls & Cowan, 2007). The mask occurred only after ample time for encoding (Vogel, Woodman, & Luck, 2006). At the end of the trial, a single probe character was to be judged as being either the same as or different from the array charac-

T. J. Ricker, tjr899@mail.mizzou.edu

Each word within this set had a familiarity rating between 545 and 700 and an imageability rating between 550 and 700. The maximum difference in volume between the sound files used to present the auditory stimuli was 1.5 dB(A) at peak.

Procedure. The experiment started with 8 practice visual array trials followed by memorization of two separate lists of four spoken words that would be used throughout the experiment. Participants were then given instructions on how to complete the auditory word task, along with examples of each auditory task condition and the correct response. Ten practice spoken-word trials were then performed. Finally, 10 dual-task practice trials were administered, each with a structure as illustrated in Figure 1, to end the practice portion of the experiment. In each block of 40 trials, one of the two lists was designated as the target list from which a presented word required a keypress response. Within each experimental block, there were an equal number of trials of each auditory task condition. In total, eight experimental blocks were completed (four for each set size, in random order).

The primary visual array task began with the presentation of a fixation cross on the screen for 1,000 msec (Figure 1). This was followed by the presentation of an array of one or three unfamiliar visual characters. During the practice trials, three English letters were presented, instead of unfamiliar characters. The array appeared for 500 msec and was followed by a blank-screen delay of 500 msec. A visual mask was then presented for 100 msec, followed by a 3,000-msec retention interval, during which the secondary auditory task was performed. After the retention interval, a single probe item was presented. It was either the same character as that which had appeared at the same location in the original array or a new character. The “S” and “D” keys were used for *same* and *different*. Within each block, the correct response was *same* on half of the trials.

For the secondary auditory task, at the beginning of the experiment, participants memorized two lists comprising 4 words each, which were introduced to the participants as List 1 and List 2. Both lists were determined randomly for each participant and were chosen from a larger pool of 44 words. The remaining 36 nonlist words were used throughout the experiment as filler items. A word recognition test was repeated until the participant could recognize from which list, if any, each of the words had come. Then, in the experiment proper, all words, including the list words, appeared equally often (twice per block, never on the same trial).

During each block of trials, one of the lists was specified as the target list for use in the auditory task. In this task, participants heard three words spoken through headphones at a pace of two words per second. At the onset of the third word, the screen turned green for 2,000 msec to signify that a response could be entered. It was the participant’s job to press the space bar during the green-screen response period if they heard a pure tone in place of any of the three

words or if they heard a word from the target list. Of the five types of auditory task conditions (blank, tone, nonlist, nontarget, target), the second and fifth conditions required a response.

Results

Primary visual array task. For one-item arrays, performance accuracy was .96 in the tone and nontarget conditions and .97 in the other three conditions; in each case, $SEM = .01$. This near-perfect performance demonstrates that the unfamiliar characters could be encoded with enough precision to allow a *same* versus *different* response. Mean visual array accuracies for Set Size 3, the set size of primary interest, were lower, as we have illustrated in Figure 2.

A one-factor within-participants ANOVA of visual array accuracy for three-item arrays was conducted with auditory task condition as a factor. There was an effect of auditory task condition [$F(4,156) = 9.06$, $MS_e = .01$, $p < .0001$, $\eta_p^2 = .19$]. Newman-Keuls post hoc tests revealed significant differences between each pair of conditions, except between the blank and tone, tone and nonlist, and nontarget and target conditions. All comparisons indicating the importance of memory retrieval were significant: In the absence of responding, blank > nonlist > nontarget, and, in its presence, tone > target. None of the comparisons indicating the importance of overt responding reached significance.

Secondary auditory word task. Mean accuracies and reaction times (RTs) for all conditions are presented in Table 1.¹ A 2 (set size) \times 5 (auditory task condition) within-participants ANOVA of auditory-verbal task accuracy indicated no effect of visual array set size [$F(1,39) < 1$], but there was a main effect of auditory task condition [$F(4,156) = 50.83$, $MS_e = .01$, $p < .0001$, $\eta_p^2 = .57$]. (Means are in Table 1; $SEMs$ ranged from .01 to .03 for the five conditions.) The interaction was not significant [$F(4,156) < 1$]. Newman-Keuls follow-up tests of auditory task condition means revealed no significant differences between the blank, tone, and nonlist trials, but, for all other comparisons, the group means were significantly different.

A 2×2 within-participants ANOVA of auditory task RTs was conducted with visual array set size and auditory

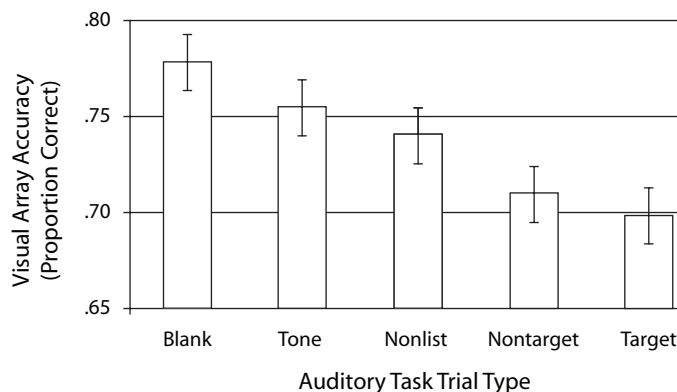


Figure 2. Mean visual array accuracy in Experiment 1 for all conditions at Set Size 3. Error bars represent $\pm 1 SEM$.

Table 1
Mean Auditory–Verbal Task Accuracies and Reaction Times (RTs, in Milliseconds)
for Each Visual Array Set Size and Auditory Task Condition

Visual Array Set Size	Auditory Task Condition						
	Blank		Tone		Nonlist	Target	
	Accuracy	Accuracy	RT	Accuracy	Accuracy	Accuracy	RT
1	.99	.99	678	.97	.92	.83	906
3	.99	.98	671	.96	.91	.84	927

Note—Standard errors are reported in the text.

task condition as factors. Only the tone and target conditions required responses, and only trials with a correct response were analyzed. There was a main effect of auditory task condition [$F(1,39) = 189.09$, $MS_e = 12,376$, $p < .0001$, $\eta_p^2 = .83$]. (Means are in Table 1; SEM was 31 msec for tones and 30 msec for target words.) There was, however, no effect of visual array set size [$F(1,39) < 1$] and no significant interaction [$F(1,39) < 1$].

Relation between tasks. There was almost no correspondence between the two tasks, either within or across individuals. One exception was a very small but significant effect of whether performance on three-item arrays was correct on accuracy in the secondary auditory task [arrays correct, $M = .94$; arrays incorrect, $M = .93$; $F(1,39) = 5.52$, $MS_e = .01$, $p < .05$, $\eta_p^2 = .12$], which did not interact with auditory task condition.

Discussion

The effect of covert verbal memory retrieval on visual array memory was observed in two different contrasts in Experiment 1. First, visual array memory was better when the retention interval was blank than when it was filled with three filler words (in the nonlist condition). This points to the potential importance of a general sort of search of long-term memory in identifying list items. Second, visual array memory was better when the spoken verbal set included no item from a learned list (in the nonlist condition) than when it did contain such an item, even when the item came from the list that did not require a response (in the nontarget condition). This finding points to the importance of the retrieval of information from a recent learning episode in interfering with visual array memory.

Response execution did not produce any significant interference with our visual array task. Together, these contrasts provide evidence for a domain-general form of memory storage that is not engaged solely to produce responses.

The interference we obtained was asymmetrical, in that secondary verbal task performance did not depend on the visual array condition; a load of three visual characters produced verbal task performance no worse than a load of only one character. This asymmetry is to be expected; other studies have shown that retrieval uses attention, but that, once the retrieval process has started, its completion is obligatory (e.g., Naveh-Benjamin, Craik, Gavrilescu, & Anderson, 2000). Thus, in the case of our study, the same amount of attention is presumably used for retrieval in the presence of a visual load of one or three items.

An estimate of the number of units in WM (Cowan, 2001) is $k = S[2 * (h - f)]$, where k is the number of items in WM, S is the set size, h is the proportion of correct change detection, and f is the proportion of spurious change detection. We find estimates that are low, as compared with simple stimuli (which typically are about 3.5 items; Cowan, 2001; Rouder et al., 2008), but are in line with prior results for unconventional stimuli (Alvarez & Cavanagh, 2004). The formula, however, underestimates capacity if each presented item is encoded as multiple chunks, as many of our stimuli might well be. Suppose, for example, that the mean capacity in the blank condition is 3.5 chunks. That capacity implies that we must multiply the average k value obtained from the blank condition (1.67 items) by a factor of 2.10 (chunks/item) to yield 3.5 chunks. In comparison, using that same multiplication factor for the other conditions yields chunk estimates of 3.21 for tone, 3.03 for nonlist, 2.64 for nontarget, and 2.50 for target ($SEMs = .18-.19$). According to this analysis comparing blank and nontarget conditions, memory retrieval without responding costs an estimated 0.86 chunks in visual WM.

EXPERIMENT 2

An alternative to the notion that visual array memory suffers interference from covert verbal retrieval is that the visual array could have been verbally recoded, wholly or in part, despite our use of unconventional characters. In Experiment 2, we examined this assumption.

The same three-item visual arrays were used as in Experiment 1, but with different secondary task conditions. Participants performed the array task (1) alone, (2) while tapping a finger at a pace of two times per second, or (3) while carrying out an articulatory suppression task, counting forward from 1 at a pace of two digits per second. Tapping is a common nonverbal control condition that consumes some attention (as suppression does) but does not prevent verbal labeling or rehearsal (again, as suppression does; Bunting, Cowan, & Colflesh, 2008). If performance suffers more from suppression than from tapping, that will indicate that memory for our unconventional character arrays relies, in part, on verbal processes, such as verbal recoding of the visual characters or verbal rehearsal.

Method

Participants. The participants were 25 students (16 female; age range = 18–21 years) who received partial course credit. Data from

1 participant were not analyzed, due to below-chance performance in one condition, leaving a final N of 24.

Design. The experiment consisted of a visual array task with three items. The participants were engaged in each of three secondary task conditions (no task, tapping, suppression) throughout the experiment. This yielded a single-factor (secondary task) within-participants design.

Apparatus and Materials. All of the materials in Experiment 2 were the same as those used in Experiment 1.

Procedure. The basic procedure was the same as in the three-item array, blank condition of Experiment 1. The participants completed six experimental blocks of 30 trials: two blocks for each of the three secondary task conditions. The two blocks of each condition were performed consecutively, beginning with a 10-trial practice block for each secondary task. During these practice blocks, for the suppression and tapping conditions, the participants were trained to pace their secondary task responses at a rate of one every 500 msec, beginning at the onset of the fixation cross and ending at the onset of the probe item. Compliance was monitored by an experimenter, who offered feedback to correct pacing when necessary.

Results

A one-factor within-participants ANOVA of visual array accuracy was conducted with secondary task condition as a factor. There was an effect of secondary task condition [$F(2,46) = 4.67$, $MS_e = .00231$, $p = .01$, $\eta_p^2 = .17$]. For the no-task condition, $M = .78$ and $SEM = .017$; for the tapping condition, $M = .74$ and $SEM = .016$; and for the articulatory suppression condition, $M = .74$ and $SEM = .018$. Newman-Keuls post hoc tests revealed a significant difference between the no-task condition and both the tapping ($p = .02$) and suppression ($p = .02$) conditions. Tapping and suppression did not differ.

Discussion

Performing a secondary paced-responding task impaired performance on the same visual array memory task used in Experiment 1. Critically, articulatory suppression did not impair performance more than simple tapping performed at the same pace did. This strongly suggests that any labeling or rehearsal of the unconventional characters used in both experiments is of little or no mnemonic value and cannot explain the interference from verbal retrieval seen in Experiment 1.

GENERAL DISCUSSION

The selection of a model of WM depends critically on the outcome of dual-task experiments. At issue is whether there is one component of WM storage that can be said to be central and general rather than modality- or code-specific (e.g., Cocchini et al., 2002; Morey & Cowan, 2005). Recent results have pointed toward the existence of a central storage component (e.g., Kane et al., 2004; Sauls & Cowan, 2007; Stevanovski & Jolicœur, 2007). The present results strengthen this conclusion by showing that interference with visual array performance from a secondary auditory-verbal task can occur even when (1) the visual stimuli are difficult to label and (2) the auditory-verbal interference requires memory retrieval (but not overt responding) during the visual array maintenance period. Many mechanisms of this effect are pos-

sible; for example, verbal retrieval could sometimes cause neglect of the goal of retaining the visual array.

The results suggest a WM system that relies on a common resource, such as attention, to maintain visual items and retrieve verbal items from long-term memory, while arguing against a more modular WM system positing only multiple modality- or code-specific storage components (Baddeley, 1986; Cocchini et al., 2002; Friedman & Miyake, 2000). Nevertheless, the fact that our quite reliable effects are numerically modest suggests a compromise model of WM that includes code-specific storage along with a central memory store. Such a model was, in fact, proposed in the seminal work of Baddeley and Hitch (1974), but the central component was omitted in subsequent work (e.g., Baddeley, 1986; Baddeley & Logie, 1999). Our findings are consistent with a revised model that includes an episodic buffer requiring attention (e.g., Repovš & Baddeley, 2006) or with a model positing feature-based activation and interference, in addition to one or more items in a limited-capacity focus of attention (Barrouillet et al., 2007; Cowan, 1999). Finally, the finding that individual differences in WM are related to both active maintenance in primary memory and controlled search from secondary memory (Unsworth & Engle, 2007) can now be explained on the grounds that both of these factors depend on the same attentional resource that was tapped in the present study by both visual array maintenance and verbal list memory retrieval.

AUTHOR NOTE

We acknowledge funding from NIH Grant R01-HD21338. Thanks to Robert Logie for suggesting unconventional stimuli and to Michael Williams for assistance. Correspondence concerning this article should be addressed to T. J. Ricker, Department of Psychological Sciences, University of Missouri, Columbia, MO 65211 (e-mail: tjr899@mail.mizzou.edu).

REFERENCES

- ALVAREZ, G. A., & CAVANAGH, 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.
- BADDELEY, A. [D.] (1986). *Working memory*. Oxford: Oxford University Press.
- BADDELEY, A. D., & HITCH, G. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 8, pp. 47-89). New York: Academic Press.
- BADDELEY, A. D., & LOGIE, R. H. (1999). Working memory: The multiple-component model. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 28-61). Cambridge: Cambridge University Press.
- BARROUILLET, P., BERNARDIN, S., PORTRAT, S., VERGAUWE, E., & CAMOS, V. (2007). Time and cognitive load in working memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *33*, 570-585.
- BUNTING, M. F., COWAN, N., & COLFLESH, G. H. (2008). The deployment of attention in short-term memory tasks: Trade-offs between immediate and delayed deployment. *Memory & Cognition*, *36*, 799-812.
- COCCHINI, G., LOGIE, R. H., DELLA SALA, S., 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. (1999). An embedded-processes model of working memory. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 62-101). Cambridge: Cambridge University Press.

- COWAN, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral & Brain Sciences*, **24**, 87-114.
- FEARNLEY, S. (1997). MRC Psycholinguistic Database search program. *Behavior Research Methods, Instruments, & Computers*, **29**, 291-295.
- FRIEDMAN, N. P., & MIYAKE, A. (2000). Differential roles for visuospatial and verbal working memory in situation model construction. *Journal of Experimental Psychology: General*, **129**, 61-83.
- KANE, M. J., HAMBRICK, D. Z., TUHOLSKI, S. W., WILHELM, O., PAYNE, T. W., & ENGLE, R. W. (2004). The generality of working memory capacity: A latent-variable approach to verbal and visuospatial memory span and reasoning. *Journal of Experimental Psychology: General*, **133**, 189-217.
- LUCK, S. J., & VOGEL, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, **390**, 279-281.
- 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, & Cognition*, **31**, 703-713.
- NAVEH-BENJAMIN, M., CRAIK, F. I. M., GAVRILESCU, D., & ANDERSON, N. D. (2000). Asymmetry between encoding and retrieval processes: Evidence from divided attention and a calibration analysis. *Memory & Cognition*, **28**, 965-976.
- PASHLER, H. (1993). Dual-task interference and elementary mental mechanisms. In D. E. Meyer & S. Kornblum (Eds.), *Attention and performance XIV: Synergies in experimental psychology, artificial intelligence, and cognitive neuroscience* (pp. 245-264). Cambridge, MA: MIT Press.
- REPOVŠ, G., & BADDELEY, A. [D.] (2006). Multi-component model of working memory: Explorations in experimental cognitive psychology. *Neuroscience*, **139**, 5-21.
- 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*, **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.
- STEVANOVSKI, B., & JOLICŒUR, P. (2007). Visual short-term memory: Central capacity limitations in short-term consolidation. *Visual Cognition*, **15**, 532-563.
- UNSWORTH, N., & ENGLE, R. W. (2007). The nature of individual differences in working memory capacity: Active maintenance in primary memory and controlled search from secondary memory. *Psychological Review*, **114**, 104-132.
- VOGEL, E. K., WOODMAN, G. F., & LUCK, S. J. (2006). The time course of consolidation in visual working memory. *Journal of Experimental Psychology: Human Perception & Performance*, **32**, 1436-1451.

NOTE

1. To verify that participants memorized both word lists used in the secondary task sufficiently to produce stable long-term memory representations, a 4 (block order) \times 2 (condition) ANOVA of secondary task reaction times (RTs) was conducted. If participants relied on stable representations, then changes in RTs across blocks should not have occurred. Analysis of correct auditory-verbal secondary task RTs revealed no effect of block [$F(3,111) = 1.13, p = .34, \eta_p^2 = .03$], demonstrating that participants did not recall and respond to the target word or respond to the tone more quickly as the experiment progressed.

(Manuscript received August 20, 2009;
revision accepted for publication January 14, 2010.)