Mental Imagery – Eyes Open and Shut

Anonymous CogSci submission

Abstract

Studies of mental imagery often ask participants to attend to a visual scene at the same time as their mental imagery. Despite the common intuition that imagery and perception interfere (known as the Perky effect), results in such experiments are not typically distinguished from those found when participants engage with mental imagery with their eyes closed. Nevertheless, studies which demonstrate the analog nature of mental images by recording the time taken for participants to scan across images have consistently found quicker scanning speeds when participants have eyes open paying attention to a visual scene as compared to with eyes closed. We show here that these results are due to the external scanning of attention across a visual scene and argue for a reevaluation of the results of such paradigms.

Keywords: mental imagery; attention; perky; projection.

Introduction

Mental imagery has enjoyed a sophisticated and rigorous investigation since the early 70s, mostly aimed at demonstrating analog, pictorial properties that can't be explained with propositional, descriptive representations. Many of the studies involve subjects closing their eyes and isolating the visualization experience from visual perceptual input (e.g., Kosslyn, 1973), but, equally, many require subjects to attend to information on a screen while employing visual mental imagery to solve a task (e.g., Finke and Pinker, 1982). Despite the fact that attending to imagery at the same time as attending to visual input is thought to inhibit one or both of these processes (Reeves & Lemley, 2012), to our knowledge no study investigating the properties of mental images in 'eyes open' scanning tasks have distinguished their findings from those of the 'eyes closed' version of this task.

In asking participants to superimpose a mental image onto an external scene, such as in Podgorny and Shepard (1978), the resulting image will have a size and position appropriate to the structure of the scene provided. For example, the 'E' in Figure 1 will be exactly 3 cells wide and 5 cells high. By contrast, when an image of the letter 'E' is not projected onto a grid on a screen, but rather held in mind with eyes closed, it has no intrinsic size or position. Even when imagining an 'E' positioned on within a 5 by 5 grid in this way with eyes closed, the 'E' incorporated into the grid has no obvious absolute or relative dimensions and position relative to a whole visual field because both the letter and the grid lack dimensional anchors that come from external scenes.

The studies outlined here seek to investigate whether mental imagery visualized in mind with eyes closed or 'projected' onto an external scaffold produce reliably different results, and in so doing demonstrate that these paradigms require further investigation before inferences can be made as to what their results contribute to the mental imagery debate.

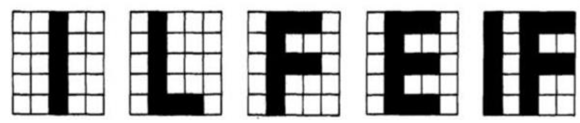

Figure 1: In Podgorny and Shepard, 1978, participants memorized letters placed in grids, and subsequently visualized the letters in empty grids which remained on screen

The distinction between eyes closed and eyes open is most apparent in mental imagery "scanning" studies (Finke & Pinker, 1982; Kosslyn, Ball, & Reiser, 1978). These types of studies seek to demonstrate that mental images have analog properties by showing that the time taken for a subject to 'scan' across the image held in mind, i.e. to move their internal attention from one item on a memorized image such as Figure 2 to another item within the same image, strongly correlated with the distance between these items in the original image.

Figure 2: An example of a memorized image used for scanning in Kosslyn, Ball, & Reiser, 1978.

There were several criticisms of this study, predominantly that performance is very sensitive to demand characteristics as researchers tell subjects to scan across the mental image and describe the scanning process in detail. Pylyshyn (1981) argued that participants were, whether consciously or not, attempting to replicate the way they would behave if they scanned the image in reality. In response to this and other criticisms, Finke and Pinker set up their 1982 study to avoid having to mention mental imagery or scanning at all. Subjects memorized a pattern of dots on the screen (Figure 3A). The dots disappeared and after a couple of seconds, an arrow appeared on the screen (Figure 3B). Their task was to respond whether the arrow was pointing at the position of any of the previously visible dots. The authors suggested that the high accuracy in this task was evidence that the subjects were visualizing the dots on the screen (Figure 3C) and scanning between the arrow and the imaginary dot. They compared the time taken to answer with the distance between the arrow and the dot, and found a strong positive correlation, similar to Kosslyn et al.'s 1978 study. This was taken as further evidence of the inherent spatiality of mental images, making the analog position even stronger due to removing the possible confounds of the 1978 study. Note that despite the fact that participants were attending to the screen the entire time, the results were taken as evidence for mental imagery, the same kind of mental imagery as evoked in the original study when subjects had their eyes closed.

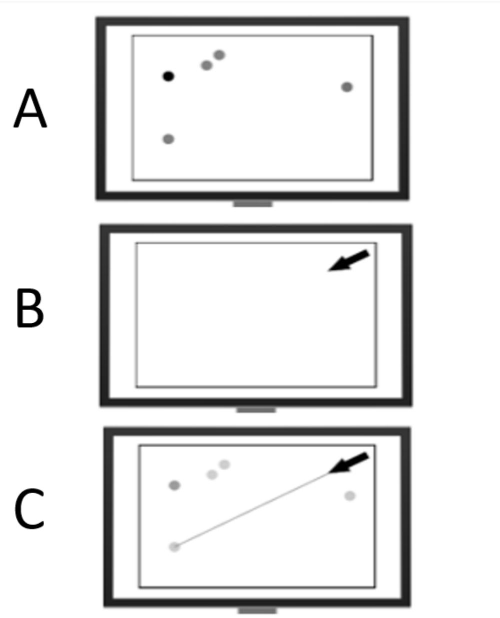

Figure 3 A-C: Finke & Pinker's arrow task (1982)

These paradigms were widely used in the intervening years (see Denis & Kosslyn, 1999) and while consistently replicating the strong correlation between distance in the image and time taken to scan, theorists questioned whether the experiments were tapping the same scanning process each time. In particular, the Kosslyn et al., 1978, scanning task (eyes closed) was thought to involve a 'transformational' process, whereby the mental image was restructured within a "visual buffer" to maintain the current area of interest in the central, high-resolution portion. The Finke and Pinker, 1982, task (eyes open) was suggested to involve shifting an internal locus of attention across the image.

In order to determine whether these paradigms were tapping two distinct types of scanning mechanisms, a side by side comparison study was ran (Borst, Kosslyn, & Denis, 2006) in which the details of the experiments were matched as much as possible. In both tasks, subjects were required to memorize an image with 5 colored dots as seen in Figure 4. In the eyes closed task, titled the KBR task after the original 1978 study, subjects mentally scanned between pairs of dots, instructed to imagine a 'small spot' moving from one dot to the other, pressing a button once they reached the second dot.

In the eyes open task, titled the FP task after the 1982 study, arrows appeared on the screen after the original image was memorized and disappeared, and subjects were asked to trace a straight line from the arrow, reporting whether the arrow pointed to a dot in the memorized image. The time taken to respond 'yes' or 'no' was taken as an indication of the time taken to scan from the arrow to the imagined dot. Both studies were matched for scanning distances between arrow and dots or pairs of dots. The speed of scanning was found to be significantly different in the FP task, and no correlations were found between performances of individual participants in each task, suggesting that different processes underly performance in each task. The authors hypothesized that the KBR task involved the transformational scanning previously mentioned, but the FP task recruited attentional scanning. Two more experiments were conducted to determine how distinct these processes were, whether specific aspects of the tasks were causing the results differential and whether they correlated with other types of tasks which may be considered to tap into similar processes. The authors found that no specific aspect of either task was responsible for the lack of within subject correlations, but did find that performance in the FP task correlated with a visual search task designed to act as a reference point for 'attentional' scanning.

The latter finding is illuminating – the visual search task involves scanning attention externally across a physical scene. This external attention is the aspect of the task that is most in common with the FP task, and yet this was not directly investigated as a potential source of the different results with the KBR task. In the FP task, participants initially attended to an arrow physically presented on screen and were then instructed to determine whether the arrow pointed to a dot that was previously presented on screen. Having memorized the dot positions, participants completed the task by tracing a line from the arrow to one of these memorized positions. This process takes place within attention to visual input whilst concurrently using imagery resources to augment this input with information previously memorized as being present within the current. That is, when memorizing and recalling the positions of the dots on the screen, the screen being attended to remains attended to *throughout the process*. In performing the task, it is uncontroversial to say that participants remain consciously attentive to visual input – they are instructed to look at the arrow, and trace a line in its direction across the screen. That they employ 'mental imagery', indistinct from that produced in eyes closed visualizing, to complete the task is the inference of decades of use of this paradigm.

Borst, Kosslyn and Denis, 2006, suggest that participants are "mentally projecting" the line of the arrow onto the screen. Kirsh (2009) investigates such projection as a collaboration of visual input and mental imagery to create shared, creative representations that are neither pure perception nor pure imagery, but a combination of both. An external scene is "augmented" by imagined information projected onto it.

In the current study we argue that 'eyes open' mental imagery scanning paradigms do not test mental imagery alone but rather representations which are formed under the integration of mental imagery and visual input. We hypothesize that any difference in the results of such studies compared with eyes closed tasks, such as that found in the Borst et al., 2006, study, will be primarily due to the integrated nature of the imagery representation, and the fact that participants are attending to visual input while employing mental imagery to complete the task.

We repeated the Borst et al. 2006 comparison study but reduced all task differences down to just one – eyes open looking at the screen versus eyes closed and picturing the image in mind. We used the 2006 KBR task in both conditions: the first condition, entitled here the 'imagination' condition, is a direct replication of the KBR task with eyes closed. In the second condition, 'projection', participants projected their mental image onto the computer screen and scanned between items within the image accordingly. If, as we suggest, the presence of visual input was driving the results differential previously found, then when visual input is the only difference between the tasks, the differential in scanning speeds should persist.

Methods

Participants

30 UC San Diego undergraduates were recruited through the SONA system and participated for course credit. 7 were excluded from data analysis after failing to report having followed the instructions at least 75% of the time in both conditions. Of the remaining 23, the average age was 21 years. 20 were native English speakers, and 3 spoke English as a second language. 19 participants were right-handed, 4 left-handed. All 23 participants reported that they had no color-blindness.

Materials

The experiment was carried out on a computer, with instructions provided on the screen and reinforced by the experimenter in person. Participants were told which colored dots to focus on and scan to by audio instructions which they heard through headphones.

Procedure

At the beginning of the experiment participants received all instructions onscreen, completed practice tasks, and then paused so that the experimenter could test their knowledge of the instructions. They then received one of two possible patterns of 5 colored dots to memorize (see Figure 4). The size and distances within the image were recreated exactly as in Borst, Kosslyn and Denis, 2006. In order to pass through the memorization phase, participants had to recreate the pattern by dragging colored dots onto the correct positions on the screen. They clicked on a button onscreen to check whether they were correct, and if each dot was overlapping with the correct position, they were allowed to move on. If not, the pattern reappeared on screen for further memorization, and when they were ready, they attempted the test again. No participant took more than 5 attempts to pass the test, and the lowest number of attempts recorded was 1.

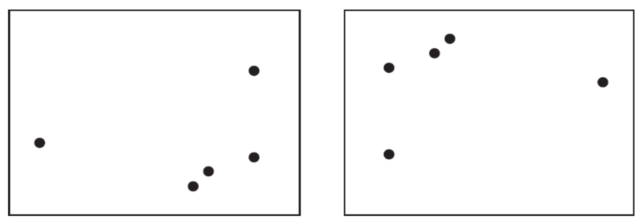

Figure 4: Pattern stimuli used in the scanning tasks. Each dot in the pattern was a different color.

The experiment had two conditions, 'imagination'; in which during test trials participants were asked to keep their eyes closed and perform the visualizations in mind, and 'projection'; in which they were asked to act as though the memorized pattern were still on the screen. Participants heard color names in pairs, 5 seconds apart, and a new color pair started 6.5 seconds after the second color was called out of the previous pair. Upon hearing the first name, participants were asked to focus on that color; either in their 'mind's eye', or on the location which that color occupied previously on the screen. They were asked to then wait until they heard the second color, at which point they scanned from the first dot to the second. Specifically, they were encouraged to imagine a small spot moving in a straight line as fast as possible (while still remaining visible) between the dots. Once they reached the second dot, they responded with a particular keypress. If the second dot was not in the pattern, they responded with a different keypress.

Of the 80 trials in each condition, half involved a color pair in which the second color was not in the pattern, and half had both colors in the pattern. All pairs began with one of the 5 colors in the pattern. No more than two pairs in a row had the second color in the pattern, but otherwise the order of pairs was randomized. Each pattern had 10 distances between dots, and each distance was encountered 8 times across the 80 trials.

Results

The results were first checked for accuracy of response indicating whether the second color of the pair was in the memorized pattern. Errors occurred in 0.4% of trials in the imagination condition and 0.6% of trials in the projection condition, consistent with findings on the KBR task in the Borst et al. 2006 condition.

Response times for all correct 'yes' responses, indicating that the second color was in the memorized pattern, were compared with the distance between that particular color pair. Following the procedure outline in the Borst et al. 2006 comparison study, RTs which were more than 2 standard deviations away from the mean of that condition for that participant were replaced with his or her mean RT for that distance. Outliers occurred in 2.2% of the trials overall.

Conducting an ANOVA on the average RT for each distance and participant and the distances in the image, we found a main effect of distance on scanning time for both the imagination ($F(9,220) = 4.22$, $p < .0001$) and projection ($F(9,220) = 1.98$, $p < 0.05$) conditions. Mean response times were 2610 ms and 2350 ms respectively. A linear regression on both tasks determined that response times increased linearly with increasing distances in both, $F(1,8) = 31.47$, $p < 0.001$ and $F(1,8) = 33.27$, $p < 0.001$ for imagination and projection respectively.

An ANOVA over data from both conditions, showed an overall effect of condition on response time, $F(1,16) = 14.3$, $p < .01$. However, unlike in the Borst et al. 2006 study, no interaction was found between distance and condition ($F<1$).

As in previous iterations of mental imagery scanning paradigms, strong correlations were found between times and distance. The time-distance correlation found for imagination was $r(8) = .893$, $p < .01$, and for the projection condition $r(8) = .898$, $p < .001$. These results indicate a successful replication of the 2006 KBR task, in both conditions.

As in the Borst et al. 2006 study, the data of primary interest were the mean slope of the best-fitting line across participants. The difference between these slopes in the original comparison, M = 82 msec/cm for the KBR task significantly different to the slope of M = 52 msec/cm for the FP task, provided the impetus for the inference of separate scanning mechanisms. The current study hypothesized that this differential would remain when task differences were removed except for external attention. The imagination condition in our experiment recorded a mean slope of M = 72.5 msec/cm, and the projection condition a mean slope of M = 40.7 msec/cm. These were significantly different from each other, $F(1,15) = 2.17$, $p < 0.05$ (see Figure 5 and 6 for a comparison of slopes across all tasks; Borst et al. 2006 and present). Where the intercepts in the original comparison were significantly different, reflecting the different time it takes to initiate responses between the two tasks, intercepts in the current comparison were almost identical, (2080 ms for imagination, 2050 ms for projection). This is consistent with the fact that the response requirements were the same for both tasks.

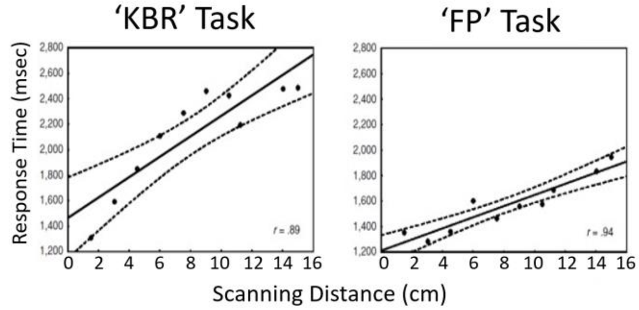

Figure 5: Comparison of linear relationships between scan time and distance between both original scanning tasks of Borst, Kosslyn and Denis (2006)

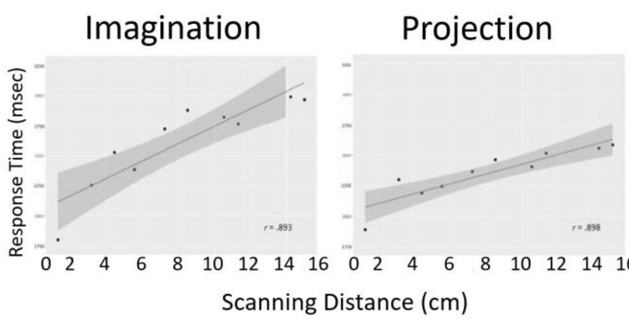

Figure 6: Comparison of linear relationships between scan time and distance between both tasks of the current study

In order to determine whether the fixed dimensions of the screen reduced variability in scanning times across participants, an ANOVA was conducted between standard deviations, distance and condition. A main effect of distance was found on standard deviation, $F(1,16) = 65.56$, $p < .0001$, as well as condition, $F(1,16) = 14.3$, $p < .01$. An interaction between condition and distance was also found, $F(1,16) = 15.94$, $p < .01$.

Discussion

When performing a mental imagery task whilst concurrently paying attention to external structure such as a screen, our results show that the characteristics of resultant behavior, normally ascribed entirely to the nature of the mental imagery held in mind, are influenced by the properties of the screen being attended to. Specifically, we show here that participants 'scan' their attention across a mental image that is projected onto a computer screen significantly quicker than they do a mental image that is held in mind with eyes closed. Unsurprisingly, there was also significantly less variance in response times across participants in the 'projection' task. In a nutshell; when subjects scan their attention externally across a physical screen, they do it much quicker than when they scan their attention across internal information; and the presence of the screen scaffolded and constrained their mental images such that the image properties were more consistent across participants. In the projection condition, we suggest, participants were working with a mental image that was integrated with the computer screen – the visual input of the white screen and black rectangular border formed part of the mental image representation, along with the dots which were no longer present on the screen and the locations of which were nonetheless remembered with high precision (as tested at the beginning of each block).

The differential in results produced here mirror that of previous studies comparing different scanning paradigms (Borst, Kosslyn, & Denis, 2006), however, in searching for the source of the differential, the presence of the computer screen in 'eyes open' scanning tasks was not considered. Why is this? It is not new to say that mental imagery can integrate with perception. Imagery and perception cause similar activations in areas of the visual cortex (Kosslyn & Thompson, 2003). Eye movements appear to be implicated in the creation and manipulation of mental imagery (Mast & Kosslyn, 2002; Johansson et al., 2005). The Perky effect (Perky, 1910), once describing the subjective similarly between perception and imagery is now taken to mean the tendency of imagery to interfere with perception (Reeves & Lemley, 2012). This interference can be constructive, as when a mental image matches a stimulus being detected (Farah, 1985), or destructive, as when projected imagery reduces visual acuity (Craver-Lemley & Reeves, 1987). These cases are all taken to demonstrate the representational similarity of perception and mental imagery, however, they are not explicitly concerned with whether representations can be formed through the integration of both perception and imagery, and still treat mental imagery as separate from visual input - even in cases when both are attended to concurrently.

Lewis, Borst and Kosslyn (2011) sought to investigate such a representational integration. Subjects attended to and memorized a pattern of dots on a grid, and subsequently were requested to visualize this pattern in an empty grid on a screen. A second pattern of dots was briefly presented for 33ms in the grid, after which participants indicated within the again-empty grid which cell of the grid remained empty after combining both the previously memorized and the just-seen pattern of dots. The high accuracy of subjects on this task provided evidence that mental images can be "integrated with percepts to create a single composite representation". However, although this representational integration is precisely what we are arguing for here, the authors missed a crucial aspect of this integration which forms the basis for the current study, and is the same aspect that, in missing, allowed the results differential between eyes open and closed scanning paradigms to elude explanation for so many years.

Here is a key line from Lewis et al. 2011 which illustrates the point being made here: "During this interval, we discouraged participants from visualizing the dots by presenting a blank gray background; this background did not contain the grid, and thus made it more difficult to visualize the dots". Why was it difficult to visualize the dots (in mind) without the grid (in the world) being present? If the mental image of the dots needs the external grid for its formation and maintenance, then is the mental image really a totally internal mental image, separate from visual input? What we are arguing for here is that, when visualizing a pattern of dots while looking at a grid on a screen, the resultant mental image is *itself* a composite of imagery and percept. The authors see this integration has happening between two distinct representations, one of which is canonical mental imagery, the same whether it involves an external grid or not. We see the integration as happening in the formation of and maintenance of the mental image, whenever visualizing is accompanied by externally-directed attention to information designed to support visualizing. If one attends to visual input as part of the process of visualizing mental imagery, then that visual input is part of the imagery representation. The properties uncovered during this process; for example - a correlation between distance scanned and time taken to scan across the mental image projected onto a screen; should not be taken to describe 'mental imagery' per se, purely internal information, as is the case when visualizing with eyes closed. The properties uncovered in eyes open, projected-imagery tasks, tell us about the combination of imagery and perception; a combination that, as previous studies have argued, can provide compelling support for the depictive view of mental imagery. The novelty of the current finding and our argument is that, where previous studies demonstrate the integration of static imagery and perceptual information, in mind and after the fact, the integration demonstrated and argued for here is an *online* integration of imagery and perception. Imagery, of say a pattern of dots, and perception, of a grid, combine under attention to form one unified representation in real time. In this case, there is no separate, purely mental image; the image is formed using visual input as a basis.

It should be noted that Pylyshyn, (1997), discusses the superimposition of images onto visual perception in his analysis of mental imagery scanning paradigms in a manner that comes closest to an 'online' account of imagery and perception. He argues that in the eyes open scanning task subjects are scanning their attention and eyes "from place to place on the display", using indexes in the "real scene" to navigate. This is intended to challenge the notion that this paradigm gives support for the analog point of view – if subjects are scanning in the 'real world', then the constraints which give rise to the distance time correlation come from the world and not from the inherent nature of the format of mental imagery.

While it is true that, under this approach, the distance-time correlation in the projected condition is likely to be a function of distance in the real world as opposed to properties of a mental image alone, the question of how imagined information can be superimposed so successfully onto visual input, in the current studies as well as the many paradigms which utilize this type of paradigm (e.g., Ongchoco & Scholl, 2019), is in itself interesting. In order to proceed into the scanning trials of the study presented here, subjects had to recreate the 5 dots on the screen so that each dot was within half a radius of the original position. Participants succeed in this memory test surprisingly quickly; within 2-4 attempts for the majority. In the "low discriminability" arrow task of the Borst et al. 2006 study, in which the arrow pointed very close to but not at a dot, participants performed with 86.5% accuracy across 80 trials. This ability to so successfully and accurately map a mental image onto the screen might well be evidence that mental images share a format with visual perception. Further investigation onto the online integration of imagery and visual input could provide fruitful insights into the nature of mental imagery, and into its top-down role in perception.

Conclusion

The aim of much of mental imagery research has been to determine the analog characteristics of imagery representations. This research commonly uses paradigms in which attention is putatively paid to mental imagery while the subject is concurrently attending to visual input. In these cases, the resulting behavior has been taken as evidence of the properties of the imagery representations alone. In this paper, we argue that the results of paradigms where participants visualize mental imagery while also attending to an external scaffold should not be taken as providing evidence as to the characteristics of pure mental imagery. However, these paradigms could, if investigated as an online integration of resources, provide new theoretical and empirical opportunities to advance our knowledge of the nature of mental imagery representations and their role in perception.

References

Borst, G., Kosslyn, S. M., & Denis, M. (2006). Different cognitive processes in two image-scanning paradigms. Memory & cognition, 34(3), 475-490.

Craver-Lemley, C., & Reeves, A. (1987). Visual imagery selectively reduces vernier acuity. *Perception*, *16*(5), 599-614.

Denis, M., & Kosslyn, S. M. (1999). Scanning visual mental images: A window on the mind. Cahiers de Psychologie Cognitive/Current Psychology of Cognition.

Farah, M. J. (1985). Psychophysical evidence for a shared representational medium for mental images and percepts. *Journal of Experimental Psychology: General*, *114*(1), 91.

Finke, R. A., & Pinker, S. (1982). Spontaneous imagery scanning in mental extrapolation. Journal of Experimental Psychology: Learning, Memory, and Cognition, 8(2), 142.

Johansson, R., Holsanova, J., & Holmqvist, K. (2005). What do eye movements reveal about mental imagery? Evidence from visual and verbal elicitations. In *Proceedings of the 27th Cognitive Science conference* (Vol. 1054). Erlbaum Mahwah, NJ.

Kirsh, D. (2009). Projection, problem space and anchoring. In Proceedings of the Annual Meeting of the Cognitive Science Society (Vol. 31, No. 31).

Kosslyn, S. M. (1973). Scanning visual images: Some structural implications. Perception & Psychophysics, 14(1), 90-94.

Kosslyn, S. M., Ball, T. M., & Reiser, B. J. (1978). Visual images preserve metric spatial information: evidence from studies of image scanning. Journal of experimental psychology: Human perception and performance, 4(1), 47.

Kosslyn, S. M., & Thompson, W. L. (2003). When is early visual cortex activated during visual mental imagery?. *Psychological bulletin*, *129*(5), 723.

Lewis, K. J., Borst, G., & Kosslyn, S. M. (2011). Integrating visual mental images and visual percepts: new evidence for depictive representations. *Psychological research*, *75*(4), 259-271.

Mast, F. W., & Kosslyn, S. M. (2002). Eye movements during visual mental imagery. *Trends in Cognitive Sciences*, *6*(7), 271-272.

Ongchoco, J. D. K., & Scholl, B. J. (2019). How to create objects with your mind: From object-based attention to attention-based objects. Psychological Science, 30(11), 1648-1655.

Perky, C. W. (1910). An experimental study of imagination. The American Journal of Psychology, 21(3), 422-452.

Podgorny, P., & Shepard, R. N. (1978). Functional representations common to visual perception and imagination. Journal of Experimental Psychology: Human Perception and Performance, 4(1), 21.

Pylyshyn, Z. W. (1981). The imagery debate: Analogue media versus tacit knowledge. Psychological review, 88(1), 16.

Pylyshyn, Z. (1997). The role of visual indexes in spatial vision and imagery. Visual Attention.

Reeves, A., & Lemley, C. (2012). Unmasking the Perky effect: Spatial extent of image interference on visual acuity. Frontiers in psychology, 3, 296.