Quantifying Aphantasia through drawing: Those without visual imagery show deficits in object but not spatial memory

Wilma A. Bainbridge[1], Zoë Pounder[2], Alison F. Eardley[2], Chris I. Baker[1]

1 – Laboratory of Brain and Cognition; National Institute of Mental Health, Bethesda, MD USA
2 – School of Social Sciences; University of Westminster, London, UK

Abstract

Congenital aphantasia is a recently identified experience defined by the inability to form voluntary visual imagery, with intact semantic memory and vision. Although understanding aphantasia promises insights into the nature of visual imagery, as a new focus of study, research is limited and has largely focused on small samples and subjective report. The current large-scale online study of aphantasics (N=63) and controls required participants to draw real-world scenes from memory, and copy them during a matched perceptual condition. Drawings were objectively quantified by 2,700 online scorers for object and spatial details. Aphantasics recalled significantly fewer object details than controls, and showed a reliance on verbal strategies. However, aphantasics showed equally high spatial accuracy as controls, and made significantly fewer memory errors, with no differences between groups in the perceptual condition. This object-specific memory impairment in aphantasics provides evidence for separate systems in memory that support object versus spatial details.

Introduction

Visual imagery, the ability to form visual mental representations, is a common human cognitive experience, yet it is has been hard to characterize and quantify. What is the nature of the images that come to mind when forming visual representations of objects or scenes? What might these representations look like if one lacks this ability? *Aphantasia* is a recently identified experience, defined by an inability to create voluntary visual mental images, although semantic memory and vision remain intact [1,2]. Aphantasia is still largely uncharacterized, with many of its studies based on case studies or employing small sample sizes. Here, using an online crowd-sourced drawing task designed to quantify the content of visual memories [3], we examine the nature of aphantasics' mental representations of visual images within a large sample, and reveal evidence for separate object and spatial systems in human imagery.

Although some cases reporting an absence of mental imagery were first identified in the 19th century [4], the term *aphantasia* has only recently been defined and investigated, within fewer than a dozen studies [1,2,4-8]. This is arguably because most individuals with aphantasia can lead functional, professional lives, with many individuals realizing their imagery experience differed from the majority only in adulthood. The current method for identifying if an individual has aphantasia is through subjective self-report, using the Vividness of Visual Imagery Questionnaire [9]. However, recent research has begun quantifying the experience using objective measures such as priming during binocular rivalry [2] and skin conductance during reading [7]. Since its identification, several prominent figures have come forth describing their experience with aphantasia, including economics professor Nicholas Watkins [10], Firefox co-creator Blake Ross [11], and former Pixar Chief Technology Officer Ed Catmull [12], leading to broader recognition of the experience.

Like congenital prosopagnosia [13], in the absence of any brain damage or trauma, aphantasia is considered to be congenital (although it can also be acquired through trauma [14]). However, beyond this, little research has examined the nature of aphantasia and the impact on imagery function and cognition more broadly. A single-participant aphantasia case study found no significant difference from controls in a visual imagery task (judging the location of a target in relation to an imagined shape) nor its matched version of a working memory task, except at the hardest level of difficulty [5]. However, aphantasics show significantly less imagery-based priming in a binocular rivalry task [2, 15], and show diminished physiological responses to fearful text as compared with controls [7]. While these studies have observed differences between aphantasics and controls, the nature of aphantasics' mental representations during visual recall is still unknown. Understanding these differences in representation between aphantasics and controls could shed light on broader questions of what information (visual, semantic, spatial) makes up a memory, and how this information compares to the initial perceptual trace. In fact, the existence of aphantasia serves as key evidence against the hypothesis that visual perception and imagery rely upon the same neural substrates and representations [16], and also suggests a dissociation of visual recognition and recall (as aphantasia only affects the latter). Examination into aphantasia thus has wide-reaching potential implications for the understanding of the way we form mental representations of our world.

The nature and content of our visual imagery has proved incredibly difficult to quantify. Several studies in psychology have developed tasks to objectively study the cognitive process of mental imagery through visual working memory or priming (e.g., [9,17,18]). One of the long-standing debates within the imagery literature has been over the nature of images, and specifically whether visual imagery representations are depictive and picture-like in nature [19,20] or symbolic, "propositional" representations [21,22]. Neuropsychological research, especially in neuroimaging, has led to large leaps in our understanding of visual imagery. Studies examining the role and activation of the primary visual cortex during imagery tasks have been interpreted as supporting the depictive nature of imagery [23-26]. However, neuropsychological studies have identified patients with dissociable impairments in perception versus imagery [27,28], and recent neuroimaging work has suggested there may be systematically related yet separate cortical areas for perception and imagery, and that the neural representation during recall may lack much of the richer, elaborative processing of the initial perceptual trace [29-31]. Combined with research identifying situations where propositional encoding dominates spatial imagery (e.g., [32]), researchers have concluded that there is a role for both propositional and depictive elements in the imagery process (e.g., [33]). In their case study, Jacobs and colleagues [5] argue that differences in performance between aphantasic participant *AI* and neurotypical controls may result from different strategies, including a heavier reliance on propositional encoding, relying on a spatial or verbal code. Thus, ideally a task that measures both depictive (visual) and propositional (semantic) elements of a mental representation could directly compare the strategies used by aphantasics and controls.

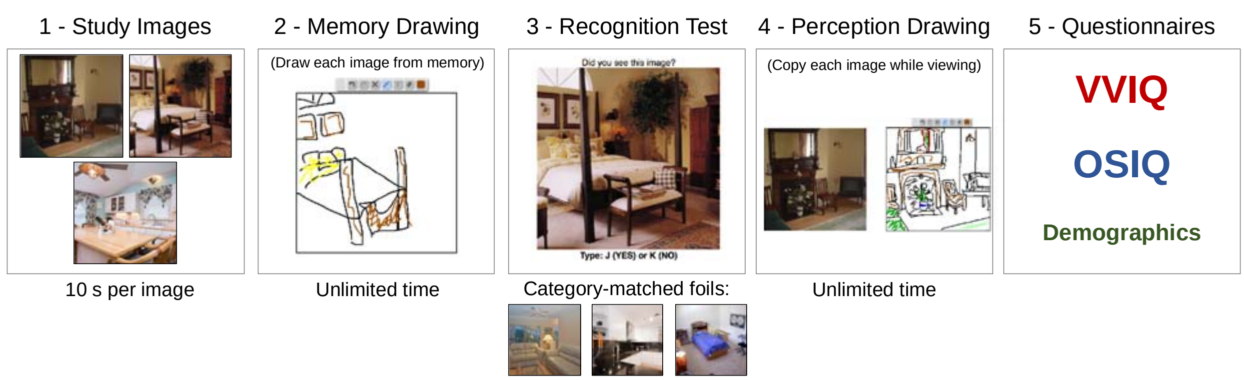
In a recent study, impressive levels of both object and spatial detail could be quantified by drawings made by neurotypical adults in a drawing-based visual memory experiment [3]. Such drawings allow a more direct look at the information within one's mental representation of a visual image, in contrast to verbal descriptions or recognition-based tasks. A drawing task may allow us to identify what fundamental differences exist between aphantasics and individuals with typical imagery, and in turn inform us of what information exists within imagery.

In the current study, we examine the visual memory representations of congenital aphantasics and individuals with typical imagery (controls) for real-world scene images. Through online crowd-sourcing, we leverage the power of the internet to identify and recruit large numbers of both aphantasic and controls for a memory drawing task. We also recruit over 2,700 online scorers to objectively quantify these drawings for object details, spatial details, and errors in the drawings. We discover a selective impairment in aphantasics for object memory, with significantly fewer visual details and evidence for increased semantic scaffolding. In contrast, for the items that they remember, aphantasics show spatial accuracy at the same high level of precision as controls. Aphantasics also show fewer memory errors and memory correction as compared to controls. These results may point to two systems that support object information versus spatial information in memory.

Results

Aphantasic (N=63) and control (N=52) participants were recruited online through social media (Facebook, Twitter, and aphantasia-specific Facebook and Reddit communities) to participate in an online memory drawing experiment (Fig. 1a). The experiment comprised of five parts. First, participants studied three real-world scene images for 10s each (Fig. 1a), all pre-selected to give maximal information in a prior memory drawing study [3]. Second, participants were instructed to draw each of the three images from memory using a basic drawing canvas web interface that included a pencil tool, different colors, and the ability to erase and undo/redo. Participants did not know they would be tested through drawing until after studying the images, to prevent drawing-targeted study strategies. Participants were given unlimited time to draw, and could draw in any order. Mouse movements were tracked during drawing in order to measure drawing time and erasing behavior. Third, participants completed a recognition task in which they indicated if they had previously seen each of six images: the three images presented for drawing as well as three category-matched foils in the experiment. Fourth, they were instructed to draw a copy of each of the first three images, while viewing them. This phase again had unlimited time, and the images were presented in a random order. Finally, participants completed the VVIQ [9] and Object-Spatial Imagery Questionnaire (OSIQ [34]), and were asked for feedback with regards to the section of the experiment they found most difficult, as well as asked several demographic questions. Only aphantasics with a VVIQ score of 25 or below and controls with a VVIQ score of 40 or above were used in the analyses [1], resulting in the exclusion of eight participants with intermediate scores between 25 and 40.

a)



b)

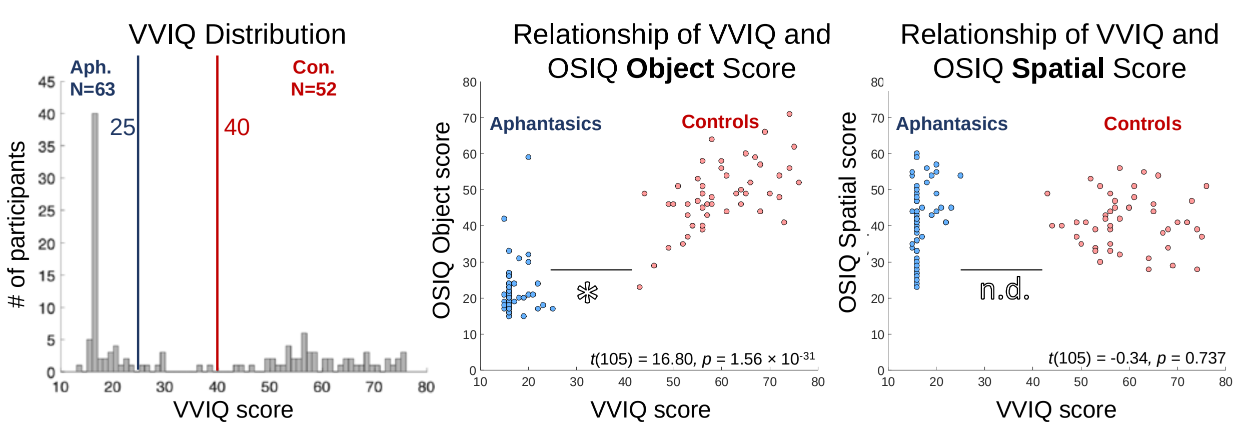


Fig. 1. Experimental paradigm and basic demographics. a) The experimental design of the online experiment. Participants studied three photographs, drew them from memory, completed a recognition task, copied the images while viewing them, and then filled out the VVIQ and OSIQ questionnaires in addition to demographics questions. You can try the experiment at http://wilmabainbridge.com/research/aphantasia/aphantasia-experiment.html. whole experiment took approximately 30 minutes. b) (Left) A histogram of the distribution of all participants across the VVIQ. Aphantasics were selected as those scoring 25 and below (N=63) and controls were selected as those scoring 40 and above (N=52), while those in between were removed from the analyses (N=8). (Middle) A scatterplot of total VVIQ score plotted against total OSIQ Object component score for participants meeting criterion. Each point represents a participant, with aphantasics in blue and controls in red. There was a significant difference in OSIQ Object score between the two groups. (Right) A scatterplot of total VVIQ score plotted against OSIQ Spatial component score. There was no difference in OSIQ Spatial score between the two groups.

No demographic differences between groups, but reported differences in object and spatial imagery

First, we analyzed whether there were demographic differences between the groups. There was a significant difference in age between groups with aphantasics generally older than controls (aphantasic: M=41.16 years, SD=14.22; control: M=32.12 years, SD=15.26). However, if we conduct all analyses with a down-sampled set of 52 aphantasic participants with a matched age distribution, no meaningful differences in the results emerge. There was no significant difference in gender proportion between the two groups (aphantasic: 63.5% female; control: 59.6% female; Pearson's chi-square test for proportions: χ^2=0.18, p=0.670), even though a previous study reported a sample comprising of predominantly males (Zeman et al., 2015). There was no significant difference between participant sets in reported artistic abilities ($t(113)$=0.71, p=0.480).

Second, we investigated the relationship of the VVIQ score and OSIQ (Fig. 1b), a questionnaire developed to separate abilities to perform imagery with individual objects versus spatial relations amongst objects [34]. Controls scored significantly higher on the OSIQ than aphantasics ($t(105) = 11.44$, $p=3.60 \times 10^{-20}$). There was a significant correlation between VVIQ score and OSIQ score for control participants (M=89.73, SD=10.97; Spearman rank-correlation test: ρ=0.54, p=7.70 × 10^{-5}), but not for aphantasics (OSIQ *M* score=63.88, *SD*=12.12; ρ=0.24, p=0.071). When broken down by OSIQ subscale, there was a significant difference between groups in questions relating to object imagery (t(105)=16.80, p=1.56 × 10^{-31}), but not spatial imagery (t(105)=-0.34, p=0.737). Indeed, a 2-way ANOVA (participant group × subscale) reveals a main effect of participant group (F(1,210)=128.87, p~0), subscale (F(1,210)=30.95, p=8.00 × 10^{-8}), and a significant interaction (F(1,210)=140.20, p~0), confirming a difference in self-reported ratings for object imagery and spatial imagery respectively. This difference in self-reported object imagery and spatial imagery has been reported in previous studies [2], and suggests a potential difference between the two imagery subsystems.

Finally, given the focus of the current experiment on visual recall, we also compared measures of visual recognition performance. Both groups performed near ceiling at visual recognition of the images they studied, with no significant difference between groups in recognition hit rate (controls: M=0.96, SD=0.12; aphantasics: M=0.98, SD=0.12; Wilcoxon rank-sum test: Z=1.16, p=0.245), or false alarm rate (controls: M=0.02, SD=0.12; aphantasics: M=0, SD=0; Wilcoxon rank-sum test: Z=1.13, p=0.260). These results indicate that there is no deficit in aphantasics for recognizing images, even with lures from the same semantic scene category.

Diminished object information for aphantasics

Next, we turned to analyzing the drawings made by the participants to reveal objective measures of the mental representations of these two groups. Looking at overall number of drawings made, while a small number of participants could not recall all three images, there was no significant difference between groups in number of images drawn from memory (control: *M*=2.92, *SD*=0.27; aphantasic: *M*=2.87, *SD*=0.38; Wilcoxon rank-sum test: Z=0.63, p=0.526). To evaluate the drawings, 2,795 unique workers from the online experimental platform Amazon Mechanical Turk (AMT) scored the drawings on a variety of metrics including object information, spatial accuracy, and memory errors, using methods previously established for quantifying memory drawings [3]. Importantly, each participant completed both a memory drawing (i.e., drawing an image from memory for an unlimited time period) and a perception drawing (i.e., copying from a drawing for an unlimited time period) for each image, allowing us to compare for each participant what is in memory versus what that individual would maximally draw given an image without memory constraints (refer to Fig. 2 for example drawings). This comparison allows us to control for differences in effort and drawing ability, which we should expect to be reflected in both types of drawings.

Fig. 2. Example drawings. Example drawings made by aphantasic and control participants from memory and perception (i.e., copying the image) showing the range of performance. Each row is a separate participant, and the memory and perception drawings connected by arrows are from the same participant. Low memory examples show participants who drew the fewest from memory but the most from perception. High memory examples show participants who drew the highest amounts of detail from both memory and perception.

These examples are all highlighted in the scatterplot of Fig. 3. The key question is whether there are meaningful differences between these two sets of participants' drawings.

To score level of object information, AMT workers (N=5 per object) identified whether each of the objects in an image was present in each drawing of that image (Fig. 3). A 2-way ANOVA of participant group (aphantasic / control) × drawing type (memory / perception drawing, repeated measure) looking at number of objects drawn per image showed no significant overall effect of participant group ($F(1,225)=0.69$, $p=0.408$), but a significant effect of drawing type ($F(1,225)=593.96$, $p\sim0$), and more importantly, a significant statistical interaction ($F(1,225)=11.08$, $p=0.0012$). Targeted post-hoc t-tests revealed that when drawing from memory, controls drew significantly more objects ($M=6.32$ objects per image, $SD=3.07$) than aphantasics ($M=5.07$, $SD=2.61$; independent samples t-test: $t(113)=2.33$, $p=0.022$) across the experiment. In contrast, when copying a drawing (perception drawing), aphantasics on average drew more objects from the images than controls, but with no significant difference (controls: $M=18.00$ objects per image, $SD=5.81$; aphantasics: $M=20.45$, $SD=6.58$; $t(113)=0.86$, $p=0.392$). These results suggest that aphantasics are showing a specific deficit in recalling object information during memory.

Given that some participants tended to draw few objects even when copying from an image, we also investigated a corrected measure, taken as the number of objects drawn from memory divided by the number of objects drawn from perception, for each image for each participant. Drawings from perception with fewer than 5 objects were not included in the analysis, to remove any low-effort trials. Aphantasics drew a significantly smaller proportion of objects from memory than control participants (aphantasic: $M=0.269$, $SD=0.173$; control: $M=0.369$, $SD=0.162$; Wilcoxon rank-sum test: $Z=3.88$, $p=1.02 \times 10^{-4}$). We also investigated the correlation within groups between the number of objects drawn from memory and the number drawn from perception. Controls show a strong correlation, where the more objects one draws from perception, the more one also tends to draw from memory (Pearson correlation: $r=0.45$, $p=7.94 \times 10^{-4}$). Aphantasics show a significant, but much weaker relationship ($r=0.27$, $p=0.038$).

We also assessed the relationship between performance in the task and self-reported object imagery in the OSIQ. Across groups, there was a significant correlation between proportion of objects drawn from memory and OSIQ object score (Spearman's rank correlation: ρ=0.31, p=9.43 × 10^{-4}), although these correlations were not significant when separated by participant group (p>0.10).

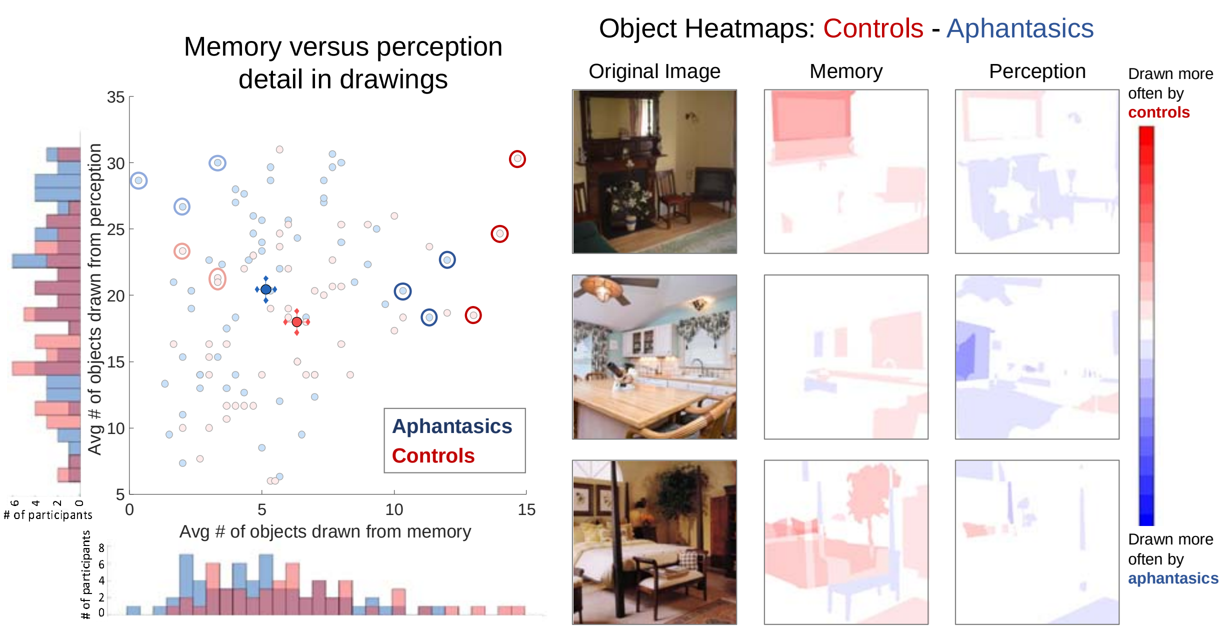


Fig. 3. Comparison of object information in drawings between aphantasics and controls. (Left) A scatterplot of each participant as a point, showing average number of objects drawn from memory across the three images (x-axis), versus average number of objects drawn from perception across the three images (y-axis). Aphantasics are in blue, while controls are in red. The bright blue circle indicates average aphantasic performance, while the bright red circle indicates average control performance, with crosshairs for both indicating standard error of the mean for memory and perception respectively. Histograms on the axes show the number of participants who drew each number of objects. Controls drew significantly more objects from memory, although with a tendency towards fewer from perception. The highlighted light blue and red points are the participants with the lowest memory performance shown in Fig. 2, while the highlighted dark blue and red points are the participants with the highest memory performance shown in Fig. 2. (Right) Heatmaps of which objects for each image tended to be drawn more by controls (red) or aphantasics (blue). Pixel value represents the proportion of control participants who drew that object in the image subtracted by the proportion of aphantasics who drew that object (with a range of -1 to 1). Controls remembered more objects (i.e., there is more red in the memory heatmaps), even though aphantasics tended to copy more objects (i.e., there is more blue in the perception heatmaps).

Next, we examined whether there was a difference in visual detail within objects, by quantifying whether participants included color in their object depictions. Significantly more memory drawings by controls contained color than those by aphantasics (control: 38.2%, aphantasics: 21.0%; Pearson's chi-square test for proportions: χ^2=11.07, p=8.78 × 10^{-4}), while there was no difference for perception drawings (control: 46.2%, aphantasic: 38.0%, χ^2=2.12, p=0.146). Control participants also spent significantly longer time on their memory drawings than aphantasics (control: M=2023.5 ms per image, SD=1383.6 ms; aphantasics: M=1002.7 ms, SD=654.7ms; $t(110)$ = 5.14, p=1.19 × 10^{-6}), possibly implying more attention to detail in their drawings. We investigated other forms of object detail, by having AMT workers (N=777) judge whether different object descriptors (e.g., material, texture, shape, aesthetics; generated by 304 separate AMT workers) applied to each drawn object. This task did not identify differences between groups for the memory drawings ($t(115)$=0.14, p=0.886), although objects were significantly more detailed when copied than when drawn from memory for both aphantasics (memory: M=42.2% descriptors per object applied, SD=5.1%; copied: M=45.7%, SD=4.0%; $t(127)$=4.31, p=3.23 × 10^{-5}) and control participants (memory: M=42.1%, SD=5.6%; copied: M=47.0%, SD=3.8%; $t(102)$=5.20, p=1.01 × 10^{-6}). However, it is possible this task may have asked for too fine-grained information than can be measured from these drawings (e.g., judging the material and texture of a drawn chair).

In sum, these results present concrete evidence that aphantasics recall fewer objects than controls, and these objects contain less visual detail (i.e., color) within their memory representations.

Aphantasics show greater dependence on symbolic representations

While aphantasics show decreased object information in their memory drawings, they are still able to successfully draw some objects from memory (5.07 objects per image on average). Do these drawings reveal evidence for alternative, non-visual strategies that may have supported this level of performance? To test this question, we quantified the amount of text used to label objects included in the participants' drawings. Note that while labeling was allowed (the instructions stated: "Please draw or label anything you are able to remember"), it was effortful as it required drawing the letters with the mouse. We found that significantly more memory drawings by aphantasics contained text than those by controls (aphantasic: 27.8%, control: 16.0%; χ^2=6.84, *p*=0.009). Further, there was no difference between groups for perception drawings (aphantasic: 2.8%, control: 0.8%; χ^2=1.66, *p*=0.197). These results imply that aphantasics may have relied upon symbolic representations to support their memory.

Comments by aphantasics at the end of the experiment supported their use of symbolic strategies. When asked what they thought was difficult about the task, one participant noted, "Because I don't have any images in my head, when I was trying to remember the photos, I have to store the pieces as words. I always have to draw from reference photos." Another aphantasic stated, "I had to remember a list of objects rather than the picture," and another said, "When I saw the images, I described them to myself and drew from that description, so I… could only hold 7-9 details in memory." In contrast, control participants largely commented on their lack of confidence in their drawing abilities: e.g., "I am very uncoordinated so making things look right was frustrating"; "I can see the picture in my mind, but I am terrible at drawing."

Aphantasics and controls show equally high spatial accuracy in memory

While aphantasics show an impairment in memory for object information, do they also show an impairment in spatial placement of the objects? To test this question, AMT workers (N=5 per object) drew an ellipse around the drawn version of each object, allowing us to quantify the size and location accuracy of each drawn object (Fig. 4). When drawing from memory, there was no significant difference between groups in object location error in the x-direction (aphantasic: *M* pixel error=63.99, *SD*=31.18; control: *M*=60.63, *SD*=28.45; *t*(113)=0.60, *p*=0.551) nor the y-direction (aphantasic: *M*=64.97, *SD*=29.90; control: *M*=69.10, *SD*=29.72; *t*(113)=0.74, *p*=0.461). However, this lack of difference was not due to difficulty in spatial accuracy; both groups' drawings were incredibly spatially accurate, with all average errors in location less than 10% of the size of the images themselves. Similarly, there was also no significant difference in drawn object size error in terms of width (aphantasic: *M* pixel error=23.06, *SD*=10.88; control: *M*=24.89, *SD*=13.58; *t*(113)=0.81, *p*=0.422) and height (aphantasic: *M*=26.80; *SD*=14.01; control: *M*=22.82; *SD*=11.05; *t*(113)=1.66, *p*=0.099), and these sizes were incredibly accurate in both groups (average errors less than 4% of the image size). There was no correlation between a participant's level of object location or size error and ratings on the OSIQ spatial questions (all *p*>0.250). In all, these results show that both aphantasics and controls have highly accurate memories for spatial location, with no observable differences between groups.

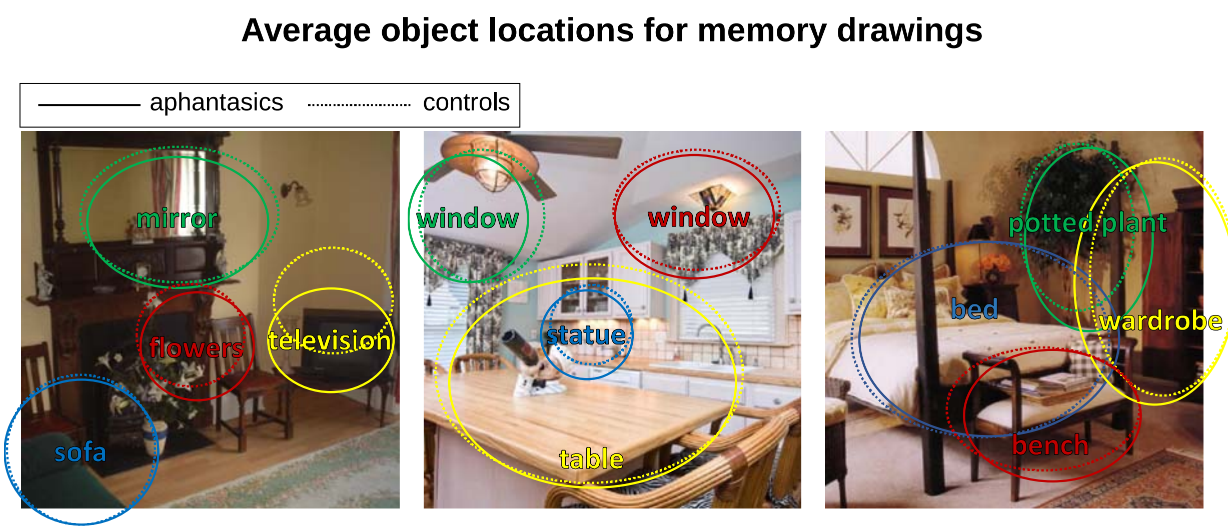


Fig. 4. Average object locations and sizes recalled by aphantasics and controls. Average object locations and sizes for memory drawings of four of the main objects from each image, made by aphantasics (solid lines) and controls (dashed lines). Even though these objects were drawn from memory, their location and size accuracy was still very high. Importantly, aphantasics and controls showed no significant differences in object location or size accuracy.

Aphantasics draw fewer false objects than controls

Finally, we quantified the amount of error in participants' drawings from memory by group. AMT workers (N=5 per drawing) viewed a drawing and its corresponding image and wrote down all objects in the drawings that were not present in the original image (essentially quantifying false object memories). Significantly more memory drawings by controls contained false objects than drawings by aphantasics (control: 12 drawings, aphantasic: 3 drawings;

Pearson chi-square test: χ^2=9.35, p=0.002); examples can be seen in Fig. 5. Similarly, significantly more objects drawn by controls were false alarms than those drawn by aphantasics (χ^2=5.09, p=0.024). This indicates that control participants were making more memory errors, even after controlling for the fewer number of objects drawn overall by aphantasics. Interestingly, all aphantasic errors (see Fig. 5) were transpositions from another image and drawn in the correct location as the original object (a tree from the bedroom to the living room, a window from the kitchen to the living room, and a ceiling fan from the kitchen to the bedroom). In contrast, several false memories from controls were objects that did not exist across any image but instead appeared to be filled in based on the scene category (e.g., a piano in the living room, a dresser in the bedroom, logs in the living room). No perception drawings by participants from either group contained false objects.

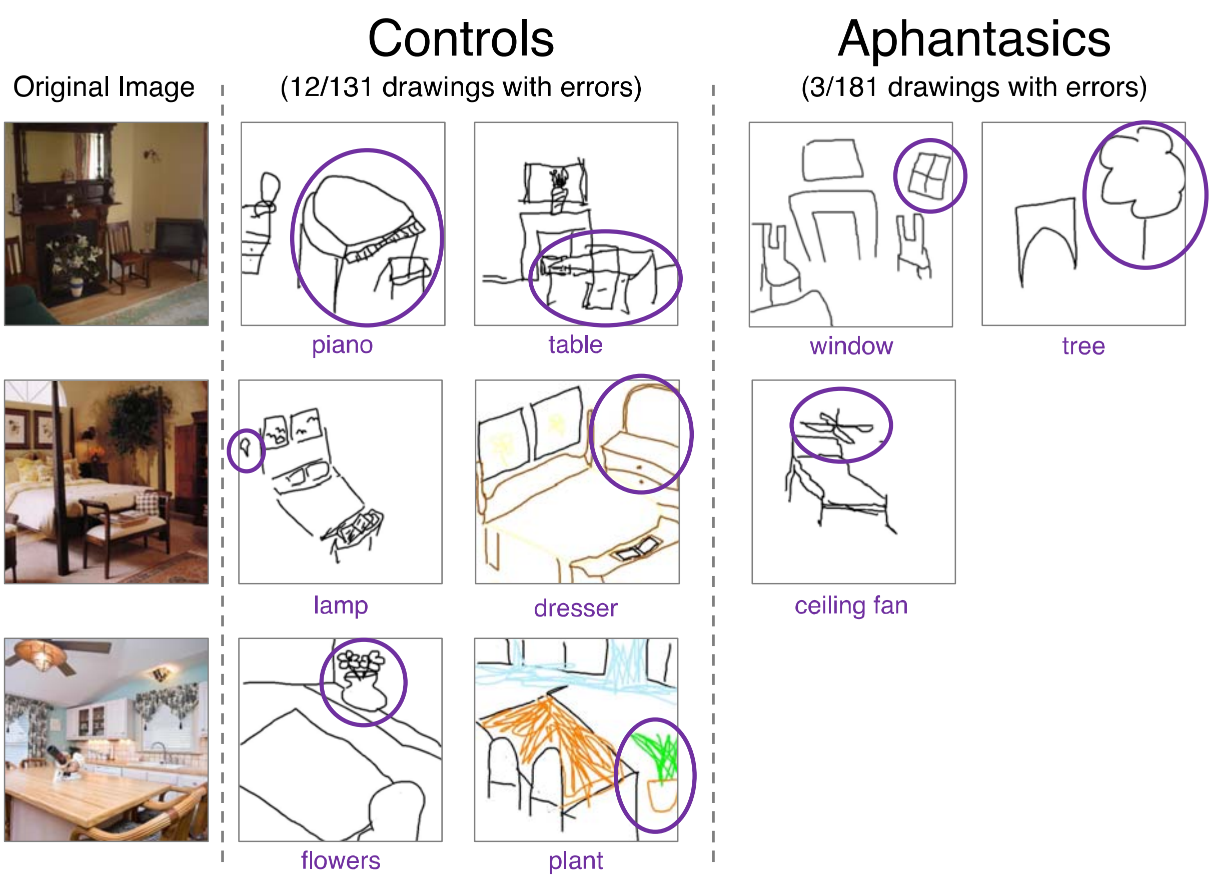


Fig. 5. False object memories in the drawings. Examples of the false object memories made by participants in their memory drawings, with the inaccurate objects circled. Control participants made significantly more

errors, with only 3 out of 181 total aphantasic drawings containing a falsely remembered object. Note, all aphantasic errors were also transpositions from other drawings.

As another metric of memory error, we also coded whether a drawing was edited or not, based on tracked mouse movements. A drawing was scored as edited if at least one line was drawn and then erased during the drawing. Significantly more memory drawings by control participants had editing than those by aphantasic participants (aphantasic: 27.6%, control: 46.6%; χ^2=11.90, p=6.63 × 10^{-4}). There was no significant difference in editing between groups for the perception drawings (aphantasic: 37.4%, control: 47.7%; χ^2=3.31, p=0.069), indicating these differences are not due to differences in effort.

Discussion

Through a drawing task with a large online sample, we conducted an in-depth characterization of the mental representations held by congenital aphantasics, a recently identified group of individuals who self-report the inability to form voluntary visual imagery. We discover that aphantasics show impairments in object memory, drawing fewer objects, containing less color. Further, we find evidence for greater dependence on symbolic information in the task, with more text in their drawings and common self-reporting of verbal strategies. However, aphantasics show no impairments in spatial memory, positioning objects at accurate locations with the correct sizes. Further, aphantasics show significantly fewer errors in memory, with fewer falsely recalled objects, and less correction of their drawings. Importantly, we observe no significant differences between controls and aphantasics when drawing directly from an image, indicating these differences are specific to memory and not driven by differences in effort, drawing ability, or perceptual processing.

Collectively, these results point to a dissociation in imagery between object-based information and spatial information. In addition to selective deficits in object memory over spatial memory, aphantasics subjectively report a lower preference for object imagery compared to spatial imagery in the OSIQ. This supports the previous findings in the smaller dataset (N=15) of Keogh & Pearson [2], which first reported differences in OSIQ measures.

Further, participants' reported object imagery abilities correlated with the number of objects they drew from memory. These consistent results both confirm the OSIQ as a meaningful measure, while also demonstrating how such deficits can be captured by a behavioral measure such as drawing. While a similar dissociation between object and spatial memory has been observed in other paradigms and populations, this is the first study to identify this in a population of individuals in the absence of trauma or changes in brain pathology. Cognitive decline from aging and dementia have shown selective deficits in object identification versus object localization [35], owing to changes in the medial temporal lobe, where the perirhinal cortex is thought to contribute to object detail recollection, while the parahippocampal cortex contributes to scene detail recollection [36]. The neocortex is also considered to be organized along separate visual processing pathways, with ventral regions primarily coding information about visual features, and parietal regions coding spatial information [37-41]. These findings also suggest interesting parallels between the imagery experience of individuals with aphantasia and individuals that are congenitally totally blind, who have been shown to perform similarly to typically sighted individuals on a variety of spatial imagery tasks [42-45]. Neuroimaging of aphantasics will be an important next step, to see whether these impairments are manifested in decreased volume or connectivity of regions specific to the imagery of visual details, such as anterior regions within inferotemporal cortex [23,31,46,47] or medial parietal regions implicated in memory recall [30,48-50].

Further investigations on aphantasics will also provide critical insight on the nature of imagery, and how it compares to different forms of memory. While aphantasics show an impairment at recall performance, no evidence has shown impairments in visual recognition, and indeed our study also observes near-ceiling recognition performance. These results support other converging evidence pointing towards a neural dissociation in the processes of quick, automatic visual recognition and slower, elaborative visual recall [3, 51-54]. Aphantasics also report fully intact verbal recall abilities, and our results suggest that they may be using semantic strategies, in combination with accurate spatial representations, to compensate for their lack of visual imagery. In fact, in the current study, aphantasics' drawings from memory contained more text than those of controls, potentially indicating a semantic propositional coding of their memories to perform the task. Imagery of a visual stimulus thus may not necessarily be visual in nature; while forming a visual representation of the scene or object may be one way to undertake the task, there may be other, non-visual strategies to complete the task. Even in neurotypical adults, imagery-based representations in the brain may differ from perceptual representations of the same items [31]. Further neuroimaging investigations will lead to an understanding of the neural mechanisms underlying these different strategies.

Further, aphantasics' lower errors in memory (e.g., fewer falsely recalled objects compared to controls) could possibly reflect higher accuracy in semantic memory versus controls, to compensate for visual memory difficulties. Aphantasics may serve as an ideal population to probe the difference between visual and semantic memory and their interaction in both behavior and the brain. Additionally, while aphantasia has thus far only been quantified in the visual domain, preliminary work suggests that the experience may extend to other modalities [1]. Using a multimodal approach, researchers may be able to pinpoint neural differences in aphantasics across other sensory modalities, for instance, the auditory domain which has shown to have several characteristics similar to the visual domain [55-57].

Finally, these results serve as essential evidence to suggest that aphantasia is a valid experience, defined by the inability to form voluntary visual images with a selective impairment in object imagery. Previous work has shown relatively intact performance by aphantasics on imagery and visual working memory tasks [5], and some researchers have proposed aphantasia may be more psychogenic than a real impairment [8]. However, in the current study, we observe a selective impairment in object imagery for aphantasics in comparison to controls. Importantly, if such an impairment were caused by intentional efforts to demonstrate an impairment, we would expect decreased performance in spatial accuracy, decreased performance in the perceptual drawing task, or low ratings in all questions of the OSIQ rather than solely the object imagery component. However, in all of these cases, aphantasics performed identically with controls. In fact, aphantasics even showed higher memory precision than controls on some measures, including significantly fewer memory errors and fewer editing in their drawings. Further, the correlations between the VVIQ, OSIQ, and drawn object information lend validity to the self-reported questionnaires in capturing true behavioral deficits. This being said, while we observed a deficit in object memory for aphantasics, it was not a complete elimination of object memory abilities. Aphantasics were still able to draw a handful of objects from memory (five per image). While this moderate performance could be due to some preserved ability at object memory, this performance could also reflect the use of verbal lists of objects combined with intact, accurate spatial memory to reconstruct a scene. Future work will need to directly compare visual and verbal strategies, and push the limits to see what occurs when there is more visual detail than can be supported by verbal strategies.

In conclusion, leveraging the wide reach of the internet, we have been able to conduct an in-depth and large scale study of the nature of aphantasics' mental representations for visual images. Aphantasics have a unique mental experience that can provide essential insights into the nature of imagery, memory, and perception. Their drawings reveal a complex, nuanced story that show impaired object memory, with a combination of semantic and spatial strategies used to reconstruct scenes from memory. Collectively, these results suggest a dissocation in object and spatial information in visual memory.

Methods

Participants

N=115 adults participated in the main online experiment, while 2,795 adults participated in online scoring experiments on Amazon Mechanical Turk (AMT) of the drawings from the main experiment. Aphantasic participants for the main experiment were recruited from aphantasia-targeted forums, including "Aphantasia (Non-Imager/Mental Blindness) Awareness Group", "Aphantasia!" and Aphantasia discussion pages on Reddit. Control participants for the main experiment were recruited from the population at the University of Westminster, online social media sites such as Facebook and Twitter pages for the University of Westminster Psychology, and "Participate in research" pages on Reddit. Scoring participants were recruited from the general population of AMT.

No personally identifiable information was collected from any participants, and participants had to acknowledge participation in order to continue, following the guidelines approved by the University of Westminster Psychology Ethics Committee (ETH1718-2345) and the National Institutes of Health Office of Human Subjects Research Protections (18-NIMH-00696).

Main Experiment: Drawing Recall Experiment

The Drawing Recall Experiment was a fully online experiment that consisted of seven sections ordered: 1) study phase, 2) recall drawing phase, 3) recognition phase, 4) copied drawing phase, 5) The Vividness of Visual Imagery Questionnaire (VVIQ), the 6) Object-Spatial Imagery Questionnaire (OSIQ), and 7) basic demographic questions. The methods of the experiment are summarized in Fig. 1a.

For the study phase, participants were told to study three images in as much detail as possible. The images were presented at 500 x 500 pixels. They were shown each image for 10 s, presented in a randomized order with a 1 s interstimulus interval (ISI). These three images (see Fig 1a) were selected from a previously validated memory drawing study [3], as the images with the highest recall success, highest number of objects, and several unique elements compared to a canonical representation of its category. For example, the kitchen scene does not include several typical kitchen components such as a refrigerator, microwave, or stove, and does include more idiosyncratic objects such as a ceramic chef, zebra-printed chairs, and a ceiling fan. This is important as we want to assess the ability to recall unique visual information beyond just a coding of the category name (e.g., just drawing a typical kitchen). Participants were not informed what they would do after studying the images, to prevent targeted memory strategies.

Next, the recall drawing phase tested what visual representations participants had for these images through drawing. Participants were presented with a blank square with the same dimensions as the original images and told to draw an image from memory in as much detail as possible using their mouse. Participants drew using an interface like a simple paint program. They could draw with a pen in multiple colors, erase lines, and undo or redo actions. They were given unlimited time and could draw the images in any order. They were also instructed that they could label any unclear items. Once a participant finished a drawing, they then moved onto another blank square to start a new drawing. They were asked to create three drawings from memory, and could not go back to edit previous drawings. As they were drawing, their mouse movements were recorded to track timing and erasing behavior.

The recognition phase tested whether there was visual recognition memory for these specific images. Participants viewed images and were told to indicate whether they had seen each image before or not. The images consisted of the three images presented in the study phase as well as three new foil images of the same scene categories (kitchen, bedroom, living room). Matched foils were used so that recognition performance could not rely on recognizing the category type alone. All images were presented at 500 x 500 pixels. Participants were given unlimited time to view the image and respond, and a fixation cross appeared between each image for 200 ms.

The copied drawing phase had participants copy the drawings while viewing them, in order to see how participants perceive each image. This phase gives us an estimate of the participant's drawing ability and ability to use this drawing interface with a computer mouse to create drawings. This phase also measures the maximum information one might draw for a given image (e.g., you won't draw every plate stacked in a cupboard). Participants saw each image from the study phase presented next to a blank square. They were instructed to copy the image in as much detail as possible. The blank square used the same interface as the recall drawing phase. When they were done, they could continue onto the next image, until they copied all three images from the study phase. The images were tested in a random order, and participants had as much time as they wanted to draw each image.

Finally, participants filled out three questionnaires at the end. They completed the VVIQ [9], which measures the vividness of one's visual mental images, and currently serves as the main tool for diagnosing aphantasia. They also completed the more recent OSIQ [34], which separately measures visual imagery for object information and spatial information. Finally, participants provided basic demographics, basic information about their computer interface, and their experience with art. In these final questions, they indicated which component of the experiment was most difficult, and were able to write comments on why they found it difficult.

Online Scoring Experiments

In order to objectively and rapidly score the 692 drawings produced in the Drawing Recall Experiment, we conducted online crowd-sourced scoring experiments with a set of 2,795 participants on AMT. None of these participants took part in the Drawing Recall Experiment. For all online scoring experiments, scorers could participate in as many trials as they wanted, and were compensated for their time.

Object Selection Study

AMT scorers were asked to indicate which objects from the original images were in each drawing. This allows us to systematically measure how many and what type of objects exists in the drawings. They were presented with one drawing and five photographs of the original image with a different object highlighted in red. They had to click on all object images that were contained in the original drawing. Five scorers were recruited per object, with 909 unique scorers in total. An object was determined to exist in the drawing if at least 3 out of 5 scorers selected it.

Object Location Study

For each object, AMT scorers were asked to place an oval around that object in the drawing, in order to get information on the location and size accuracy of the objects in the drawings. AMT scorers were instructed on which object to circle in the drawing by the original image with the object highlighted in red, and only objects selected in the Object Selection Study were used. Five scorers were recruited per object, with 1,310 unique scorers in total. Object location and size (in both the x and y directions) were taken as the median pixel values across the five scorers.

Object Details Study

AMT scorers here indicated what details existed in the specific drawings. In a first AMT experiment, five scorers per object (N=304 total) saw each object from the original images and were asked to list 5 unique traits about the object (e.g., shape, material, pattern, style). A list of unique traits was then created for each object in the images. In a second AMT experiment, scorers were then shown each object in the drawings (highlighted by the ellipse drawn in the Object Location Study), and had to indicate whether that trait described the object or not. Five scorers were recruited per trait per drawn object, with 777 unique scorers in total.

False Memories Study

AMT scorers were asked to indicate "false memories" in the drawings—what objects were drawn in the drawing that didn't exist in the original image? Scorers were shown a drawing and its corresponding image and were asked to write down a list of all false objects. Nine scorers were recruited per drawing, with 337 unique scorers in total. An object was counted as a false memory if at least three scorers listed it.

Additional Drawing Scoring Metrics

In addition to the Online Scoring Experiments, other attributes were collected for the drawings. A blind scorer (the corresponding author) went through each drawing presented in a random order (without participant or condition information visible) and had to code *yes* or *no* for if the drawing 1) contained any color, 2) contained any text, and 3) contained any erasures. Erasures were quantified by viewing the mouse movements used for drawing the image, to see if lines were drawn and then erased, and did not make it into the final image.

References

[1] Zeman, A.Z.J., Dewar, M.T., & Della Sala, S. Lives without imagery – congenital aphantasia. *Cortex* **73**, 378-380 (2015).

[2] Keogh, R., & Pearson, J. The blind mind: No sensory visual imagery in aphantasia. *Cortex* **105**, 53-60 (2017).

[3] Bainbridge, W.A., Hall, E.H., & Baker, C.I. Highly diagnostic and detailed content of visual memory revealed during free recall of real-world scenes. *Nat. Commun.* **10**, 5. (2019).

[4] Galton, F. Statistics of Mental Imagery. *Mind* **5**, 301-318 (1880).

[5] Jacobs, C., Schwarzkopf, D.S., & Silvanto, J. Visual working memory performance in aphantasia. *Cortex* **105**, 61-73 (2018).

[6] Brons, L. Aphantasia, SDAM, and Episodic Memory. *Japan Assoc. Philos. Sci.* 28, 9-32 (2019).

[7] Wicken, M., Keogh, R., & Pearson, J. The critical role of mental imagery in human emotion: insights from Aphantasia. Preprint at https://www.biorxiv.org/content/10.1101/726844v1 (2019).

[8] de Vito, S., & Bortolomeo, P. Refusing to imagine? On the possibility of psychogenic aphantasia. A commentary on Zeman et al. (2015). *Cortex* **74**, 334-335 (2016).

[9] Marks, D.F. Visual imagery differences in the recall of pictures. *Br. J. Psychol.* **64**, 17-24 (1973).

[10] Watkins, N.W. Visual working memory performance in aphantasia. *Cortex* **105**, 41-52 (2018).

[11] Ross, B. Aphantasia: How it feels to be blind in your mind. *Facebook*, at https://www.facebook.com/notes/blake-ross/aphantasia-how-it-feels-to-be-blind-in-your-mind/10156834777480504/ (2016).

[12] Gallagher, J. Aphantasia: Ex-Pixar chief Ed Catmull says “my mind’s eye is blind”. *BBC News,* at https://www.bbc.com/news/health-47830256 (2019).

[13] Behrmann, M., & Avidan, G. Congenital prosopagnosia: face-blind from birth. *Trends Cogn. Sci.* **9**, 180-187 (2005).

[14] Zeman, A., et al. Loss of imagery phenomenology with intact visuo-spatial task performance: a case of 'blind imagination'. *Neuropsychologia* **48**, 145-155 (2010).

[15] Pearson, J. The human imagination: the cognitive neuroscience of visual mental imagery. *Nat. Rev. Neurosci.* **20**, 624–634 (2019).

[16] Dijkstra, N., Bosch, S.E., & van Gerven, M.A.J. Shared neural mechanisms of visual perception and imagery. *Trends Cogn. Sci.* **23**, 423-434 (2019).

[17] Marmor, G.S., & Zaback, L.A. Mental rotation by the blind: Does mental rotation depend on visual imagery? *J. Exp. Psychol. Hum. Percept. Perform.* **2**, 515-521 (1976).

[18] Brandt, S.A., & Stark, L.W. Spontaneous eye movements during visual imagery reflect the content of the visual scene. *J. Cogn. Neurosci.* **9**, 27-38 (1997).

[19] Kosslyn, S. *Image and Mind*. (Harvard Univ. Press, 1980).

[20] Kosslyn, S. M. Mental images and the brain. *Cogn. Neuropsychol.* **22**, 333–347 (2005).

[21] Pylyshyn, Z. The imagery debate: Analogue media versus tacit knowledge. *Am. Psychol. Assoc.* **88**, 16–45 (1981).

[22] Pylyshyn, Z. Return of the mental image: are there really pictures in the brain? *Trends Cogn. Sci.* **7**, 113–118 (2003).

[23] Ishai, A., Ungerleider, L., & Haxby, J. Distributed neural systems for the generation of visual images. *Neuron* **28**, 979–990 (2000).

[24] Kosslyn, S.M., Ganis, G., & Thompson, W.L. Neural foundations of imagery. *Nat. Rev. Neurosci.* **2**, 635-642 (2001).

[25] Schacter, D.L., et al. The future of memory: Remembering, imagining, and the brain. *Neuron* **76**, 677-694 (2012).

[26] Pearson, J., & Kosslyn, S.M. The heterogeneity of mental representation: Ending the imagery debate. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 10089-10092 (2015).

[27] Behrmann, M. The mind's eye mapped onto the brain's matter. *Curr. Dir. Psychol. Sci.* **9**, 50-54 (2000).

[28] Bartolomeo, P. The neural correlates of visual mental imagery: An ongoing debate. *Cortex* **44**, 107-108 (2008).

[29] Xiao, X., et al. Transformed neural pattern reinstatement during episodic memory retrieval. *J. Neurosci.* **37**, 2986-2998 (2017).

[30] Silson, E.H., et al. A posterior-anterior distinction between scene perception and scene construction in human medial parietal cortex. *J. Neurosci.* **39**, 705-717 (2019).

[31] Bainbridge, W.A., Hall, E.H., Baker, C.I. Distinct representational structure and spatial distribution for visual encoding and recall. Preprint at https://www.biorxiv.org/content/10.1101/842120v1 (2019).

[32] Stevens, A., & Coupe, P. Distortions in judged spatial relations. *Cogn. Psychol.* **10**, 422-437 (1978).

[33] Denis, M., & Cocude, M. Scanning visual images generated from verbal descriptions. *Eur. J. Cogn. Psychol.* **1**, 293-307 (1989).

[34] Blajenkova, O., Kozhevnikov, M., & Motes, M.A. Object-spatial imagery: A new self-report imagery questionnaire. *Appl. Cogn. Psychol.* **20**, 239-263 (2006).

[35] Reagh, A.M., et al. Greater loss of object than spatial mnemonic discrimination in aged adults. *Hippocampus* **26**, 417-422 (2016).

[36] Staresina, B.P., Duncan, K.D., & Davachi, L. Perirhinal and parahippocampal cortices differentially contribute to later recollection of object- and scene-related event details. *J. Neurosci.* **31**, 8739-8747 (2011).

[37] Farah, M., Hammond, K., Levine, D., & Calvanio. Visual and spatial mental imagery: dissociable systems of representation. *Cogn. Psychol.* **20**, 439–462 (1988).

[38] Ungerleider, L.G., & Haxby, J.V. 'What' and 'where' in the human brain. *Curr. Opin. Neurobiol.* **4**, 157-165 (1994).

[39] Corballis, M. Mental rotation and the right hemisphere. Brain Lang. **57**, 100–121 (1997).

[40] Carlesimo, G., Perri, R., Turriziani, P., Tomaiuolo, F., & Caltagirone, C. Remembering what but not where: independence of spatial and visual working memory in the human brain. *Cortex* **37**, 519–534 (2001).

[41] Kravitz, D.J., Saleem, K.S., Baker, C.I., & Mishkin, M. A new neural framework for visuospatial processing. *Nat. Rev. Neurosci.* **12**, 217-230 (2011).

[42] Eardley, A. F., & Pring, L. (2007). Spatial processing, mental imagery, and creativity in individuals with and without sight. European Journal of Cognitive Psychology, 19(1), 37-58.

[43] Kerr, N. H. The role of vision in "visual imagery" experiments: Evidence from the congenitally blind. *J. Exp. Psychol. Gen.* **112**, 265 (1983).

[44] Zimler, J., & Keenan, J. M. Imagery in the congenitally blind: How visual are visual images?. *J. Exp. Psychol. Learn. Mem. Cogn.* **9**, 269 (1983).

[45] Cattaneo, Z., et al. Imagery and spatial processes in blindness and visual impairment. *Neurosci. Biobehav. Rev.* **32**, 1346-1360 (2008).

[46] O'Craven, K.M., & Kanwisher, N. Mental imagery of faces and places activates corresponding stimulus-specific brain regions. *J. Cogn. Neurosci.* **12**, 1013-1023 (2000).

[47] Lee, S.H., Kravitz, D.J., & Baker, C.I. Disentangling visual imagery and perception of real-world objects. *Neuroimage* **59**, 4064-4073 (2012).

[48] Buckner, R.L., Andrews-Hanna, J.R., & Schacter, D.L. The brain's default network: anatomy, function, and relevance to disease. *Ann. N. Y. Acad. Sci.* **1124**, 1-38 (2008).

[49] Vilberg, K.L., & Rugg, M.D. Memory retrieval and the parietal cortex: A review of evidence from a dual-process perspective. *Neuropsychologia* **46**, 1787-1799 (2008).

[50] Ranganath, C., & Ritchey, M. Two cortical systems for memory-guided behaviour. *Nat. Rev. Neurosci.* **13**, 713-726 (2012).

[51] Jacoby, L.L. A process dissociation framework: Separating automatic from intentional uses of memory. *J. Mem. Lang.* **30**, 513-541 (1991).

[52] Holdstock, J. S., et al. Under what conditions is recognition spared relative to recall after selective hippocampal damage in humans? *Hippocampus* **12**, 341-351 (2002).

[53] Staresina, B. P., & Davachi, L. Differential encoding mechanisms for subsequent associative recognition and free recall. *J. Neurosci.* **26**, 9162-9172 (2006).

[54] Barbeau, E. J., Pariente, J., Felician, O., & Puel, M. Visual recognition memory: A double anato-functional dissociation. *Hippocampus* **21**, 929-934 (2011).

[55] Bunzeck, N., Wuestenberg, T., Lutz, K., Heinze, H.-J., & Jancke, L. Scanning silence: mental imagery of complex sounds. *Neuroimage* **26**, 1119–1127 (2005).

[56] Clarke, S., Bellmann, A., Meuli, R. A., Assal, G., & Steck, A. J. Auditory agnosia and auditory spatial deficits following left hemispheric lesions: evidence for distinct processing pathways. *Neuropsychologia* **38**, 797–807 (2000).

[57] Halpern, A. Mental scanning in auditory imagery for songs. *J. Exp. Psychol. Learn. Mem. Cogn.* **14**, 434–443 (1988).