# Visual mental image generation does not overlap with visual short-term memory: A dual-task interference study

Gregoire Borst · Elaine Niven · Robert H. Logie

Published online: 12 October 2011 © Psychonomic Society, Inc. 2011

**Abstract** Visual mental imagery and working memory are often assumed to play similar roles in high-order functions, but little is known of their functional relationship. In this study, we investigated whether similar cognitive processes are involved in the generation of visual mental images, in short-term retention of those mental images, and in shortterm retention of visual information. Participants encoded and recalled visually or aurally presented sequences of letters under two interference conditions: spatial tapping or irrelevant visual input (IVI). In Experiment 1, spatial tapping selectively interfered with the retention of sequences of letters when participants generated visual mental images from aural presentation of the letter names and when the letters were presented visually. In Experiment 2, encoding of the sequences was disrupted by both interference tasks. However, in Experiment 3, IVI interfered with the generation of the mental images, but not with their retention, whereas spatial tapping was more disruptive during retention than during encoding. Results suggest that the temporary retention of visual mental images and of visual information may be supported by the same visual short-term memory store but that this store is not involved in image generation.

**Keywords** Working memory · Visual Mental Imagery · Visual cache · Visual Buffer

G. Borst (⋈)
University Paris Descartes, Sorbonne Paris Cité,
46 rue Saint Jacques,
75005 Paris, France
e-mail: gregoire.borst@parisdescartes.fr

E. Niven · R. H. Logie University of Edinburgh, Edinburgh, UK

When one wants to navigate an environment, to remember where one left one's keys, or, more generally, to anticipate or to emulate an action or a situation, visual mental imagery (MI) is at play. In contrast, keeping track of rapidly changing nearby traffic while driving on a busy highway or remembering whether one has just taken the round blue pill or the oval white pill relies rather more on temporary retention of visual information, with updating moment to moment. As has been suggested by Kosslyn, Thompson & Ganis (2006; Kosslyn, 1994), visual MI is likened to the experience of perception in the absence of the imaged stimulus. As has been suggested by Logie (1995; 2003, 2011a; Logie & van der Meulen 2009), visual short-term retention supports limited capacity temporary storage of the products of perception and imagery but involves a different set of cognitive functions. However, very few studies have directly addressed the relationship between MI and visual short-term memory (VSTM). In the three experiments described here, we focused on whether image generation, the short-term retention of visual mental images, and the short-term retention of visual information rely on the same or on different cognitive processes

Kossyln et al. (2006; Kosslyn, 1994) argued that visual MI and visual perception draw on much of the same cognitive processes and representations and recruit broadly the same brain areas, including area V1 in the primary visual cortex. For example, Ganis, Thompson, and Kosslyn (2004) demonstrated that more than 90% of the same brain areas—including physically topographically organized structures—are activated in visual MI and in visual perception. Topographically organized areas in the occipital lobe are grouped in a single functional structure—the *visual buffer*, which is the medium that supports visual mental images. Many of the properties of mental images are directly related to the fact that they rely on early visual

brain areas. For example, (1) mental images are difficult to maintain in the visual buffer, and (2) the visual buffer is sensitive to any visual input, given that in order to avoid visual persistence during eye movements, patterns of activations are transient in these brain areas during perception. MI has been shown to be important for a number of functions, such as learning (e.g., Paivio, 1971), memory (e.g., Schacter, 1996), and reasoning (e.g., Kosslyn, 1983).

In a parallel line of research, working memory (Baddeley, 1986, 2007), and particularly visuospatial working memory (VSWM; Logie, 1995, 2003, 2011b; Logie & van der Meulen 2009) has been conceived as a multicomponent mental workspace, with different combinations of components playing crucial roles in visual MI and visual short-term retention. According to this view, generation of MI involves activation of stored knowledge relevant to the stimuli, coupled with executive functions to process and manipulate that activated knowledge. The generation of images overlaps with the processes of visual perception. In this sense, the view is compatible with the computational model developed by Kosslyn and colleagues. In contrast, visual short-term retention is supported by a different, limited capacity component of VSWM that stores the products of visual perception and MI in a temporary visual cache or VSTM system. This visual cache is independent of visual imagery and of visual perception, but its contents comprise temporary representations formed after a single mental image has been created and after the perceptual process is complete (e.g., Logie, Brockmole, & Jaswal, 2011; Logie, Brockmole, & Vandenbroucke, 2009; van der Meulen, Logie, & Della Sala, 2009). Therefore, the visual working memory perspective is entirely compatible with the Kosslyn et al. (2006) model of visual imagery but it points to evidence that VSTM and generation of visual MI might rely on different sets of cognitive functions.

Previous studies have demonstrated that VSTM and MI draw on partially distinct cognitive processes. On the one hand, the presentation of irrelevant visual input (IVI) has no effect on visual short-term retention of unfamiliar Chinese characters (Andrade, Kemps, Werniers, May & Szmalec, 2002) or visual patterns (Zimmer & Speiser, 2002). This suggests that visual perceptual input does not have direct access to the memory store involved to disrupt its contents. On the other hand, performance in MI tasks, such as the

pegword imagery mnemonic that involves generation of a series of bizarre mental images during encoding and during retrieval, is affected by the concurrent presentation of IVI (e.g., Andrade et al., 2002; Logie, 1986; Quinn & McConnell, 1996; Zimmer & Speiser, 2002; Zimmer, Speiser, & Siedler, 2003). The effect of IVI is not restricted to the pegword mnemonic, nor is the disruptive effect dependent on the type of IVI used. IVI affects not only the pegword mnemonic, but also animal mental size comparisons (Dean, Dewhurst, Morris, & Whittaker, 2005; Logie, 1981), rating vividness of mental images (Baddeley & Andrade, 2000) and visualization of a route during rock climbing (Smyth & Waller, 1998). Disruption of performance in MI tasks has been reported with IVI, such as changing plain-colored squares (Logie, 1986), randomly changing black and white dots (dynamic visual noise; e.g., Quinn & McConnell, 1996), color matching (Zimmer & Speiser, 2002), or line drawings (e.g., Andrade et al., 2002; Logie, 1986; van der Meulen et al., 2009). IVI also has been shown to interfere with the recognition of the precise size of a circle (McConnell & Quinn, 2004), of the precise font of a letter (Darling, Della Sala, & Logie, 2007, 2009), of the precise color of a dot (Dent, 2010), or of a matrix (Dean, Dewhurst & Whittaker, 2008). Thus, IVI might selectively affect temporary retention of visual information when precise, detailed information needs to be maintained, and Darling et al. (2009) suggested that maintenance of such a high level of visual detail might require repeated generation of an image of the items, rather than retention in a VSTM system. Kosslyn and Thompson (2003) also argued that if a stimulus is novel or precise visual details are to be retained, this may involve image generation within the visual buffer.

Taken together, the results described above suggest (1) that IVI selectively interferes with the generation of visual mental images during encoding and retrieval, but not with temporary retention of recently presented visual information unless this involves a high level of detail, and (2) that MI and VSTM are partially distinct functions.

Some authors have proposed that visual MI and VSTM involve distinct sets of cognitive processes (e.g., Andrade et al., 2002; Logie, 1995, 2003; 2011a; 2011b; Logie & van der Meulen, 2009; Pearson, 2001). According to this view, MI generation would be supported by an active store (or visual buffer) that overlaps with perception and involves the primary visual areas in the posterior occipital cortex, including V1 (Ganis, Thompson, & Kosslyn, 2009; Kosslyn, 1994; Kosslyn et al., 1993; Kosslyn et al., 2006; Kosslyn & Thompson, 2003). Because it overlaps with visual perception, it would be sensitive to disruption from IVI. However, on this same view, short-term retention of visual information and, possibly, retention of a single image would be achieved through a passive VSTM store or visual cache (Logie, 1995). Visual input does not directly access



In the literature on visual attention, visual perception, and visual feature binding, the terms *visual short-term memory* and *visual working memory* tend to be used interchangeably. As is noted here, we view VSTM as comprising a temporary store that is one of a range of functions of visual working memory, which, in turn, is a set of functions within a broader, multicomponent working memory, and we use the terms in this way throughout the article. See Logie and van der Meulen (2009) for a review and detailed discussion.

the visual cache, and therefore, the contents would not be sensitive to disruption from irrelevant visual stimuli. The visual cache holds the products of initial perceptual processing and is conceived of as being implemented in the posterior parietal cortex (Todd & Marois, 2004, 2005). This view makes sense given what we know of the neural architecture of the visual buffer and of the visual cache. The visual cache is thought to support retention of the visual appearance of a single visual pattern that might comprise a small number (e.g., four) of visual object representations in an array (e.g., Logie, 1995; Logie et al., 2011; Logie et al., 2009; Logie, Zucco, & Baddeley, 1990; Luck & Vogel, 1997; Phillips & Christie, 1977; Vogel, Woodman, & Luck, 2001). The representation has a low visual resolution, so precise visual details are not retained, but the overall shapes and colors of the objects are retained. The visual cache does not overlap with perception or MI generation but stores the outputs from those processes after they are complete in the form of a visual temporary representation. This explains why VSTM performance appears to be unaffected by IVI.

In a recent study, van der Meulen et al. (2009) demonstrated that the generation of visual mental images and VSTM were not sensitive to the same interference tasks. When participants were asked to retain, for 15 s, a series of four letters shown in a mixture of upper- and lowercase (e.g., Q-h-D-r) presented visually one at a time, their performance was affected by spatial tapping, but not by IVI (a series of unrelated line drawings of objects and animals). Conversely, when participants were asked to generate visual mental images from aural presentation of the names of letters in order to judge the figural properties of these letters (e.g., presence of curved line), their performance was affected by IVI, but not by spatial tapping. This double dissociation suggests that image generation and short-term retention of visual information do not rely on the same underlying cognitive systems. The disruptive effect of spatial tapping on short-term retention of visual information was interpreted by van der Meulen et al. as a consequence of having to simultaneously retain in VSTM the (unseen) tapping pattern to be followed and the visual form of the sequence of letters to be retained (see also, e.g., Della Sala, Gray, Baddeley, Allamano, & Wilson, 1999; Engelkamp, Mohr & Logie, 1995; Logie et al., 1990; Quinn & Ralston, 1986; Salway & Logie, 1995). According to Kosslyn's model, IVI disrupts generation processes because it interferes with the content of the visual buffer. According to the Logie (1995, 2011a) model, IVI disrupts the activation of stored knowledge associated with a stimulus, and this is wholly compatible with the Kosslyn view. In the case of the van der Meulen et al. study, this content would comprise the letter form being imaged in the visual buffer. Again, according to the Logie model, IVI did not disrupt memory for the letter form because this visual information is held in the visual cache, which is a temporary memory store in which the visual form is stored after it has been generated and is separate from the visual buffer, which supports image generation and low-level visual processing.

The new experiments reported here address a question arising from the differential pattern of interference between image generation and short-term memory reported by van der Meulen et al. (2009) and from the contrasts between the concept of the visual buffer for MI and the concept of the visual cache for VSTM. We investigated whether the shortterm retention (not the generation) of visual mental images (not only recently presented visual information) relies on the visual cache. We used a dual-task methodology to study the effect of IVI and spatial tapping during the retention of series of letters presented either aurally (i.e., auditory task) or visually (i.e., visual task). In both tasks, participants were asked to remember visual properties of letters. On each trial, a sequence of four letters was presented either aurally (for generation of visual mental images of the letters) or visually. Letters could be lower- or uppercase letters. An articulatory suppression task was performed during the entire duration of each trial to prevent phonological encoding of the identity of the letters. After the presentation of the fourth letter, participants were required to retain the letters for 12 s. During the retention interval, participants were instructed to fixate their gaze on a dot at the center of the screen (control condition), to tap with one finger each of nine numeric keys (unseen) in a 3 × 3 arrangement, following a figure-of-eight pattern (spatial tapping condition), or to press the space bar each time a line drawing was presented on the screen (IVI condition). We selected line drawings over DVN for the IVI task given that (1) DVN failed to interfere with temporary memory for visual information in some studies (e.g., Andrade et al., 2002) and with the generation of mental images in others (e.g., Pearson, Logie, & Gilhooly, 1999; Zimmer & Speiser, 2002) and that (2) van der Meulen et al. showed that presentation of a series of line drawings interfered selectively with image generation—not with retention of visual information—whereas their pilot data showed that DVN failed to interfere with any of their tasks. Immediately after the retention interval, participants were asked for written recall of the identity, the case, and the position of the letters in the series. Written recall was used to further emphasize the requirement to retain the visual form of the presented letters, whether presented visually or aurally. Aural presentation would require generation of an image of each letter and retention of a visual representation of all letters in the sequence. Visual presentation would allow for a more direct formation of a visual representation for retention, without the need for visual imagery generation. We reasoned that if a mental image from aural presentation



is not retained within the visual buffer but, rather, is retained within the visual cache of VSTM, spatial tapping, and not IVI, should interfere with retention of an image, given that spatial tapping interferes with information held in the visual cache, but not with that generated in the visual buffer (see van der Meulen et al., 2009). On the other hand, if the retention of visual mental images based on auditory input relies on the visual buffer and the retention of visually presented information relies on the visual cache in VSTM, we would expect IVI to interfere with the retention of a visual image and spatial tapping to interfere with the retention of letters presented visually.

# **Experiment 1**

## Method

Participants We recruited 32 native English-speaking volunteers with normal or corrected-to-normal vision from Harvard University (18 females and 14 males, with an average age of 22.7 years; 11 of them were left-handed). Participants received either a cash payment or course credit. All participants provided written consent and were tested in accordance with national and international norms governing the use of human research participants. The research was approved by the Harvard University Faculty of Arts and Sciences Committee on the Use of Human Subjects.

*Materials* Stimuli were presented on a 19-in. IBM monitor  $(1,280 \times 1,024 \text{ pixel resolution})$  and refresh rate of 75 Hz) using E-Prime software running under Windows XP. All stimuli were presented on a uniform white background throughout the entire experiment.

In both the visual task and the auditory task, participants were instructed to remember a series of four letters presented one at a time (Logie, Della Sala, Wynn, & Baddeley, 2000; van der Meulen et al., 2009). We selected six letters with visually dissimilar lower- and uppercase forms (Dd, Hh, Ll, Mm, Qq, and Rr). In the visual task, letters were presented one at a time in a 240-point saturated black Arial font in the center of the screen—for example, R–d–h–Q. In the auditory task, we prepared 600-ms audio files of the spoken name of each letter. The name of the letter was preceded by "low" or "up" when the letter was to be visualized in its lower- or uppercase form, respectively—for example, "up–R," "low–d," "low–d," "up–Q."

We designed two versions of 48 sequences of four mixed upper- and lowercase letters. In each version, each of the six letters in their uppercase and lowercase forms was presented 2 times in each of the four possible serial positions. Within each sequence, any one of the six letters could appear only once.

Following the procedure from van der Meulen et al. (2009), for the IVI task, we selected 120 black-and-white line drawings  $(341 \times 341 \text{ pixels})$  from the Snodgrass and Vanderwart (1980) pictures. Pictures were presented in the center of the screen. In the spatial tapping task, participants used a numeric pad to perform the hand tapping. The numeric pad was placed out of sight of the participant, but within easy arm reach.

In all tasks, a black fixation point ( $10 \times 10$  pixels) was presented at the center of the screen when no other stimuli were displayed. Participants were instructed to maintain their gaze on the fixation point at all times when it appeared on the screen in any of the experimental conditions (control, IVI, and spatial tapping).

Procedure Participants were tested individually sitting approximately 60 cm away from the computer screen. Audio files for the MI task were presented through a headset. Participants were randomly assigned to two groups. In each group, participants performed the auditory task and the visual task. In one group, participants performed the two tasks under a control condition and with the IVI task. In the other group, participants performed both tasks under a control condition and with the spatial tapping task. The order of the control and the dual-task conditions was counterbalanced across participants. Participants were randomly assigned to one of the two versions of 48 sequences. Twenty-four sequences were used in the control condition, and 24 sequences in the dual-task condition. The allocation of sequences to conditions was counterbalanced across participants. Participants performed 2 practice sequences of letters before the first experimental trial in the control condition and in the dual-task condition.

In both the visual and auditory tasks, participants were asked to memorize the identity, letter case, and presentation order of the four letters in each sequence. Each trial started with the presentation of a fixation point (1,000 ms), on which participants were instructed to focus their gaze. Then letters of the sequence were presented one at a time either visually or aurally for 600 ms, with a 1,500-ms interstimulus interval (ISI) during which the fixation point was displayed. The duration of the ISI was set in order to allow participants to form the visual mental images of the letters in the MI task. Finally, 1,500 ms after the presentation of the fourth letter of the sequence, a tone indicated the start of the 12-s retention phase. After 12 s, a tone prompted the recall of the sequence of letters. Participants wrote their responses on a response sheet that comprised four boxes arranged horizontally for each sequence. Participants were given unlimited time to recall the four letters.

In order to increase the likelihood of visual coding of the sequences of letters, participants performed articulatory suppression during the presentation and the retention of the



letters—repeating "the" 3 times per second. In addition, participants were explicitly instructed to remember the letters visually, and not as letter names.

In both the visual and the auditory tasks, on each trial during the 12-s retention of the letters, participants were asked either to fixate a point (i.e., control condition) or to perform one of the two interference tasks. In the spatial tapping task, on each trial, participants performed tapping with the index finger of their dominant hand on nine keys of a 3 × 3 arrangement on the numeric pad. Participants were requested to follow a figure-of-eight pattern starting with the bottom right key (i.e., the 3 key) and to tap a key every 500 ms. We asked participants to focus on the fixation point displayed at the center of the screen while they performed spatial tapping. Participants adjusted their spatial tapping speed in two 1-min sessions in which a tone prompting a keypress was presented every 500 ms. Two practice sessions were performed before the two practice trials of the dual-task condition. We recorded the sequence of keys tapped and the intervals between each tap.

In the IVI task, on each trial, a series of 24 different line drawings was presented on the screen. Pictures were presented for 150 ms, with a 350-ms ISI. During the ISI, participants focused their gaze on a fixation point. Participants were required to press the space bar each time a picture was displayed on the screen. Thus, the tapping rate was similar in the two interference tasks. We recorded the time between each space bar press.

## Results

For both the visual and auditory tasks, the performance on memory for the letters was determined by the proportion of sequences correctly recalled (see Table 1). We analyzed the results in a three-way task (visual vs. auditory)  $\times$  condition (control vs. dual task)  $\times$  interference type (IVI vs. spatial tapping) mixed design analysis of variance (ANOVA). When we compared two means, we computed one-tailed t tests in accordance with our hypotheses, and we corrected the alpha level with a Bonferroni procedure. In addition, for each of the analyses, we report the effect size of the

Table 1 Experiment 1: Mean proportions of sequences accurately recalled in the auditory and the visual tasks in the control and interference (IVI vs. spatial tapping) conditions for the three types of errors (identity, case form, and serial position)

IVI Spatial Tapping Identity Case Position Identity Case Position SD SD SD M SD M SD M SD M M M .24 MI task Control .86 .14 .81 .20 .77 .19 .88 .14 .83 .16 .73 .74 .77 .21 Interference .88 .18 .23 .18 .82 .15 .66 .67 .20 WM task Control .81 .24 .77 .23 .79 .24 .91 .08 .83 .15 .85 .11 Interference .83 .22 .71 .28 .76 .21 .86 .14 .69 .22 .78 .20

ANOVA (partial eta squared) or of the difference of the means (Cohen's d).

The three-way interaction was not significant, F < 1, nor was the main effect of task, F(1, 30) = 1.49, p = .23, but we found a significant main effect of condition, F(1, 30) = 21.6, p < .0001,  $\eta_p^2 = .42$ , and a marginal two way interaction between condition and interference type, F(1, 30) = 3.46, p = .07,  $\eta_p^2 = .10$ . Specific comparisons revealed that in the auditory task, the retention of the sequences of letters was significantly disrupted by spatial tapping, t(15) = -3.55, p < .01, d = 0.67, but not by IVI, t < 1 (see Fig. 1). Similarly, in the visual task, the retention of letters was affected by spatial tapping, t(15) = -5.86, p < .0001, t = 0.72, but not by IVI, t < 1 (see Fig. 1).

In order to determine whether the spatial tapping task affected equally the recall of the identity, the case, and the serial position of the letters, we conducted a new series of analyses (see Table 1). The analyses revealed that participants were selectively impaired on the retention of the case of the letters when they performed spatial tapping in the auditory task and in the visual task, t(15) = -4.39, p < .005, d = 0.91, and t(15) = -3.91, p < .01, d = 0.74, respectively. Conversely, spatial tapping had no effect on the retention of the identity or the serial position of the letters (all ps > .15). Thus, spatial tapping affected selectively the retention of the case forms of the letters, which was the property of the letter least likely to be coded phonologically and most likely to be coded visually.

Finally, although the same pattern of results was observed in the two primary tasks, in order to conclude that the results were actually similar, we needed to show that in both tasks, performance in the two interference tasks was also similar. Participants were as accurate—with errors defined as omissions or extra taps and change in tapping direction—in spatial tapping in the visual task (M = .91) as in the auditory task (M = .92), t < 1. In addition, neither the mean intertap intervals (M = 483 ms in the auditory task and M = 496 ms in the visual task) nor the variance in intertap intervals differed between the two tasks, all t < 1. Finally, participants performed the IVI task as accurately during the auditory task (M = .91) as during the visual task (M = .94), t < 1, and required the same amount of time to



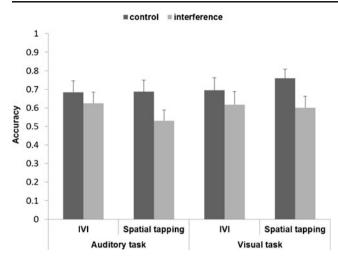


Fig. 1 Experiment 1: Mean proportions of sequences accurately recalled in the auditory and the visual tasks in the control and interference (IVI vs. spatial tapping) conditions. Error bars denote standard errors of the means

detect the presentation of a picture during the two tasks (M = 281 ms during the auditory task and M = 274 ms during the visual task), t < 1. Thus, it is unlikely that the similarity of the interference effects observed in the visual and the auditory tasks simply reflect different levels of trade-off between the primary tasks (i.e., visual and auditory tasks) and the interference tasks (i.e., IVI and spatial tapping tasks).

### Discussion

Spatial tapping selectively interfered with the retention of sequences of letters presented visually or aurally. Conversely, the concurrent presentation of a series of line drawings during the retention of sequences of letters had no effect on the recall of these letters, regardless of the nature of the presentation (visual vs. auditory). The pattern of interference reported in our experiment on the short-term retention of visually presented sequences of letters replicated previous findings (van der Meulen et al., 2009). Our data are consistent with the van der Meulen et al. interpretation that tapping disrupts the visual retention of the sequence of letters because keeping track of the progress of the tapping pattern and maintaining the visual property of the letters compete for similar resources in a passive VSTM system, referred to as the *visual cache* (Logie, 1995).

In this study, by making the procedures used in the visual and the auditory tasks more directly comparable, we extended previous results by showing that spatial tapping disrupts short-term retention of visual information not only when stimuli are presented visually, but also when they are encoded by relying on visual mental images (in the auditory task). Interestingly, only the retention of the case form—not the identity or the serial position—of the letters was

affected by spatial tapping. Also, van der Meulen et al. (2009) found that the case form was most sensitive to visuospatial interference, although identity and serial position information were also affected, but to a lesser extent in that earlier study. This greater sensitivity of case form to visuospatial interference probably arises because it is the property of the letters most likely to be represented and retained as a visual code, whereas letter identity and serial order might rely more on a phonological code, despite the use of articulatory suppression. Consistent with this view, Logie et al. (2000) and Saito, Logie, Morita, and Law (2008) have provided evidence for the use of both visual codes and phonological codes in written recall of visually presented letter and word sequences.

In the auditory task, the lack of interference produced by the IVI task suggests that visual mental images are not retained in the visual buffer after they have been generated—in contrast to the hypothesis proposed by Kosslyn et al. (2006). An alternative explanation is that visual mental images may be stored in the visual buffer only when they are novel and when detailed visual information must be retained (see Kosslyn & Thompson, 2003), neither of which was a property of the stimulus materials we used here.

Taken together, the results rather suggest that after they have been generated in the visual buffer, retention of visual mental images is supported by a VSTM system in the same way as are other visual representations of recently presented visual stimuli.

However, two major objections can be raised. First, one could argue that participants did not necessarily generate visual mental images of the letters in the auditory version of the task. Second, one could argue that part of our interpretation relies, at least in part, on a null result—the lack of an effect of the IVI task, despite the evidence for its effectiveness in disrupting generation of visual images in the van der Meulen et al. (2009) study. Thus, in order to demonstrate that our IVI task was not simply an insensitive manipulation for exploring visual interference, it would be important to demonstrate that our version of IVI can interfere with image generation, even if it does not disrupt retention of images. Experiment 2 addressed this by exploring (1) whether participants would generate visual mental images of the letters when the letters were presented aurally (i.e., auditory task) and (2) whether IVI would affect performance in the auditory task when the IVI was presented during image generation, rather than during the retention period for the image.

## **Experiment 2**

In this experiment, we asked a new group of participants to perform the same auditory task as in Experiment 1, with the



same control and interference conditions. However, in this experiment, the interference tasks were performed during the encoding phase (i.e., when the name of the letters were presented aurally), not during the retention phase. We reasoned that if visual mental images are generated during the encoding of the sequences of letters, the presentation of a series of line drawings (i.e., IVI task) should affect the generation of mental images of the letters and, consequently, the recall of the sequences, as has been shown previously (e.g., Logie, 1986; Quinn & McConnell, 2006; van der Meulen et al., 2009). We also reasoned that as the visual image of each letter was generated, it would then be encoded into the visual cache, and so we expected that spatial tapping during encoding would also interfere with memory performance. That is, there should be a main effect of interference, although the causes of that interference would differ, but there would be no interaction with interference type.

## Method

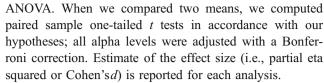
Participants Twenty-four native English-speaking students from Harvard University volunteered to take part in this experiment (11 females and 13 males; average age of 20.4 years; all right-handed but 2). None of them had taken part in Experiment 1. Participants all had normal or corrected-to-normal vision and received payment or course credit for their participation. The research was approved by the Harvard University Faculty of Arts and Sciences Committee on the Use of Human Subjects. All participants provided written consent and were tested in accordance with national and international norms governing the use of human research participants.

*Materials* The materials were the same as those used in Experiment 1.

Procedure The procedure was strictly identical to that used in Experiment 1, except that (1) participants performed only the auditory task, not the visual task, and (2) the interference tasks were performed during the encoding phase—during the sequential aural presentation of the four letters. As in Experiment 1, participants were randomly assigned to one of two groups. The two groups differed in the type of interference task performed (i.e., IVI or spatial tapping task).

# Results

The proportion of sequence of letters correctly recalled was analyzed within a condition (control vs. dual task) × interference type (IVI vs. spatial tapping) mixed design



The two-way interaction failed to reach significance, F (1, 22) = 2.91, p = .10, as well as the main effect of type of interference, F < 1, but the main effect of condition was significant, F(1, 22) = 53.31, p < .0001,  $\eta_p^2$  = .71. As is evident in Fig. 2, the proportion of sequences correctly recalled was lower in the IVI task (M = .47) than in the control condition (M = .75), t(11) = 3.55, p < .005, d = 1.3. In the same way, the retention of the sequences of letters was lower in the spatial tapping task (M = .42) than in the control condition (M = .86), t(11) = 7.31, p < .0001, d = 2.52. These results are consistent with our hypotheses.

An analysis of the type of errors (i.e., identity, case form, serial position; see Table 2) in the retention of the sequences of letters revealed no two-way interaction between condition and type of interference, F(1, 22) = 1.16, p = .29 for identity, F(1, 22) = 2.05, p = .17 for case form, and F(1, 22) = 1.53, p = .23 for serial position, and no main effect of interference type, all Fs < 1. However, the main effect of condition was significant for identity, F(1, 22) = 46.59, p < .0001,  $\eta_p^2 = .53$ , and for serial position, F(1, 22) = 27.26, p < .0001,  $\eta_p^2 = .53$ , and for serial position, F(1, 22) = 40.62, p < .0001,  $\eta_p^2 = .65$ . IVI and spatial tapping affected the retention of the identity, t(11) = 3.98, p < .01, d = 1.33, and t(11) = 5.71, p < .0001, d = 1.83, the case form, t(11) = 3.03, p < .05, d = 1.11, and t(11) = 4.27, p < .005, d = 1.88, and the serial positions of the letters, t(11) = 2.99, p < .05, d = 1.21, and t(11) = 7.41, p < .0001, d = 2.32, respectively.

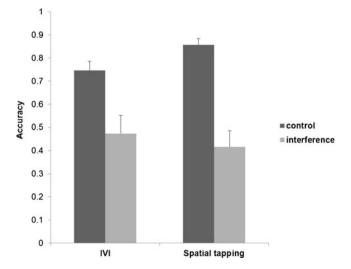


Fig. 2 Experiment 2; Mean proportions of sequences accurately recalled in the auditory and the visual tasks in the control and interference (IVI vs. spatial tapping) conditions. Error bars denote standard errors of the means



Table 2 Experiment 2: Mean proportions of sequences accurately recalled in the auditory and the visual tasks in the control and interference (IVI vs. spatial tapping) conditions for the three types of errors (identity, case form, and serial position)

		MI Task							
		Identity		Case		Position			
		M	SD	M	SD	M	SD		
IVI	Control	.90	.11	.88	.09	.81	.15		
	Interference	.71	.17	.70	.21	.56	.25		
Spatial tapping	Control	.94	.07	.95	.05	.89	.09		
	Interference	.69	.18	.65	.22	.53	.20		

Finally, participants were accurate in spatial tapping (M = .97) and in the IVI task (M = .99). In addition, the mean intertap intervals (M = 496 ms) did not differ significantly with the expected interval (M = 500 ms), t < 1. Thus, we found no hint of performance trade-off between the primary task and the interference tasks.

## Discussion

We replicated previous findings in showing that the presentation of series of line drawings during the encoding phase, in which participants needed to generate visual mental images of letters presented aurally, affected the retention of the visual properties of sequences of letters (van der Meulen et al., 2009). Thus, it is unlikely that participants did not generate mental images of the letters in Experiment 1 when prompted to do so, and the results of Experiment 1 cannot readily be explained by suggesting that the imagery required in the main memory task was simply insensitive to disruption from our version of IVI. As was expected, spatial tapping also impaired recall when performed concurrently with the aural presentation of the letters. This is consistent with our hypothesis that during the encoding phase, participants not only need to generate visual mental images of each letter, but also need to encode and retain in the visual cache each letter as it is presented and then also retain that letter representation while subsequent letters in the sequence for that trial are presented and added to the visual cache, in turn, until the representation of the sequence of four letters is complete.

If mental images are held in the visual buffer only for a very brief period of time (i.e., the time taken for the generation process to occur) but, once generated, are maintained in the visual cache, it should be possible to interfere, during the encoding phase, either with the generation of the visual mental image or with the maintenance of these images in the visual cache. Experiment 3 was designed to determine whether each mental image of the letter would be maintained in the visual cache during encoding—while generation of subsequent letters was performed.

## **Experiment 3**

In Experiment 3, a new group of participants performed the auditory task. Half of the participants performed the auditory task with spatial tapping, and the other half with IVI. The interference tasks were presented during the encoding phase, as in Experiment 2, but in one condition, the interference occurred only when the letters were presented aurally (image generation), and in another condition, the interference occurred only during the ISI. If participants generated mental images of the letters at the time each letter was presented aurally, concurrent presentation of pictures during this period should affect the ability to generate the images and, hence, should impair subsequent recall of the letters. In addition, if mental images of letters were not maintained in the visual buffer but were held in the visual cache. IVI should not interfere with sequence recall when it was presented during the ISI, which would occur after image generation had taken place and after that particular letter had been encoded in the visual cache. Spatial tapping should interfere with sequence retention when performed during the ISI and might also disrupt performance when performed during the time devoted to image generation, assuming that the image was encoded in the visual cache as the image was generated.

### Method

Participants Participants were 48 (30 female, 18 male; mean age, 19.9 years; 43 right-handed and 5 left-handed) native English speakers, students from Harvard University (n = 32) and the University of Edinburgh (n = 16). None of them had taken part in Experiment 1 or 2, and all reported normal or corrected-to-normal vision. Volunteers received payment or course credit for their participation. All participants provided written consent and were tested in accordance with national and international norms governing the use of human research participants.

*Materials* The materials were the same as those used in Experiment 1.



Procedure The procedure was the same as that used in Experiment 1, with the following exceptions. Participants were assigned to one of four groups depending on the type of interference task and when each was performed—namely, IVI task during image generation, IVI during ISI (or image retention), spatial tapping during image generation, and spatial tapping during ISI. There were 12 participants per group. All participants also completed the control (no-interference) condition.

On each trial, first a fixation point (1,000 ms) was displayed, and participants were instructed to focus their gaze on it. Then a letter was presented aurally (600 ms), and participants were allowed 1,900 ms to visualize the letter. After 2,500 ms (ISI), another letter was presented, so letters were presented at a rate of one every 4.4 s. The specific timing for the generation phase was chosen on the basis of the results reported by Thompson, Kosslyn, Hoffman, and van der Kooij (2008) showing that participants take, on average, less than 1,000 ms to generate a mental image of a single letter. The ISI timing was set to 2,500 ms to mirror the timing used for the generation of the image. The retention phase was limited to 6 s, given that we extended the duration of the encoding phase. This ensured that the overall time for each trial was approximately the same as for Experiments 1 and 2.

During the 2,500-ms time for presentation and visualization of the letters or during the 2,500-ms ISI, participants were asked either to fixate a point (i.e., control condition) or to perform one of the two interference tasks. The fixation point turned red when participants needed to perform the interference task. Participants were explicitly instructed to visualize the letter as soon as they heard the name of the letter, and not during the ISI. Participants performed two practice series in each experimental condition (control vs. dual task).

In the spatial tapping task, participants performed tapping with the index finger of their dominant hand on nine keys of the numeric pad in a  $3 \times 3$  arrangement. Participants were required to follow a 3-1-5-9-7 sequence at a rate of one key pressed every 500 ms. Participants were requested to focus on the fixation point displayed at the center of the screen during spatial tapping. As in previous experiments, participants practiced spatial tapping before the experimental phase. In the IVI task, five different line drawings were presented on the screen. Pictures were presented for 150 ms, with a 350-ms ISI. Participants were required to press the space bar each time a picture was displayed on the screen. Thus, the tapping rate was similar in the two interference tasks, and the interference tasks were the same whether they were performed during the presentation time or during the retention interval.



We analyzed the proportion of correctly recalled sequences of letters within a condition (control vs. dual task)  $\times$  interference type (IVI vs. spatial tapping)  $\times$  encoding phase (image generation vs. image retention) mixed design ANOVA. When two means were compared, we reported paired sample one-tailed t tests and corrected alpha levels with a Bonferroni procedure. For each analysis, we reported the effect size (i.e., partial eta square or Cohen's d).

The three-way interaction was significant, F(1, 44) =6.13, p < .05,  $\eta_p^2 = .12$ , but none of the two-way interactions reached significance, F < 1 for the interaction between condition and interference type, F(1, 44) = 1.67, p = .20 for the interaction between condition and encoding phase, and F(1, 44) = 1.87, p = .17 for the interaction between interference type and encoding phase (see Fig. 3). The main effect of experimental condition was significant, F(1, 44) = 35.89, p < .0001,  $\eta_p^2 = .45$ , but not the main effects of interference type and encoding phase, Fs < 1. Critically, the recall of the letters was impaired, as compared with the control condition, when the IVI task was performed during the generation of the mental images of the letters, t(11) = 4.96, p < .0001, d = 1.14, but not when performed during the ISI (i.e., the retention of the mental images), t < 1. However, the retention of sequences of letters was affected by spatial tapping regardless of when spatial tapping was performed, t(11) = 3.19, p < .05, d = 0.71during image generation and t(11) = 4.01, p < .01, d = 1.23during ISI.

One could argue that the pattern of interference reported in the present study simply reflects a difference in attentional resources between the two interference tasks. We reasoned that if one of the interference tasks involves a

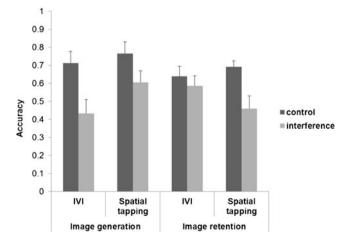


Fig. 3 Experiment 3: Mean proportions of sequences accurately recalled during image generation or image retention in the control and interference (IVI vs. spatial tapping) conditions. Error bars denote standard errors of the means



greater amount of attentional resources than does the other. this task should produce more interference regardless of the primary task (image generation vs. image retention). To test this hypothesis, we compared the magnitude of the interference effect produced by spatial tapping with the one produced by IVI on the retention of the letters when spatial tapping and IVI were performed during either the generation of the image or the ISI (i.e., the retention of the image). The magnitude of the interference effect was computed by subtracting participants' performance in the interference tasks from their performance in the control condition during the generation of the image or the ISI. Critically, IVI (M = .28) produced more interference than did spatial tapping (M = .16) when performed during image generation, t(22) = 1.71, p < .05, d = 0.70, but spatial tapping (M = .23) produced more interference than did IVI (M = .05) when performed during the ISI, t(22) = 1.88, p <.05, d = 0.80. Thus, differences in amounts of attentional resources of the interference tasks cannot account for the pattern of interference reported in the present experiment.

Finally, we further analyzed the type of errors in the retention of the sequences of letters (see Table 3) for the three planned comparisons on which significant differences were observed between the control and the interference conditions. IVI task performed during the image generation selectively affected the retention of the case form of the sequences of letters, t(11) = 2.85, p < .025, d = 0.79, but not the identity or serial position. Spatial tapping during the ISI selectively impaired the retention of the case form of the letters, t(11) = 3.34, p < .01, d = 1.13. Spatial tapping performed during the generation of the letters impaired the identity, t(11) = 4.25, p < .005, d = 0.86, the case form, t(11) = 3.72, p < .01, d = 0.73, and the serial position, t(11) = 3.81, p < .01, d = 0.85, of the letters.

Finally, participants were as accurate in spatial tapping during image generation (M = .96) as during ISI (M = .95), t < 1; the intertap intervals were similar during image generation (M = 496 ms) and during ISI (M = 505 ms), t < 1, and the intertap variance did not differ, t < 1. In the same way, performance in the IVI task did not significantly differ when performed during image generation (M = .99 for accuracy and M = 257 ms for RTs) or during ISI (M = .99 for accuracy and M = 247 ms for RTs), all t < 1.

## Discussion

The results converged in showing that presenting irrelevant pictures selectively disrupted the recall of aurally presented sequences of letters—and more specifically, the case form of these letters—by interfering with visual mental image generation processes. Conversely, IVI had no effect when performed during the ISI. Thus, mental images are unlikely to have been held in the visual buffer during the ISI,

because the visual buffer is assumed to be sensitive to any visual input.

Spatial tapping impaired the recall of the sequences of letters when performed during the ISI, which might suggest that mental image retention takes place in the VSTM system that we refer to as the visual cache. However, spatial tapping also interfered with the recall of the sequences of letters when performed during generation of the mental images of letters. The additional effect of spatial tapping during image generation could be due to interference with the encoding of the items in the visual cache. With auditory presentation of the letters, a visual mental image of each letter first has to be generated, and then, as it is generated, the resulting images would be encoded in the visual cache. It is therefore entirely likely that at least some participants generated an image of the letter within or very soon after the 600-ms aural presentation time and proceeded to encode the images in the visual cache during that time—therefore, effectively increasing the ISI between image generation and presentation of the next stimulus. Alternatively, one could argue that the spatial tapping task required the generation of a mental image to keep track of the unseen pattern of tapping. If so, spatial tapping might interfere with the retention of the series of letters when performed during the generation of the image of the letters, because both tasks load the visual buffer. However, we reasoned that spatial tapping should not interfere with image generation, because when one needs to represent the keys to tap on the numeric pad (to retain the tapping pattern), one needs to generate a spatial mental image, not a visual mental image, as a mental image of a letter. Critically, spatial mental images elicit activation not within early visual areas, but within the parietal cortex and, thus, should not interfere with the generation of visual mental images of letters within the visual buffer (see Aleman et al., 2002; Kosslyn et al., 2006).

One could also argue, consistent with a unitary account of working memory that does not assume a separate VSTM system (e.g., Cowan, 2005), that spatial tapping simply demands more attention than do irrelevant pictures, given that spatial tapping requires both motor planning and tracking the pattern of hand tapping. Thus, the more demanding task of spatial tapping would affect both initial encoding and maintenance of visual representations of letters, whereas the less demanding task involving irrelevant pictures would affect only initial encoding. Furthermore, the unitary model could also account for our results by suggesting that disruption of initial encoding by either of the secondary tasks (spatial tapping or IVI) occurs simply because encoding is more demanding than is maintenance. This is a plausible account but is rather post hoc, because it was not entirely clear in advance of the experiment that one of the secondary tasks would be any more demanding than



Table 3 Experiment 3: Mean proportions of sequences accurately recalled during image generation or image retention in the control and interference (IVI vs. spatial tapping) conditions for the three types of errors (identity, case form, and serial position)

			MI Task							
			Identity		Case		Position			
			M	SD	M	SD	M	SD		
Generation	Spatial tapping	Control	.91	.10	.85	.17	.84	.15		
		Interference	.82	.13	.71	.21	.70	.18		
	IVI	Control	.81	.18	.79	.21	.74	.24		
		Interference	.71	.23	.60	.27	.63	.26		
Maintenance	Spatial tapping	Control	.90	.07	.84	.12	.79	.13		
		Interference	.76	.17	.63	.21	.64	.19		
	IVI	Control	.85	.13	.84	.14	.74	.16		
		Interference	.77	.18	.75	.17	.69	.20		

the other. It would be equally plausible to predict that viewing meaningful line drawings of recognizable objects would be much more distracting and demanding of attention than performing a relatively automatic motor task involving repetition of a simple movement sequence.<sup>2</sup> So, the multicomponent model of VSWM predicts the results that we obtained on the basis of the nature (not the overall cognitive demand) of the tasks chosen, whereas a model based on overall demand on attention could explain some aspects of the results after the fact but could not readily predict the experimental outcomes.

A further reason for suggesting that our data cannot readily be explained in terms of overall demands on attention comes from the fact that IVI produced more interference than did spatial tapping when performed during image generation, whereas spatial tapping affected the retention of the letters during the ISI but IVI did not. Finally, consistent with the pattern of interference reported in the present study, complementary data patterns reported by van der Meulen et al. (2009) demonstrated that IVI selectively interferes with image generation, whereas spatial tapping selectively interferes with maintenance of information in VSTM independently of the attentional demand of the task. In fact, in their experiments, spatial tapping had no effect on the more attentional demanding task (i.e., the MI task). Thus, it is more likely that spatial tapping disrupted the recall of the sequences of letters not because of an interference with generation processes, but because of its disruptive effect on the encoding of visual representations in the visual cache.

<sup>&</sup>lt;sup>2</sup> Note that we designed the IVI task to limit such an effect: Participants were instructed simply to press a space bar whenever a line drawing appeared, while disregarding the meaning of the line drawing.



## General discussion

The three experiments reported in this study converged in showing that (1) IVI selectively interferes with the generation of visual mental images, but not the maintenance of visual representations in VSTM, (2) mental images are generated in a buffer sensitive to IVI, but (3) once generated, they are encoded and maintained in a VSTM store (or visual cache) that is insensitive to IVI but is sensitive to disruption by concurrent unseen spatial tapping. Critically, our results suggest that the visual cache both can mediate the short-term retention of visual representations after they have been created on the basis of retinal inputs and can also mediate the short-term retention of visual representations generated from information stored in longterm memory, such as visual mental images. The disruptive effect on the recall of sequences of letters of spatial tapping performed during the generation of mental images might suggest that mental images cannot be maintained in the visual buffer and are fed into the visual cache as soon as they are formed. In fact, mental images are unlikely to be maintained even for short periods of time within the visual buffer, given that the visual buffer relies on early visual areas in which activations have to be transient in order to avoid perceptual persistence when the eyes are moved.

A potential limitation of the pattern of results reported in this study is the lack of a full double dissociation in Experiment 3. However, image generation in that experiment was disrupted more by the IVI task than by spatial tapping, and conversely, the retention of mental images was interfered with by spatial tapping, but not by IVI. IVI did not interfere with the retention of visual information in Experiment 1 but did interfere with image generation in Experiments 2 and 3. In addition, a previous study did report the expected double dissociation between image generation and retention of visual information, using

similar interference tasks (van der Meulen et al., 2009). Thus, the balance of evidence suggests that the lack of a clear double dissociation arises not because the two interference tasks differed in attentional demands, but because visual images based on auditory input are generated within a visual buffer only very briefly and, as soon as they are generated, are fed into the visual cache to be maintained for the period of a single trial.

Finally, given the structural equivalence between visual mental images and visual percepts (e.g., Borst & Kosslyn, 2008), we would argue that the visual cache can be accessed by visual input and mental images in a similar way, but only after they have been processed in the visual perceptual system. Furthermore, our data complement and extend previous evidence and a theoretical perspective (Logie, 1986, 1995, 2003; 2011a; 2011b; Logie et al., 2009; van der Meulen et al., 2009) suggesting that VSTM retains the products of visual imagery and visual perception but is a separate component of a multicomponent VSWM system that supports online cognition.

**Acknowledgements** We are grateful to Stephen M. Kosslyn for his contribution and inputs on the design and the interpretation of the data collected in the three experiments. We also wish to thank Graeme Standing for his help in recruiting participants and collecting data. Robert Logie's research is supported by the Leverhulme Trust, Grant F/00 158/W.

#### References

- Aleman, A., Schutter, D. J. L. G., Ramsey, N. F., van Honk, J., Kessels, R. P. C., Hoogduin, J. M., et al. (2002). Functional anatomy of top-down visuospatial processing in the human brain: Evidence from rTMS. Cognitive Brain Research, 14, 300-302
- Andrade, J., Kemps, E., Werniers, Y., May, J., & Szmalec, A. (2002). Insensitivity of visual short-term memory to irrelevant visual information. *Quarterly Journal of Experimental Psychology*, 55A, 753–774.
- Baddeley, A. D. (1986). Working memory. Oxford: Oxford University Press.
- Baddeley, A. D. (2007). Working memory, thought and action. Oxford: Oxford University Press.
- Baddeley, A. D., & Andrade, J. (2000). Working memory and vividness of imagery. *Journal of Experimental Psychology:* General, 129, 126–145.
- Borst, G., & Kosslyn, S. M. (2008). Visual mental imagery and visual perception: Structural equivalence revealed by scanning processes. *Memory & Cognition*, 36, 849–862.
- Cowan, N. (2005). Working memory capacity. New York: Psychology Press.
- Darling, S., Della Sala, S., & Logie, R. H. (2007). Behavioural evidence for separating components of visuo-spatial working memory. *Cognitive Processing*, 8, 175–181.
- Darling, S., Della Sala, S., & Logie, R. H. (2009). Segregation within visuo-spatial working memory: Appearance is stored separately from location. *Quarterly Journal of Experimental Psychology*, 62, 417–425.

- Dean, G. M., Dewhurst, S. A., Morris, P. E., & Whittaker A. (2005). Selective interference with the use of visual images in the symbolic distance paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*, 1043–1068.
- Dean, G. M., Dewhurst, S. A., & Whittaker, A. (2008). Dynamic visual noise interferes with storage in visual working memory. *Experimental Psychology*, 55, 283–289.
- Della Sala, S., Gray, C., Baddeley, A., Allamano, N., & Wilson, L. (1999). Pattern span: A tool for unwielding visuo-spatial memory. *Neuropsychologica*, 37, 1189–1199.
- Dent, K. (2010). Dynamic visual noise affects visual short-term memory for surface color but not spatial location. *Experimental Psychology*, 57, 17–26.
- Engelkamp, J., Mohr, G., & Logie, R. H. (1995). Memory for size relations and selective interference. European Journal of Cognitive Psychology, 7, 239–260.
- Ganis, G., Thompson, W. L., & Kosslyn, S. M. (2004). Brain areas underlying visual mental imagery and visual perception: An fMRI study. Cognitive Brain Research, 20, 226–241.
- Ganis, G., Thompson, W. L., & Kosslyn, S. M. (2009). Visual mental imagery: More than "seeing with the mind's eye.". In J. R. Brockmole (Ed.), *Representing the visual world in memory* (pp. 215–249). Hove, UK: Psychology Press.
- Kosslyn, S. M. (1983). Mental representation. In J. R. Anderson & S. M. Kosslyn (Eds.), *Tutorials in learning and memory: Essays in honor of Gordon Bower*. San Francisco, CA: Freeman.
- Kosslyn, S. M. (1994). Imagery and brain: The resolution of the imagery debate. Cambridge, MA: MIT Press.
- Kosslyn, S. M., Alpert, N. M., Thompson, W. L., Maljkovic, V., Weise, S. B., Chabris, C. F., et al. (1993). Visual mental imagery activates topographically organized visual cortex: PET investigations. *Journal of Cognitive Neuroscience*, 5, 263–287.
- Kosslyn, S. M., & Thompson, W. L. (2003). When is early visual cortex activated during visual mental imagery? *Psychological Bulletin*, 129, 723–746.
- Kosslyn, S. M., Thompson, W. L., & Ganis, G. (2006). The case for mental imagery. New York: Oxford University Press.
- Logie, R. H. (1981). The symbolic distance effect: A study of internal psychophysical judgements. Unpublished doctoral thesis, University College, University of London.
- Logie, R. H. (1986). Visuo-spatial processing in working memory. Ouarterly Journal of Experimental Psychology, 38A, 229–247.
- Logie, R. H. (1995). Visuo-spatial working memory. Hove, U.K.: Erlbaum.
- Logie, R. H. (2003). Spatial and visual working memory: A mental workspace. In D. E. Irwin & B. H. Ross (Eds.), *Cognitive vision: The psychology of learning and motivation, vol. 42* (pp. 37–78). San Diego, CA: Academic Press.
- Logie, R. H. (2011a). The visual and the spatial of a multicomponent working memory. In A. Vandierendonck & A. Szmalec (Eds.), Spatial working memory (pp. 19–45). Hove, UK: Psychology Press
- Logie, R. H. (2011b). The functional organisation and the capacity limits of working memory. Current Directions in Psychological Science.
- Logie, R. H., & van der Meulen, M. A. (2009). Fragmenting and integrating visuo-spatial working memory. In J. R. Brockmole (Ed.), *Representing the visual world in memory* (pp. 1–32). Hove, U.K.: Psychology Press.
- Logie, R. H., Brockmole, J. R., & Jaswal, S. (2011). Feature binding in visual working memory is unaffected by task-irrelevant changes of location, shape and color. *Memory & Cognition*, 39, 24–36.
- Logie, R. H., Brockmole, J. R., & Vandenbroucke, A. (2009). Bound feature combinations in visual short-term memory are fragile but influence long-term learning. *Visual Cognition*, 17, 160–179.



- Logie, R. H., Della Sala, S., Wynn, V., & Baddeley, A. (2000). Visual similarity effects in immediate verbal serial recall. *Quarterly Journal of Experimental Psychology*, 53A, 626–646.
- Logie, R. H., Zucco, G., & 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.
- McConnell, J., & Quinn, J. G. (2004). Cognitive mechanisms of visual memories and visual images. *Imagination, Cognition and Personality*, 23, 201–207.
- Paivio, A. (1971). Imagery and verbal processes. New York: Holt, Rinehart and Winston.
- Pearson, D. G. (2001). Imagery and the visuo-spatial sketchpad. In J. Andrade (Ed.), *Working memory in perspective* (pp. 33–59). Hove, UK: Psychology Press.
- Pearson, D. G., Logie, R. H., & Gilhooly, K. J. (1999). Verbal representations and spatial manipulation during mental synthesis. *European Journal of Cognitive Psychology*, 11, 295–314.
- Phillips, W. A., & Christie, D. F. M. (1977). Components of visual memory. Quarterly Journal of Experimental Psychology, 29, 117–133.
- Quinn, J. G., & McConnell, J. (1996). Irrelevant pictures in visual working memory. *Quarterly Journal of Experimental Psychol*ogy, 49A, 200–215.
- Quinn, J. G., & McConnell, J. (2006). The interval for interference in conscious visual imagery. *Memory*, 14, 241–252.
- Quinn, J. G., & Ralston, G. E. (1986). Movement and attention in visual working memory. *Quarterly Journal of Experimental Psychology*, 38A, 689–703.
- 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.
- Salway, A. F. S., & Logie, R. H. (1995). Visuo-spatial working memory, movement control, and executive demands. *Quarterly Journal of Experimental Psychology*, 41A, 107–122.

- Schacter, D. L. (1996). Searching for memory: The brain, the mind, and the past. New York: Basic Books.
- Smyth, M. M., & Waller, A. (1998). Movement imagery in rock climbing: Patterns of interference from visual, spatial and kinaesthetic secondary tasks. *Applied Cognitive Psychology*, 12, 145–157.
- Snodgrass, J., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory*, 6, 174–215.
- Thompson, W. L., Kosslyn, S. M., Hoffman, M. S., & van der Kooij, K. (2008). Inspecting visual mental images: Can people "see" implicit properties as easily in imagery and perception? *Memory & Cognition*, 36, 1024–1032.
- Todd, J. J., & Marois, R. (2004). Capacity limit of visual short-term memory in the 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, and Behavioral Neuroscience, 5, 144–155.
- van der Meulen, M., Logie, R. H., & Della Sala, S. (2009). Selective interference with image retention and generation: Evidence for the workspace model. *Quarterly Journal of Experimental Psychology*, 62, 1568–1580.
- Vogel, E. K., Woodman, G. F., & Luck, S. J. (2001). Storage of features, conjunctions, and objects in visual working memory. *Journal of Experimental Psychology: Human Perception and Performance*, 27, 92–114.
- Zimmer, H. D., & Speiser, H. R. (2002). The irrelevant picture effect in visuospatial working-memory: Fact or fiction? *Psychologische Beitrage*, 44, 223–247.
- Zimmer, H. D., Speiser, H. R., & Seidler, B. (2003). Spatio-temporal working memory and short-term object location tasks use different memory mechanisms. Acta Psychologica, 114, 41–65.

