Mind the Gap: The Cost of Looking at Nothing, or the Performance Implications of Memory-induced Attention Shifts

Ivan Vankov (i.i.vankov@cogs.nbu.bg)

Central and East European Center for Cognitive Science
New Bulgarian University, 21 Montevideo Street, Sofia 1618, Bulgaria

Abstract

The visual indexing theory proposed by Zenon Pylyshyn predicts that visual attention mechanisms are employed when projecting mental images 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 remained unclear what performance implications this behavior has. The present study aimed to explore how the relative location of presentation affects performance in tasks involving recalling whether an object has already been seen. Experiment 1 demonstrated that recalling objects which have been presented away from the current focus point results in slower reaction times. Experiment 2 showed that a spatial compatibility effect arises if the relative location of the memorized object is incompatible with a spatial task at hand.

Keywords: attention; memory; space; visual imagery; spatial compatibility; Simon effect

Introduction

The visual indexing theory proposed by Zenon Pylyshyn (Pylyshyn, 1989) stipulates that there is a pre-attentive 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 which was bound to a certain location in the visual field and guide visual attention.

One of the major implications of visual indexing theory is that visual indexes are used to direct attention not only when attending directly perceivable stimuli, but also when projecting mental images 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 (Richardson & Spivey, 2000; Spivey & Geng, 2001; Hoover & Richardson, 2008). For example, Spivey and Geng (2001) asked participants to memorize a set of four simple shapes displayed at the corner cells of a 3x3 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 been previously presented. 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 towards that location.

An important question that remained unexplored by the eye-tracking studies was whether the attention shifts towards locations of previously seen objects were related to memory performance. If attention shifts are employed in the retrieval process, then one would expect that their processing costs would be reflected in the speed with which objects are recalled. In other words, the question is whether one would need more time to retrieve information which was presented away from the current focus point compared to when it was at the same location. The results of previous studies focused on the interplay of memory and visual imagery suggest that indeed the relative location of a target item within a visual image affects memory performance. The seminal work of Kosslyn, Ball and Reiser (1978) demonstrated that people needed more time to mentally scan further distances across visual images, even when the same amount of material fell between the initial focus point and the target. However, Kosslyn and colleagues explicitly asked the participants to base their responses on their visual mental images and when this instruction was omitted, the effect of distance disappeared. Thus, the first goal of the present study was to investigate whether it is possible to show an effect of the relative location of a memorized item on performance in a purely memory task even if no explicit instructions for employing visual imagery are given.

A second issue which remains is what consequences the memory-driven redirection of attention would have if subjects are given a task requiring them to respond on the basis of available spatial information. To this end, a task was selected which has both memory retrieval and spatial judgment components. If the memory retrieval aspect of the task induces attention shifts, then they may happen to be incompatible with some of the responses and thus a spatial compatibility effect would be observed.

A spatial compatibility effect, usually referred to as a Simon effect (Simon & Rudell, 1967), denotes faster and more accurate responses when the task-irrelevant stimulus' location is compatible with a response action than when it is not. Several researches have advocated that the Simon effect is related to attention shifts (Nicoletti & Umilta, 1994; Rubichi at el, 1997). According to the attention shift account, directing attention to a location in space facilitates actions towards that position. Hence, there are good reasons to expect that attention shifts generated in the process of memory retrieval may cause Simon-like effects.

Simon-like effects on memory retrieval have already been reported (Hommel, 2002; Zhang & Johnson, 2004; Wühr & Ansorge, 2007). For example, Hommel (2002) asked participants to first memorize a display of four stimuli that varied in color, shape, and location. Then, a cue stimulus required to retrieve one of the stimuli on the basis of its color, and to respond to stimulus shape by pressing a left or right key. The irrelevant horizontal location of the to-be-retrieved stimulus in the encoding display produced a Simon effect in subsequent responses to the cue.

The existing studies of memory-induced Simon-like effects have demonstrated that a spatial compatibility effect could be invoked not only by a present spatially incongruent stimulus, but also by one which is retrieved from memory. The second aim of this study was to investigate whether a similar effect would be found even in the presence of directly available spatial cues which determine the correct response action.

Experiment 1

The goal of this experiment was to investigate whether the location of previously presented objects is related to the speed with which they are later recalled. During the experiment, participants viewed images of four objects located on a 2x2 grid and tried to remember them. Afterwards the objects disappeared and a single word appeared in one of the 4 cells of the grid. The task of the participants was to press a button if the word denoted some of the objects and do nothing otherwise (a go/no-go task).

Method

Participants 24 participants (14 females) took part in the experiment. Their average age was 25.79 (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 ‘cuing’ 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)[1]. For ease of exposition, throughout this paper 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 (Figure 1a), and a corresponding cuing word (dominant response of a picture) denoting the target object (Figure 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 x 200 pixels and displayed on a 2x2 gird which covered the whole screen. The justification for presenting the objects within a grid came from the study Spivey and Geng (2001), which showed that such a spatial context enhances attention shifts towards the empty spaces where objects have been previously presented.

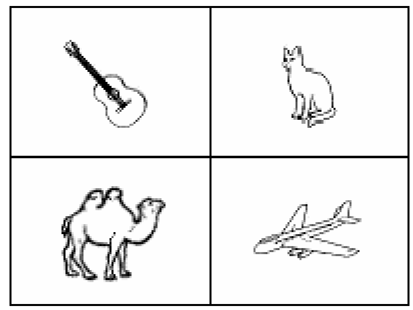

Figure 1a: Sample stimulus. The guitar is a target object.

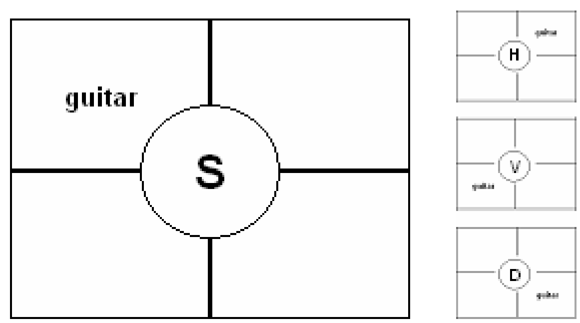

Figure 1b: Sample stimulus. The relative position of the cueing word with respect to the position of the target object determines the condition. Enlarged is the position of the cue in condition S. The circled letters in the center of the grid indicate the condition; they were not displayed in the real stimuli. The actual cuing words used in the study were Bulgarian.

Design The independent variable was the relative location of the cuing 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

1 For more details on participants, procedures, and pictorial stimuli in the picture naming norming study, see Szekely et al. (2004) or visit the online database at http://www.crl.ucsd.edu/~aszekely/ipnp/.

The dependent measure in this experiment was reaction time.

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 pseudo-randomized, 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 pseudo-randomized, so that a stimulus or filler trial would appear no more than 2 consecutive times. Each subject 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 being presented on a 17” computer monitor with a resolution of 800x600 pixels. The experimental session started with 10 practice trials, none of which appeared in the experimental part. Each trial began with 4 objects displayed simultaneously in the four cells of the grid for 2000 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 2500 ms or until a subject’s response was generated by pressing a predefined button of 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 at the previous screen. The inter-trial interval was 1500 ms. Reaction time (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 minutes.

Results and Discussion

One subject had more than 40% errors and his data were excluded from the analysis. Trials with errors were excluded from the analysis (5.0%). Response times 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 non-filler RT data were included in further analyses.

As mentioned above, each subject went through 4 lists containing the same stimuli. A 4x4 repeated measures ANOVA performed on subject RT means failed to show an interaction between the list serial number and condition ($F(9, 189) = 0.863$; $p > .5$), so the data was collapsed across all lists for further analysis.

A one-way repeated measures ANOVA was performed for subjects RT means and revealed significant main effect of the word location factor ($F(3, 63) = 2.71$; $p < .05$) – Figure 2.

A set of planned orthogonal contrast tests revealed significant difference between reaction times in condition S and in condition H ($F(1, 22) = 5.46$, $p < .05$) and marginally significant difference between conditions S and D ($F(1, 22) = 3.778$; $p = .065$). The difference between conditions S and V was not significant ($F(1, 22) = 2.88$; $p > .1$). Another planned contrast test revealed a significant difference between reaction times in condition S and the mean of the other conditions ($F(1, 22) = 4.80$; $p < .05$).

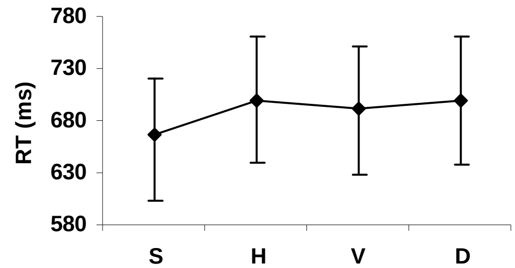

Figure 2: Experiment results. Subjects were faster when the target object and cuing word appeared in the same location (S), compared to mean of the other conditions (H – in the same grid row, V – in the same grid column, D – diagonally to each other). The error bars stand for confidence intervals.

The results demonstrated that the relative location of the target objects affected task performance, although object locations were irrelevant to the task. Subjects responded more slowly when they had to retrieve an object which had been presented away from the position of the cue. This result bears two alternative explanations. First, it is possible that in condition S the location of the cuing word directly activated the corresponding object representation, which was bound to the same location through visual indexing. Another possibility is to attribute the slower mean reaction times in the distant conditions to an attention shift, which was required in order to retrieve an object that had been presented at a different location. The eye-tracking results of Richardson and Spivey (2000), Spivey and Geng (2001) and Hover and Richardson (2008) are consistent with the latter explanation.

Although the main effect of relative location was significant, its size effect was rather small ($h^2 = .11$). One reason for this could be that the manipulation of the factor was not strong enough. The experiment lacked enough power to reveal more clearly an effect which was expected to be subtle.

Experiment 2

Experiment 2 aimed to investigate if retrieving target objects which had been presented away from the cuing word would invoke attention shifts and whether this memory-driven redirection of attention will affect performance in a spatial judgment task. Also it aimed to differentiate performance effects that were due to memory retrieval from those due to spatial incompatibility. The same stimuli from Experiment 1 were used, but the task was modified, so that it had both a memory and a spatial component.

Method

Participants 20 participants (15 females) took part in the experiment. Their average age was 22.50 (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 cuing word with respect to the target objects were used. Unlike Experiment 1, error rates were also considered as a dependent measure.

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 cuing 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; 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 minutes.

Results and Discussion

One subject had more than 40% errors and her data were excluded from the analysis. All other subjects had less 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%). Response times 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 non-filler RT data were included in the RT analyses.

A one-way repeated measures ANOVA was performed for subject RT means and error rates (ER) and revealed significant main effects (RT: $F(3, 54) = 11.24$, $p < .001$; ER: $F(3, 54) = 13.81$, $p < .001$).

A set of planned orthogonal contrast tests revealed significant difference between conditions S and H (RT: $F(1, 18) = 11.16$, $p < .001$; ER: $F(1, 18) = 22.58$, $p < .001$), conditions H and V (RT: $F(1, 18) = 7.57$, $p < .01$; ER: $F(1, 18) = 14.25$, $p < .001$) and conditions V and D (RT: $F(1, 18) = 19.70$, $p < .001$; ER: $F(1, 18) = 14.19$, $p < .001$). The RT results are shown in Figure 3.

An additional contrast test compared the mean of conditions S and V against the mean of conditions H and D. The idea was to contrast conditions in which the cue and the target appeared on the same horizontal side of the screen against conditions in which they did not. The result of this comparison was significant (RT: $F(1, 18) = 25.32$, $p < .001$; ER: $F(1, 18) = 37.85$, $p < .001$). Another contrast test, non-orthogonal to the previous one, failed to find an RT effect of a vertical separation of the screen by comparing the mean of conditions S and V to the mean of H and D ($F(1, 18) = 2.78$, $p > .1$), but found it in error rates data ($F(1, 18) = 6.74$, $p < .05$).

The results indicate that a robust Simon-like effect was found, most clearly demonstrated by the difference between the horizontal (H) and the vertical (V) conditions. Experiment 1 failed to find a significant difference between these two conditions and the only change in Experiment 2 was the task, so the difference can only be attributed to a spatial interference which arose in condition H. In this condition the cuing words appeared in the opposite horizontal (left/right) part of the screen with respect to the target objects.

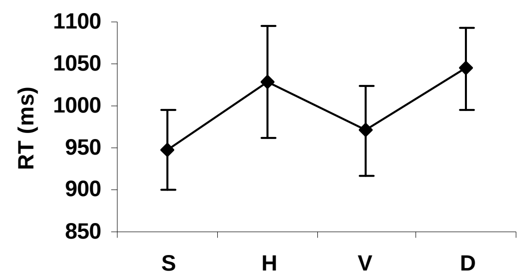

Figure 3: Experiment 2, RT results. Subjects were significantly slower when the cuing word was located on the opposite horizontal side of the screen, compared to the object it denoted.

The outcome of Experiment 2 can be accommodated within the response-discrimination account of the Simon effect (Ansorge and 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 revealed when comparing conditions S and V, as well as H and D, because in these pairs the position of the cueing word was varied only along the vertical axis, which was irrelevant to choosing the correct response action. That was exactly the pattern found in Experiment 2 (Figure 3).

The response-discrimination account makes no claims about the particular stimulus and response spatial features which are employed in the process of discrimination. Thus, it is consistent with the attention shift explanation of the Simon-like effects. Combining these accounts with visual indexing theory leads to an explanation of the results of Experiment 2. Retrieving target objects required directing attention to the location at which they had been previously presented. If in this process attention was shifted along the critical axis which discriminated responses, a spatial compatibility effect was exhibited. In this experiment only a non-correspondence effect was observed, as there was no condition in which attention was shifted in a direction compatible with the correct response action.

The significant difference in error rates between the mean of conditions S and H compared to the mean of conditions V and D can be explained with the results of Experiment 1, which showed that having to shift attention by itself worsens performance. Two additional orthogonal contrast tests were performed on RT and ER data and revealed that only the difference between conditions S and V was approaching significance (RT: $F(1, 18) = 3.75$, $p = 0.087$, $h^2 = .15$; ER: $F(1, 18) = 3.74$, $p = 0.069$, $h^2 = .17$), while comparing H and D failed to show an effect of a vertical discrimination (RT: $F(1, 18) = .52$, $p > .4$, $h^2 = .03$; ER: $F(1, 18) = 2.82$, $p > .1$, $h^2 = .13$). Thus, the effects of vertical discrimination, if any, could be attributed to facilitation in condition S, in which there was no need to shift attention in any direction.

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

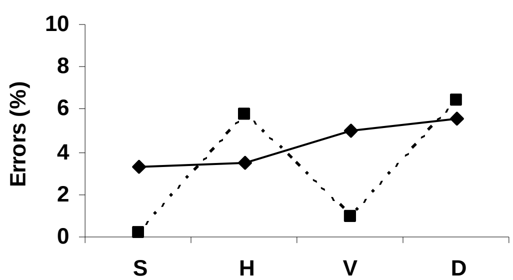

Figure 4: Error rates results. The pattern of spatial errors closely resembles the pattern of reaction times displayed in Figure 3.

A one-way repeated measures ANOVA was performed for subjects’ 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 Figure 4. No main effect or any significant difference between conditions was 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 conditions S and H ($F(1, 18) = 18.56$, $p < .001$), H and V ($F(1, 18) = 20.94$, $p < .001$), V and D ($F(1, 18) = 13.47$, $p < .01$). A set of non-parametric Friedman tests for related samples were performed on spatial error numbers as the number of spatial errors in some of the conditions was very low. A significant difference between all conditions was found ($c^2(3, 19) = 31.62$, $p < .001$), as well as between the conditions S and H ($c^2(1, 19) = 15.00$, $p < .001$), H and V ($c^2(1, 19) = 11.28$, $p < .001$), V and D ($c^2(1, 19) = 8.07$, $p < .01$).

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 performance. Participants made an equal number of errors in recalling objects that were located on either side of the response discriminating axis. This result is compatible with the eye-tracking studies by Richardson and Spivey (2000) and Spivey and Geng (2001) which also did not find an effect of the relative location of fixation on memory performance. However, the relative location of the correctly recalled objects did interfere with the spatial judgment component of the task.

General Discussion

The two experiments presented in this paper demonstrated that the relative location of presentation with respect to the current focus of attention is an important factor when recalling previously seen objects.

Experiment 1 revealed that subjects were slower when recalling an object that was presented away from the cue position, although they were not instructed to employ visual imagery. One of the reasons the experiment succeeded in showing this effect could be that the pictures of objects were displayed within a grid which allowed for better visual indexing. Also, the pictures were visually rich enough and referred to real-world objects, rather than to abstract entities, such as geometric shapes or letters. Thus, it was difficult for participants to represent the objects solely as symbolic structures by verbalizing or classifying them.

The results of Experiment 1 are in favor of the view that spatial information is automatically encoded and exploited by memory retrieval processes. It was demonstrated that irrelevant spatial information may interfere into a task which does not involve any kind of spatial discrimination – a memory recall go/no-go task. This finding reveals the ubiquity of spatial information and shows that it can intrude into cognitive processes, which, at first glance, seem to have nothing to do with space. In broader terms, the results of Experiment 1 are consistent with the externalist view of cognition (O'Regan, 1992; Spivey, Richardson & Fitneva, 2004), according to which cognitive processes such as memory retrieval can not be fully explained in terms of neural activity, but are also related to bodily states and properties of the environment.

The goal of Experiment 2 was to investigate the possible consequences of redirecting attention when retrieving objects which had been located at a distant position with respect to the current focus of attention. Several eye-tracking studies have shown that when asked to retrieve information from memory people look at the empty spaces where it has been previously presented. The results of Experiment 2 indicated that such an embodied memory retrieval process may drastically affect performance if it is running in parallel with a spatial judgment task. They imply that the reported eye-movements to empty spaces reflect attention shifts which are similar in nature to directing attention towards currently displayed objects. However the experiment failed to find evidence in favor of the view that looking at nothing facilitates memory retrieval (Ferreira, Apel & Henderson).

Although other studies have already reported Simon-like effects by memory retrieval, the novelty of the present study is in demonstrating that irrelevant spatial cues generated by attention shift to the locations of absent objects may affect performance even in the presence of directly (visually) available relevant spatial cues. This finding has important methodological implications for the studies of spatial cognition. One must always be aware that if a given stimulus is related to information previously presented at a different location it could invoke attention shifts towards this location and thus affect performance in tasks involving spatial discrimination of responses, even if the relevant spatial features are directly available. Also, Experiment 2 proposed a new experimental paradigm, which could be useful for studying the interplay of cognitive processes involving memory retrieval and spatial discrimination. Finally, the fact that memory processes may intervene in spatial tasks has to be taken into account when designing human-computer interfaces.

Acknowledgements

I would like to thank Armina Janyan for her invaluable guidance and suggestions. I would also like to thank Marina Hristova and the ANALOGY team for useful discussions and Kristina Nenova for collecting most of the data.

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(2),* 365-377.

Ferreira, F., Apel, J., & Henderson, J. (2008). Taking a new look at looking at nothing. *Trends in Cognitive Science, 11*, 405–410,

Hommel, B. (2002). Responding to object files: Automatic integration of spatial information revealed by stimulus–response compatibility effects. *The Quarterly Journal of Experimental Psychology, 55A(2),* 567-580.

Hoover, M., & Richardson (2008). When facts go down the rabbit hole: Contrasting features and objecthood as indexes to memory. *Cognition, 108,* 533-542.

Kosslyn, S., Ball, T., & Reiser, B. (1978). Visual images preserve metrical spatial information: evidence from studies of image scanning. *Journal of Experimental Psychology: Human Perception and Performance, 4*, 46-60.

Nicoletti, R., & Umilta, C. (1994). Attention shifts produce spatial stimulus codes. *Psychological Research, 56*, 144-150.

O'Regan, J. (1992). Solving the ‘real’ mysteries of visual perception: The world as an outside memory. *Canadian Journal of Psychology, 46,* 461-488.

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*. New York: Oxford University Press.

Richardson, D., & Spivey, M. (2000). Representation, space and Hollywood Squares: looking at things that aren't there anymore. *Cognition, 76(3),* 269-295.

Rubichi, S., Nicoletti, R., Iani, C., & Umilta` , C. (1997). The Simon effect occurs relative to the direction of an attention shift. *Journal of Experimental Psychology: Human Perception and Performance, 23*, 1353–1364.

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., Geng, J. (2001). Oculomotor mechanisms activated by imagery and memory: eye movements to absent objects. *Psychological Research, 65*, 235-241.

Spivey, M., Richardson, D., & Fitneva, S. (2004). Thinking outside the brain: Spatial indices to linguistic and visual information. In J. Henderson and F. Ferreira (Eds.), *The Interface of Vision Language and Action*. New York: Psychology Press.

Szekely, A., Jacobsen, T., D’Amico, S., Devescovi, A., Andonova, E., Herron, D., Lu, C-C., Pechmann, T., Pléh, C., Wicha, N., Federmeier, K., Gerdjikova, G., Gutierrez, G., Hung, D., Hsu, J., Iyer, G., Kohnert, K., Mehotcheva, T., Orozco-Figueroa, A., Tzeng, A., Tzeng, O., Arévalo, A., Vargha, A., Butler, A.C., Buffington, R., & 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(5),* 984-988.

Zhang, J., & Johnson, M. (2004). A memory-based, Simon-like, spatial congruence effect: Evidence for persisting spatial codes. *The Quarterly Journal of Experimental Psychology, 57A(3)*, 419–436.