The interplay of spatial attention and memory: evidence for memory-driven spatial compatibility effects, but not for a facilitating role of attention in memory retrieval

Ivan Vankov

Published online: 14 June 2011
© Psychonomic Society, Inc. 2011

Abstract The visual indexing theory proposed by Zenon Pylyshyn (*Cognition*, 32, 65–97, 1989) predicts that visual attention mechanisms are employed when mental images are projected onto a visual scene. Recent eye-tracking studies have supported this hypothesis by showing that people tend to look at empty places where requested information has been previously presented. However, it has remained unclear to what extent this behavior is related to memory performance. The aim of the present study was to explore whether the manipulation of spatial attention can facilitate memory retrieval. In two experiments, participants were asked first to memorize a set of four objects and then to determine whether a probe word referred to any of the objects. The results of both experiments indicate that memory accuracy is not affected by the current focus of attention and that all the effects of directing attention to specific locations on response times can be explained in terms of stimulus–stimulus and stimulus–response spatial compatibility.

Keywords Attention and memory · Spatial memory

Introduction

The visual indexing theory proposed by Zenon Pylyshyn (1989) stipulates that there is a preattentive mechanism in the visual system for individuating features and indexing their locations within the visual field. Once such location

indexes are formed, they serve to bind together the features of an object and maintain its identity. The indexes are also used to refer to information that is bound to a certain location in the visual field and to guide visual attention.

One of the major implications of visual indexing theory is that visual indexes are used to direct attention not only when directly perceivable stimuli are attended, but also when mental images are projected onto some visual scene (Pylyshyn, 1998). Recent eye-tracking studies have rendered support for this hypothesis by showing that people look at empty spaces where requested information has previously been presented (Hoover & Richardson, 2008; Laeng & Teodorescu, 2002; Richardson & Spivey, 2000; Spivey & Geng, 2001). For example, Spivey and Geng asked participants to memorize a set of four simple shapes displayed at the corner cells of a 3×3 grid. The display then vanished and reappeared with one of the objects missing. When asked about the properties of the missing object, participants spontaneously looked at the blank region where the object had previously been presented. Laeng and Teodorescu found not only that participants looked at empty spaces when imagining a previously perceived scene, but also that the order of scanning these locations was correlated with the original order during perception. These results were interpreted in terms of the visual indexing theory: The missing object identity had been bound to a specific location in the visual field, and attempting to access any information about the object resulted in an attention shift toward that location.

An important question that has remained underexplored by the eye-tracking studies is whether the attention shifts toward the locations of previously seen objects are related to memory performance. In a recent article, Ferreira, Apel, and Henderson (2008) suggested that directing attention to an empty location should facilitate memory retrieval for any

I. Vankov (✉)
Department of Cognitive Science and Psychology,
New Bulgarian University,
21 Montevideo Street,
Sofia 1618, Bulgaria
e-mail: i.i.vankov@cogs.nbu.bg

information associated with that location, including visual and linguistic details. This hypothesis was justified by the study of Laeng and Teodorescu (2002), which reported that the strength of the relatedness of the eye movement patterns generated during perception of a scene and its recollection in front of a blank screen predicts memory accuracy. Laeng and Teodorescu, however, admitted that this finding is only correlational and does not constitute unambiguous evidence that eye movements subserve memory processes. Other studies have failed to show any relationships between looking behavior and memory performance (Hoover & Richardson, 2008; Richardson & Spivey, 2000). Richardson, Altmann, Spivey, and Hoover (2009) argued that the existing data are not enough to resolve this controversy and that new experiments are needed in which eye movements are manipulated as an independent variable.

The aim of the present study was to investigate whether the reorienting of attention to locations where memorized objects have previously been presented would facilitate retrieving information about them. If attention shifts are, indeed, employed in the retrieval process, one would expect their processing costs to be reflected in the speed and accuracy with which objects are recalled. In other words, the question is whether one would need more time and would make more errors when retrieving information that was presented away from the current focus point, as compared with when it was at the same location. The hypothesis put forward by Ferreira et al. (2008) predicts that memory performance will be improved in the cases in which spatial attention is directed to the location where the requested information has previously been presented.

Experiment 1

The goal of this experiment was to investigate whether the location of previously presented objects is related to

the speed and accuracy with which they are later recalled. During the experiment, participants viewed images of four objects located on a 2×2 grid and tried to remember them. Afterward, the objects disappeared, and a single word appeared in one of the four cells of the grid. The task of the participants was to press a button if the word denoted some of the objects and to do nothing otherwise (a go/no-go task).

Method

Participants Twenty-four participants (14 females) took part in the experiment. Their average age was 25.79 years (age range from 15 to 47, $SD = 7.28$). Participants were paid for participation.

Stimuli The stimulus set was constructed out of 128 images of various objects and 32 probe words denoting 32 of the objects. The images and words were taken from a picture-naming database with 520 pictures and their dominant responses (Szekely et al., 2004). For ease of exposition, throughout this article, objects denoted by words will be referred to as *target objects*. Each stimulus consisted of four pictures of objects, one of them being a target object (Fig. 1a), and a corresponding probe word (dominant response to a picture) denoting the target object (Fig. 1b). The number of stimuli was 32. The target-object—word pairs were controlled for image agreement ($M = 5.61$, $SD = 0.65$), name agreement ($M = 85.43$, $SD = 11.09$), word length ($M = 5.28$, $SD = 0.72$), and objective frequency ($M = 1.76$, $SD = 0.35$). All images were controlled for visual complexity. An equal number of fillers were compiled using 128 different images and 32 unrelated words from the same database.

All images used in the experiment were resized to 200×200 pixels and were displayed on a 2×2 grid that covered the whole screen. The grid was used in order to setup a rich spatial context that would facilitate the formation of visual

Fig. 1 **a** Sample stimulus. The guitar is a target object. **b** Sample stimulus. The relative position of the probe word with respect to the position of the target object determines the condition. The position of

the probe in the same (S) condition is enlarged. The circled letters in the center of the grid indicate the condition; they were not displayed in the real stimuli. The actual words used in the study were Bulgarian

indexes binding together the semantic and spatial characteristics of presented objects (Spivey & Geng, 2001). According to Ferreira et al. (2008), the tight integration of conceptual, visual, linguistic, and spatial representations is a prerequisite for the facilitating effect of looking at nothing on memory retrieval.

Design The independent variable was the relative location of the probe word with respect to the target object. Thus, four possible conditions were defined:

S (Same): The word appeared in the same grid location as the target object.

H (Horizontal): The word appeared in the position to the left or right of the target object.

V (Vertical): The word appeared in the position below or above the target object.

D (Diagonal): The word appeared in the position diagonal to the position of the target object.

The dependent measure in this experiment was response time (RT).

Procedure Four lists were generated, each of them containing all the stimuli and the same number of fillers. Each stimulus was included in each of the lists in a different condition by varying the position of the word (the object images were always located at the same grid locations). Target object locations were counterbalanced across stimuli. The lists were pseudorandomized, so that the same condition appeared no more than 2 consecutive times and there were at least 15 trials between two appearances of the same stimulus, no matter what the list order was. The order of stimuli and filler trials was also pseudorandomized, so that a stimulus or filler trial would appear no more than 2 consecutive times. Each participant went through all the lists in varying order. Thus, the total number of trials was 256, including 128 filler trials.

Participants were tested in a sound-proof booth. The stimuli were presented on a 17-in. computer monitor with a resolution of 800×600 pixels. The experimental session started with 10 practice trials, none of which appeared in the experimental part. Each trial began with four objects displayed simultaneously in the four cells of the grid for 2,000 ms. Then the objects disappeared, and the grid stayed blank for 300 ms. After that, a word was presented in one of the grid cells for 2,500 ms or until a participant's response was generated by pressing a predefined button on a serial response button box. The participants were instructed to react as soon as possible if the presented word denoted any of the objects shown in the previous screen. The intertrial interval was 1,500 ms. RT was measured from the onset of each word stimulus. Stimulus presentation and response recordings were controlled by

E-Prime software (Schneider, Eschman, & Zuccolotto, 2002). The experiment took about 25 min.

Results and discussion

One participant had more than 40% errors, and his data were excluded from the analysis. Trials with errors were excluded from the analysis (5.0%). RTs below 200 ms and those lying more than ± 2.5 standard deviations from the RT mean were removed as well, which resulted in the removal of 2.7% of overall responses. Thus, a total of 92.4% of the originally collected nonfiller RT data were included in further analyses.

Each participant went through four lists containing the same stimuli presented in different experimental conditions. In order to make sure that the multiple presentations of the same stimulus did not alter the effect of the independent variable under consideration, we conducted a 4×4 repeated measures ANOVA on participants' RT means, with the list serial order and the relative locations of the probe word as factors. It failed to show an interaction between the list serial order and the relative location factor, $F(9, 198) = 0.863$, $p > .5$, so the data were collapsed across all lists for further analysis.

A one-way repeated measures ANOVA was performed for participants' RT means and revealed significant main effect of the word location factor, $F(3, 66) = 2.71$, $p < .05$ (see Fig. 2). The analysis of error rates failed to show an effect, $F(3, 66) = 0.61$, $p = .61$.

A set of planned orthogonal contrast tests revealed a significant difference between RTs in the S and H conditions, $F(1, 22) = 5.46$, $p < .05$, and marginally significant difference between the S and D conditions, $F(1, 22) = 3.78$, $p = .065$. The difference between the S and V conditions was not significant, $F(1, 22) = 2.88$, $p > .1$. Another planned contrast test revealed a significant difference between RTs in the S condition and the mean for the other conditions, $F(1, 22) = 4.80$, $p < .05$.

Fig. 2 Experiment results. Participants were faster when the target object and the word appeared in the same location (S), as compared with the means for the other conditions (H, in the same grid row; V, in the same grid column; D, diagonal to each other). The error bars stand for confidence intervals

The results demonstrated that the relative location of the target objects affected task performance, although object locations were irrelevant to the task. Participants responded more slowly when they had to retrieve an object that had been presented away from the position of the probe word. This finding speaks in favor of the hypothesis that memory retrieval is facilitated by directing attention to the empty space where requested information has previously been presented. There are, however, two aspects of the data that preclude such an interpretation of the results. First, the analysis did not reveal an effect on error rates. This result is consistent with the eye-tracking studies that failed to show any relation between looking at nothing and memory accuracy. The second problem is that the planned orthogonal comparisons failed to yield all the predicted effects. In particular, there was no significant difference between the S and V conditions.

A closer look at the data suggests an alternative explanation of the results that can account for the lack of an effect on error rates and the insignificant difference between the S and V conditions. It is evident from Fig. 2 that the shortest RTs occurred in the S and V conditions, in which both the target object and the probe word were displayed on the same side (left or right) of the screen. Therefore, it is possible to attribute the observed effects to the spatial compatibility between the locations of the target object and the probe word. On this view, the participants responded more quickly when the probe word appeared on the same side of the screen as the target object and reacted more slowly when it appeared on the opposite side. There was no effect on error rates, because the task required the execution of the same action in all the experimental conditions. In order to test this stimulus–stimulus spatial compatibility explanation of the results, we conducted an additional contrast test that aimed to compare the S and V conditions against the H and D conditions. It revealed a significant difference, $F(1, 22) = 4.64$, $p < .05$.

In general, Experiment 1 did not manage to yield conclusive evidence that the direction of attention toward empty locations critically constrains memory performance. The results, however, clearly indicate that the retrieval of remembered objects activates irrelevant spatial information about the setting in which they have originally been perceived. The additional analysis suggests that the performance implications of memory-driven attention can be explained in terms of spatial compatibility and do not necessarily entail a functional role of spatial attention in memory retrieval. In order to further investigate this issue, we conducted a second experiment whose aim was to dissociate more clearly the impact of redirecting attention on memory performance from spatial compatibility effects.

Experiment 2

The aim of Experiment 2 was to extend the findings of Experiment 1 by showing that retrieving a previously seen object not only activates irrelevant spatial information, but also can induce a Simon-like effect (Simon & Rudell, 1967). The prediction was that the people would respond more quickly and more accurately when the location of the remembered object was spatially compatible with the response action required by the task. The second goal of the experiment was to test again whether the manipulation of the relative location of the target object would affect the accuracy of memory retrieval.

Method

Participants Twenty participants (15 females) took part in the experiment. Their average age was 22.50 years (age range from 19 to 28, $SD = 2.58$). Participants were paid for participation.

Stimuli The stimulus set was the same as in Experiment 1.

Design The same four conditions standing for the relative location of the probe word with respect to the target object were used.

Procedure The procedure was the same as in Experiment 1, except for the task. Participants were instructed to respond by pressing the leftmost button of the response box with the left hand when the word denoted any of the objects and appeared in the left part of the screen, press the rightmost button with their right hand when the word denoted any of the objects and appeared in the right part of the screen, or press the middle button when the word did not denote any of the objects. The response box was located just in front of the screen, and its width roughly corresponded to the distance between the centers of the left- and right-side grid cells. The experiment took about 30 min.

Results and discussion

One participant had more than 40% errors, and her data were excluded from the analysis. All other participants had fewer than 20% errors, including filler trials. Filler trials were not included in the analysis. Trials with errors were excluded from the RT analysis (8.1%). RTs lying more than ± 2.5 standard deviations from the mean were removed as well, which resulted in the removal of 2.1% of overall responses. Thus, a total of 89.7% of the originally collected nonfiller RT data were included in the RT analyses.

A one-way repeated measures ANOVA was performed for participants' RT means and error rates (ERs) and revealed significant main effects [RT, $F(3, 54) = 11.24$, $p < .001$; ER, $F(3, 54) = 13.81$, $p < .001$]. A set of pairwise Tukey HSD posthoc tests revealed significant differences between the S and H, S and D, V, and H, and V and D conditions (all $ps < .05$). The RT results are shown in Fig. 3.

A planned contrast test was used to compare conditions in which the locations of the target object and the probe word were spatially compatible with respect to the response discrimination axis against conditions in which it was not. In other words, we compared the mean of the S and V conditions, in which the target object and the probe word were on the same side of the screen (left or right), against the mean of the H and D conditions. The difference was statistically significant [RT, $F(1, 18) = 25.32$, $p < .001$; ER, $F(1, 18) = 37.85$, $p < .001$].

The results indicate that a strong stimulus–response compatibility effect was found. Participants responded more quickly when the location of the target object was spatially compatible with the correct response action. The overall pattern of results is consistent with the response discrimination account of the Simon effect (Ansorge & Wühr, 2004). According to this theory, stimulus-triggered response activation shows up in Simon effects only when stimulus locations match the top-down selected spatial codes used to discriminate between alternative responses. Hence, the response discrimination account predicts that no spatial compatibility effect will be found when the S and V conditions are compared, as well as the H and D conditions, because in these pairs the position of the probe word is varied only along the vertical axis, which is irrelevant to choosing the correct response action. That was exactly the pattern found in Experiment 2 (Fig. 3).

The lack of difference between the S and V conditions replicated the pattern of results found in Experiment 1 and suggests that all the effects on RTs are due to spatial compatibility only. In other words, the directing of attention toward the location where requested information has

previously been presented does not seem to facilitate memory retrieval in terms of RTs.

Two types of errors were analyzed separately in order to set apart the effects that were due to the memory and to the spatial aspects of the task. A *memory* error was counted when a participant wrongly pressed the middle button, as if the word did not denote any of the objects. An error was regarded as a *spatial* one if a participant responded by pressing the right button instead of the left one and vice versa.

A one-way repeated measures ANOVA was performed for participants' memory and spatial error rates, computed separately for each condition as a percentage of the total number of target trials. The results are presented in Fig. 4. No main effect and no significant difference between conditions were found for memory error rates. There was a main effect of relative location for spatial errors, $F(3, 54) = 12.95$, $p < .001$, as well as significant differences between the S and H, $F(1, 18) = 18.56$, $p < .001$, H and V, $F(1, 18) = 20.94$, $p < .001$, and V and D, $F(1, 18) = 13.47$, $p < .01$, conditions. A set of nonparametric Friedman tests for related samples were performed on spatial error numbers, since the number of spatial errors in some of the conditions was very low. A significant difference between all conditions was found, $\chi^2(3, 19) = 31.62$, $p < .001$, as well as between the S and H, $\chi^2(1, 19) = 15.00$, $p < .001$, H and V, $\chi^2(1, 19) = 11.28$, $p < .001$, and V and D, $\chi^2(1, 19) = 8.07$, $p < .01$, conditions.

The failure to find an effect of relative location on memory error rates suggests that the spatial aspect of the task did not influence memory accuracy. Participants made an equal number of errors in recalling objects that were located on either side of the response discriminating axis.

General discussion

The two experiments reported in this article explored the interplay between memory and attention processes in the

Fig. 3 Experiment response time (RT) results. Participants were significantly slower when the probe word was located on the opposite horizontal side of the screen, as compared with the object it denoted. Error bars represent confidence intervals

Fig. 4 Error rates results. The pattern of spatial errors closely resembles the pattern of response times displayed in Fig. 3

context of a task that required recalling a previously seen object prompted by a probe word.

Experiment 1 revealed that participants responded more slowly when recalling an object that was presented away from the cued position. This result favors the view that irrelevant spatial information about remembered objects is automatically encoded and that it can affect performance in a task that does not involve any kind of spatial discrimination—a memory recall go/no-go task. The lack of effect on error rates, however, suggests that memory accuracy is not affected by directing attention toward empty locations of remembered objects.

The goal of Experiment 2 was to further investigate the performance implications of redirecting attention when previously seen objects are recalled. It confirmed the finding of Experiment 1 that the presentation locations of remembered objects, although irrelevant for the task, may affect RTs through stimulus–response compatibility. Although other studies have already reported Simon-like effects by memory retrieval (Hommel, 2002; Wühr & Ansorge, 2007; Zhang & Johnson, 2004), the novelty of our finding is in demonstrating that irrelevant spatial cues generated by memorized objects may affect performance even in the presence of relevant spatial cues. In Experiment 2, the location of the memorized object induced a stimulus–response compatibility effect even though participants were instructed to choose a response action corresponding to the location of the visually available probe word.

The present study failed to show any evidence in support of the hypothesis that directing attention toward locations where requested information has previously been presented facilitates retrieving memory representations (Ferreira et al., 2008). The outcomes of both experiments suggest that memory accuracy is not affected by the manipulation of presentation location and that all the effects of directing attention can be explained in terms of compatibility between the spatial codes generated by memorized objects, probes, and response actions. Our results are, therefore, consistent with the view defended by Richardson et al. (2009), according to which looking at nothing does not support memory retrieval but is a side effect of offloading cognitive load onto the environment. In the real world, objects do not tend to disappear suddenly, and looking back to their locations will, in most cases, bring them into sight and help one to retrieve requested information about them. However, if the requested information is no longer available, looking at nothing not only cannot improve memory performance, but also can worsen general cogni-

tive performance, due to the interference of task-irrelevant spatial codes.

Author Note This work was supported by the Project ANALOGY: Humans—the Analogy-Making Species, financed by the FP6 NEST Programme of the European Commission. (Contr. No 029088).

References

- Ansorge, U., & Wühr, P. (2004). A response-discrimination account of the Simon effect. *Journal of Experimental Psychology: Human Perception and Performance*, 30, 365–377.
- Ferreira, F., Apel, J., & Henderson, J. (2008). Taking a new look at looking at nothing. *Trends in Cognitive Sciences*, 11, 405–410.
- Hommel, B. (2002). Responding to object files: automatic integration of spatial information revealed by stimulus–response compatibility effects. *Quarterly Journal of Experimental Psychology*, 55A, 567–580.
- Hoover, M. A., & Richardson, D. C. (2008). When facts go down the rabbit hole: contrasting features and objecthood as indexes to memory. *Cognition*, 108, 533–542.
- Laeng, B., & Teodorescu, D. S. (2002). Eye scanpaths during visual imagery re-enact those of perception of the same visual scene. *Cognitive Science*, 26, 207–231.
- Pylyshyn, Z. (1989). The role of location indexes in spatial perception: a sketch of the FINST spatial-index model. *Cognition*, 32, 65–97.
- Pylyshyn, Z. (1998). The role of visual indexes in spatial vision and imagery. In R. Wright (Ed.), *Visual attention* (pp. 215–231). New York: Oxford University Press.
- Richardson, D. C., Altmann, G. T., Spivey, M. J., & Hoover, M. A. (2009). Much ado about eye movements to nothing: a response to Ferreira et al.: taking a new look at looking at nothing. *Trends in Cognitive Sciences*, 13, 235–236.
- Richardson, D. C., & Spivey, M. J. (2000). Representation, space and Hollywood Squares: looking at things that aren't there anymore. *Cognition*, 76, 269–295.
- Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime user's guide*. Pittsburgh: Psychology Software Tools Inc.
- Simon, J., & Rudell, A. (1967). Auditory S–R compatibility: the effect of an irrelevant cue on information processing. *Journal of Applied Psychology*, 51, 300–304.
- Spivey, M. J., & Geng, J. J. (2001). Oculomotor mechanisms activated by imagery and memory: eye movements to absent objects. *Psychological Research*, 65, 235–241.
- Szekely, A., Jacobsen, T., D'Amico, S., Devescovi, A., Andonova, E., Herron, D., . . . Bates, E. (2004). A new on-line resource for psycholinguistic studies. *Journal of Memory and Language*, 51, 247–250.
- Wühr, P., & Ansorge, U. (2007). A Simon effect in memory retrieval: evidence for the response-discrimination account. *Psychonomic Bulletin & Review*, 14, 984–988.
- Zhang, J., & Johnson, M. (2004). A memory-based, Simon-like, spatial congruence effect: evidence for persisting spatial codes. *Quarterly Journal of Experimental Psychology*, 57A, 419–436.