



Original Articles

The perceptual and phenomenal capacity of mental imagery



Rebecca Keogh*, Joel Pearson

School of Psychology, University of New South Wales, Sydney 2052, Australia

ARTICLE INFO

Article history:

Received 20 July 2016

Revised 6 February 2017

Accepted 8 February 2017

Keywords:

Visual imagery

Capacity limits

Mental representations

Binocular rivalry

ABSTRACT

Despite the brain's immense processing power, it has finite resources. Where do these resource limits come from? Little research has examined possible low-level sensory contributions to these limitations. Mental imagery is a fundamental part of human cognition that bridges cognition with sensory representations. Hence, imagery serves as a good candidate sensory process for probing how low capacity limitations might extend down the processing hierarchy. Here we introduce a novel technique to measure the sensory capacity of mental imagery, while removing the need for memory and any direct subjective reports. Contrary to our dynamic phenomenological experience, we demonstrate that visual imagery is severely limited by the perceptual and phenomenal bottleneck of visual representation. These capacity limits appear to be independent of generation time, depend on visual feature heterogeneity, are attenuated by concurrent retinal stimulation and are endowed with good metacognition. Additionally, the precision of visual representation declines rapidly with the number of stimuli, which is governed by a simple power law. We anticipate that this assay will be important for mapping the limits of human information processing.

© 2017 Elsevier B.V. All rights reserved.

1. Introduction

Despite the great processing power of the human brain, when we are asked to remember or process multiple things at once our performance tends to decline with more items (Cowan, 2001; Franconeri, Alvarez, & Cavanagh, 2013; Halford, Cowan, & Andrews, 2007; Miller, 1956). These capacity limits are found across most cognitive domains, such as general intelligence (Neubauer & Fink, 2009), multi-tasking (Monsell, 2003; Rogers & Monsell, 1995), auditory and visual short-term memory (Bays & Husain, 2008; Grimault et al., 2014; Luck & Vogel, 1997; Vogel & Machizawa, 2004) and visual attention (Fougnie & Marois, 2006; Palmer, 1990). Much of the research into human cognitive capacity limits to date has focused on high-level working memory and attentional capacity limits, however exactly where these capacity limits originate is still an open question. Surprisingly, relatively little research has examined any low-level sensory contributions to these limitations, such as the inherent two-dimensional map-like representation of the visual cortex, which likely intrinsically limits the amount of visual information that can be concurrently held. Mental imagery is a primary part of human cognition that bridges high-level cognition with low-level sensory representations via functional sensory simulations. Hence, imagery serves as a good

visual process to utilize for probing how low these capacity limitations might extend down the processing hierarchy.

Mental imagery research suggests that both the vividness and sensory strength of mental imagery plays an important role in almost any cognitive function that involves some form of sensory simulation. For example, evidence suggests visual imagery is utilized during visual working memory maintenance (Albers, Kok, Toni, Dijkerman, & de Lange, 2013; Keogh & Pearson, 2011, 2014), when remembering the past or thinking about the future (Byrne, Becker, & Burgess, 2007; D'Argembeau & Van der Linden, 2006), making moral decisions (Gaesser & Schacter, 2014), language comprehension (Bergen, Lindsay, Matlock, & Narayanan, 2007; Zwaan, Stanfield, & Yaxley, 2002), spatial navigation (Ghaem et al., 1997), affective forecasting and eye witness memory (Dobson & Markham, 1993; Gilbert & Wilson, 2007). Surprisingly imagery vividness and strength is somewhat elevated in many psychiatric and neurological populations (Matthews, Collins, Thakkar, & Park, 2014; Sack, van de Ven, Etschenberg, Schatz, & Linden, 2005; Shine et al., 2015). However, despite the overarching importance of visual imagery in daily life, very little research has investigated the capacity limits to what can be imagined. Here we attempt to examine the capacity limits of creating and maintaining mental images in mind in isolation of overt memory.

Much mental imagery research is dependent on self-reported vividness ratings, sensory strength measures, performance on a mental rotation or manipulation task, or through the indirect impact of imagery on other stimuli. Many early visual imagery

* Corresponding author.

E-mail address: rebeckakeogh@gmail.com (R. Keogh).

studies also asked participants to imagine real-world objects and make comparative judgments about the images. For example, a classic imagery study by Kosslyn, Ball, and Reiser (1978) found that scanning larger images in mind took longer than smaller items, similar to how scanning larger images presented perceptually takes longer than smaller ones. Numerous studies since then have found that imagining a simple picture results in very similar neural and behavioral processes to perception (Ishai & Sagi, 1995; Kosslyn, 1999, 2005; Kosslyn, Alpert, & Thompson, 1997; Kosslyn, Thompson, & Alpert, 1997; Pearson, Clifford, & Tong, 2008). Some early research also delved into the construction of complex visual images, and found evidence that when participants were instructed to imagine an image of an animal as a whole, or to construct the same image which had been broken into parts or ‘units’ to be ‘glued’ back together, they took longer in the ‘gluing’ condition, suggesting it was possible for people to combine multiple units of an image in the mind’s eye (Kosslyn, Reiser, Farah, & Fliegel, 1983).

Another experiment found that participants took longer to imagine identical geometric shapes when they were described as being composed of multiple shapes rather than only a few. A further experiment in this paper found that participants were able to construct a scene of multiple images placed close or far from each other based on a verbal description. They found that when a participant scanned from one image to another in the imagined scene it took longer for far versus close images. These studies show that individuals have the capacity to imagine multiple images or ‘units’ at once however, they do not provide any information about the quality, or capacity limits, of the units the individuals in these studies imagined.

To objectively assess potential capacity limits of visual imagery, independently of limits to working memory, we devised a novel version of the binocular rivalry paradigm, previously used to measure the sensory strength of a single mental image (Chang, Lewis, & Pearson, 2013; Keogh & Pearson, 2011, 2014; Pearson, 2014; Pearson et al., 2008; Sherwood & Pearson, 2010). This method has previously been used to assess the sensory strength of a single mental image through its effect on subsequent binocular rivalry. In this paradigm individuals are cued to imagine one of two binocular rivalry patterns for a few seconds prior to a brief rivalry presentation. Following the imagery formation, there is a higher probability of the imagined pattern being dominant during this brief rivalry presentation. This effect is known as priming and allows us to obtain an objective measure of sensory imagery strength (measured as the percentage of trials primed by imagery, see Pearson, 2014 for a review of the method). This measure of imagery allows us to avoid a reliance on self-report, reaction times, simple featural judgments and any possible effects of concurrent visual attention (Pearson, 2014; Pearson, Naselaris, Holmes, & Kosslyn, 2015).

In the new technique for measuring visual imagery capacity, instead of only imagining a single pattern, participants are cued to imagine from one to seven colored Gabor patches simultaneously for 6 s. To eliminate an overt reliance on memory for the location and structure of each imagined item we presented participants with multiple imagery ‘placeholders’. These placeholder cues were two small dark grey lines that informed participants of the horizontal or vertical orientation, color and location for each to-be-imagined colored Gabor pattern around an invisible circular array (Fig. 1A; color cues shown). The logic behind using such cue placeholders was to negate contributions of memory for the location and orientation of the imagined patterns. Following a period of image generation, participants were presented with a brief (750 ms), small, single binocular rivalry display at only one of the many placeholder locations (chosen at random), to probe the prior image strength at that single location, participants then reported the dominant rivalry pattern (red-horizontal, green-vertical or a mix; see Section 2 for stimulus details).

As in prior work (Chang et al., 2013; Keogh & Pearson, 2011, 2014; Pearson, 2014; Pearson, Rademaker, & Tong, 2011b; Pearson et al., 2008; Sherwood & Pearson, 2010), the strength of the mental image was taken as percent primed (i.e. the percent of trials in which the imagined pattern matched the reported pattern in subsequent rivalry), compared to the chance score of 50% (equal number of red and green patterns) collapsed across the multiple placeholder locations. We then grouped the data based on the set size of the imagined array to look for any capacity-like set size effects. If there are limitations to what we can imagine we should expect that when subjects are required to imagine multiple images the priming effect of imagery should decrease, while if imagery is limitless, priming should remain constant across all set sizes.

2. Materials and methods

2.1. Participants

A total of 72 participants participated in these experiments (aged from 18 to 35): experiment 1 (N = 4, 2 female), experiment 2 (N = 4, 2 female), experiment 3 (Analyzed participants: N = 13, 9 female, 2 participants were not used in analysis due to attrition and 2 for very high mock priming), experiment 4 (Analyzed participants: N = 7, 3 female, 4 participants were removed from the analysis due to too many mixed percept reports (more than 33%, N = 3) and attrition (N = 1)), experiment 5 (N = 15, 9 female, 7 participants removed due to too many mixed percepts (33%) or low priming (less than 50% for one item)), and experiment 6 (Analyzed participants: N = 6, 4 female, 8 participants were removed due to attrition (N = 2) and due to too many mixed percepts or low priming (more than 33% or priming for one image less than 55%, N = 6).

The majority of participants were students who completed the experiments in exchange for course credit; five of the participants were experienced psychophysical subjects and one of the authors (RK) participated in all of the experiments (except for the background luminance experiment). To ensure the data are not driven by the inclusion of one of the authors all experiments were also analyzed without RK’s data in the [Supplementary material](#). All experiments were approved by the UNSW Human Research Ethics Advisory Panel (Psychology) and written consent was obtained from all participants.

2.2. Statistical analysis

For experiments 1–5 repeated measures ANOVA’s were carried out in SPSS. All post hoc analysis were two-tailed and controlled for multiple comparisons using the Bonferroni correction.

For experiment 3 the data was normalized using the following equation:

$$\text{Percent Primed}_{(\text{set size } (n))} / \text{Percent Primed}_{(\text{set size } (1))}$$

To analyze the homogeneity data only set sizes 2, 3 and 4 were used. A participant’s data was discarded if for any of homogeneity values (100, 75, 66.67 or 50%) there were less than 3 data points. This resulted in a total of 28 participant’s homogeneity data being included in the analysis.

For experiment 6 all data functions were fit in MATLAB using a sum of Gaussians and all data was first anchored to 50% priming at the 37 degrees point.

2.3. Apparatus

All experiments were performed in a blackened room on a 27 in. iMac with a resolution of 2560 × 1440 pixels, with a frame rate of 60 Hz. A chin rest was used to maintain a fixed viewing distance of

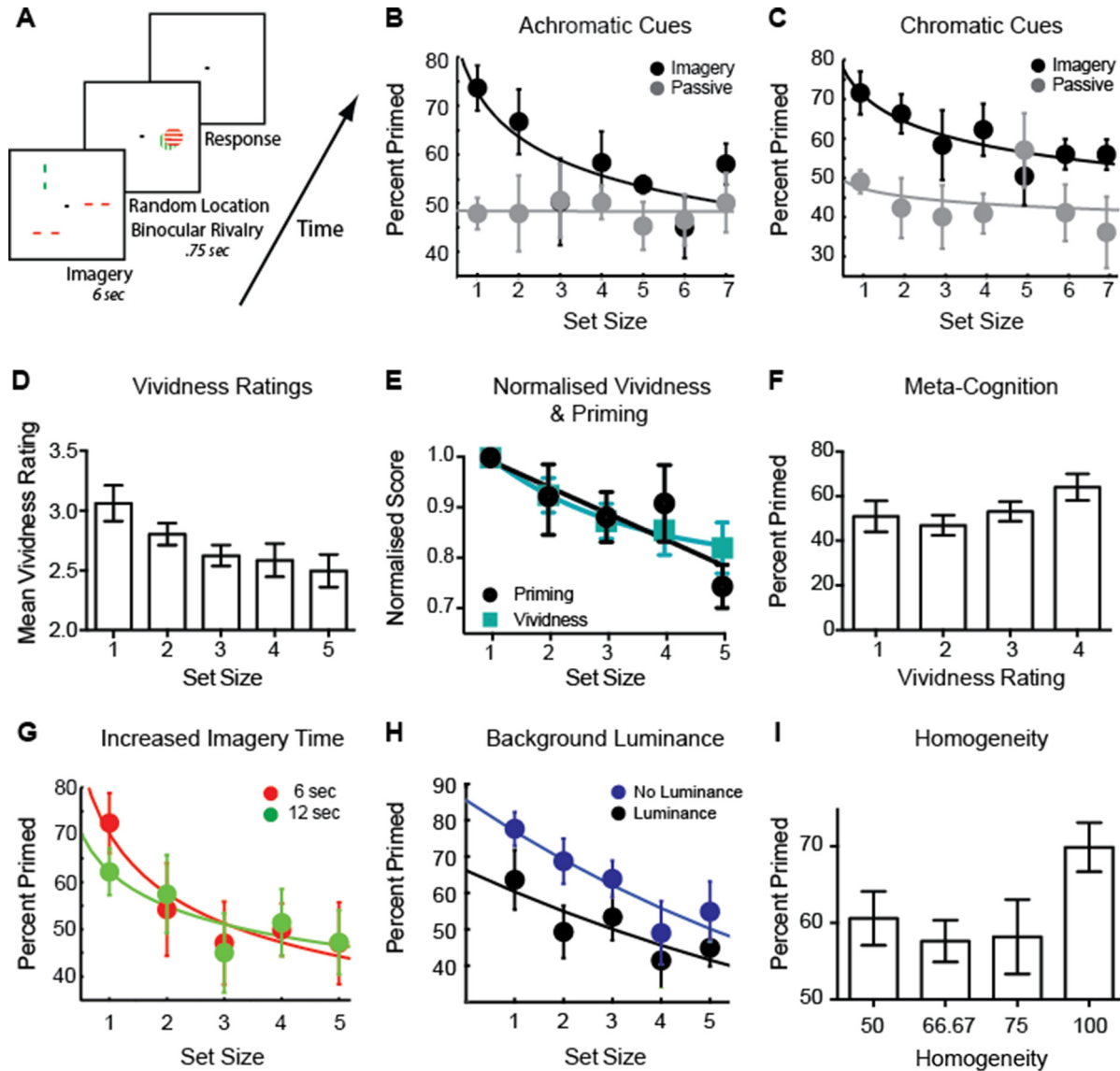


Fig. 1. (A) Visual mental representational capacity timeline. Participants were cued to imagine either red horizontal or green vertical lines between the placeholders. Following this, participants were presented with a single binocular rivalry display at any of the previous placeholder locations (randomly chosen). (B & C) Low-level sensory mental representation capacity (Achromatic cues and Chromatic cues respectively). Data shows percent primed on the Y-axis and set size on the X-axis. The black plot shows the data when participants are imagining the stimuli and the grey plots are for passive viewing. $N = 4$ for both achromatic and chromatic conditions. (D–F) Low-level sensory mental representation capacity for participants completing both the objective and subjective imagery measures. Panel D shows decreases in subjective ratings of mental representations vividness with increases in set size ($N = 13$). Panel E shows similar decreases in both imagery priming and mean vividness ratings with increases in set size (normalized data $N = 13$). Panel F shows participant's metacognition of their mental representations averaged across all set sizes ($N = 13$). (G) Data shows priming for 6 s (red line) or 12 s (green line) of imagery generation time ($N = 7$). (H) Effect of background luminance on low-level sensory representation capacity. Data shows percent primed on the Y-axis and set size on the X-axis ($N = 16$). The blue dots represent priming without background luminance and the black dots represent priming with background luminance. When luminance was present participants showed less priming than when the background remained black through the imagery period. (I) Percent primed for different levels of image homogeneity ($N = 18$). Panel shows level of priming for decreases in image homogeneity (from left to right). Decrease in image homogeneity resulted in reductions in the percentage of images primed. All error bars show \pm SEMs. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

57 cm. Participants wore red-green anaglyph glasses throughout all experiments.

2.4. Stimuli

A bulls-eye fixation point was presented in the middle of the screen for all experiments with a diameter of 0.02° and a mean luminance of 2.06 cd m^{-2} .

The binocular rivalry displays consisted of one green vertical (CIE $x = 0.268$ $y = 0.628$) and one red horizontal (CIE $x = 0.618$ $y = 0.358$) Gabor patch, 1 cycle° , Gaussian $\sigma = 2^\circ$, with a mean luminance of 13.1 cd m^{-2} for all experiments, except in the preci-

sion experiment where all the parameters were the same except the red and green displays were oriented either 0 , 11.5 , 22.5 , 30 , 37 deg from horizontal or vertical (always maintaining orthogonality).

Imagery placeholders consisted of two short lines (0.03°) separated by 1.4° degrees tilted either horizontally or vertically. The horizontal place holders were red (CIE $x = 0.568$ $y = 0.361$) and the vertical place holders were green (CIE $x = 0.272$ $y = 0.612$), except for in the first experiment and the last precision experiment where the lines were grey (luminance 0.41 cd m^{-2}). We used these stimuli as placeholders as we wanted stimuli that were small enough to not interfere with the creation of the images in mind

(previous work has shown that incoming visual information that overlaps with the location of the imagined patterns can corrupt imagery generation). The placeholders also needed to be able to supply both the orientation and color information to the participants so they did not need to overtly use memory. The reasoning behind our attempt to remove overt memory from the experiment is that memory is known to have strict capacity limits, and for this reason it would be difficult to tell which part of our observed capacity limits would be driven by general memory and which parts were specific to the creation of low-level sensory images. The placeholders were presented in a circular array (radius of 5.5°) around a bull's eye fixation point. The placeholders were placed on this circular array as far from each other as possible to avoid crowding. The distance between each Gabor patch depended on the set size and was set to 360/number of stimuli. i.e. when there were 3 placeholders each of the 3 images were presented 120 deg from each other in the circular array. The background was black in all of the experiments except for the luminance condition of the fifth experiment. During the luminance condition the background was ramped up to yellow (a mix of the green and red colors used in binocular rivalry patterns, with a luminance of 44 cd m²). This was done smoothly over the first and last second of the imagery period to avoid visual transients.

Mock trials were included for 25% of trials in all experiments (except the precision experiment due to time constraints). Mock trials were included to assess demand characteristics. Mock trials consisted of a spatial mix of a half red-horizontal and green-vertical Gabor patch. These displays had blurred edges and the exact split varied on each trial to resemble actual piecemeal rivalry. These mock patterns shared the same color, and size parameters as the real binocular rivalry displays.

2.5. Procedure

2.5.1. Eye dominance task

It is well documented that most people have one eye that is more 'dominant' during binocular rivalry, meaning that they will see one of the two images more often than the other. This eye dominance can have an effect on our measure of mental imagery, because if a participant primarily sees just one image, this eye-dominance will prevent any priming effect due to even strong imagery. To minimize this possibility, the perceptual strength of each rivalry pattern was adjusted to null any eye dominance, using a test-procedure prior to the actual imagery task, as previously documented (see Pearson, 2014 for discussion; Pearson, Rademaker, & Tong, 2011a; Pearson et al., 2008; Rademaker & Pearson, 2012; Sherwood & Pearson, 2010).

This eye dominance procedure involves adjusting the contrast of the two patterns until the perceptual competition between the two is 'balanced'. During this procedure, participants are presented with a brief binocular rivalry display (750 ms) followed by the selected dominant image at full contrast for 4 s, which results in an increased likelihood that an individual will see the opposite pattern in the subsequent rivalry presentation (due to weakened neural responses from the full contrast adaptor pattern e.g. due to adaptation). If this did not lead to a perceptual switch (if participants reported green and were presented with a full contrast green and then reported green again), the dominant rivalry pattern contrast was decreased by 2% and the suppressed pattern's contrast increased by 2%, a simple type of staircase procedure. This was done until the participants had perceptual switches on 80–100% of trials.

Following this pre-task test, the intensity (contrast and mean luminance) values from this first test were put into a peripheral eye dominance trial-run. In this trial-run participants were pre-

sented with a binocular rivalry display at a random location on an invisible circle with a radius of 5.5 deg for 750 ms and were asked to report which color they saw by pressing '1' for green, '2' for mixed, and '3' for red. They were instructed to keep their eyes on the bull's eye fixation point throughout the task and only view the binocular rivalry display in their peripheral vision. Participant's completed a block of 12 trials. If there was still heavy eye/pattern dominance for one of the two colors (i.e. the participant saw one of the two colors more than approximately 60% of the time) the prior eye-dominance test was run again until the perceptual competition between the two was more balanced in the peripheral vision. If a participant only ever saw one of the two colors, they did not continue onto the imagery capacity experiments.

2.5.2. Imagery capacity experiments

For the first two experiments participants were presented with 1–7 place holders for 6 s, the place holders were presented intermittently 1.5 s on/1.5 s off to prevent peripheral fading from occurring, which was an issue in pilot trials. Following this imagery period, binocular rivalry was displayed randomly at one of the prior placeholder locations for 750 ms and participants reported which pattern was dominant (1 = green vertical, 2 = equally mixed, 3 = red horizontal). Participants completed 5 blocks of 42 trials resulting in a total of 210 trials. In the passive condition of experiment 1 and 2 the procedure and timeline were the same except during the imagery condition participants were instructed to just passively fixate on the bulls-eye in the center of the placeholders. For the passive condition participants completed 5 blocks of 42 trials resulting in a total of 210 trials.

2.5.3. Phenomenal capacity limits and metacognition experiment

For the third experiment the procedure and stimuli was exactly the same as experiment one and two plus the inclusion trial-by-trial vividness ratings and the number of images to imagine were 1–5 (the number of images participants had to imagine was reduced as data from experiment 1 and 2 suggested that participants could not imagine more than 5 items simultaneously). In this condition after the 6 s imagery period participants were presented with the word 'vividness' which prompted them to rate the vividness of the image(s) they had just imagined on a scale of 1–4, with 1 = least vivid and 4 = most vivid. Participants completed 3 blocks of 40 trials, resulting in a total of 120 trials.

2.5.4. Increased time experiment

For experiment four the procedure was the same as experiment one and two (with no vividness rating and only 1–5 placeholders) with the addition of a 12 s imagery period as well as the original 6-s imagery period. Participants completed a total of 2 blocks of 40 trials per imagery time (2 six-second and 2 twelve-second imagery blocks), resulting in a total of 160 trials.

2.5.5. Irrelevant visual information (background luminance) experiment

For experiment five the procedure was the same as experiment one and two except there was the inclusion of a condition with a luminous background. During the imagery generation period the background ramped up to a yellow color and then back down to black over the first and last second of the imagery interval. Other than this all other parameters were the same. Participants completed 40 trials for both the luminance and no luminance conditions, resulting in a total of 80 trials.

2.5.6. Precision and imagery capacity experiment

For the precision experiment the parameters/time course of the experiment was exactly the same as experiment one and two except the binocular rivalry display that was presented was tilted

either 0, 11.5, 22.5, 30, 37 from horizontal or vertical (always maintaining orthogonality) and participants were only presented with 1–4 placeholder cues. Participants completed 12 blocks of 40 trials resulting in a total of 480 trials.

3. Results

3.1. Imagery capacity

We found that our measure of visual imagery (percent of trials primed) decreased with increases in set size (Fig. 1B, black plot), with a main effect of set size ($F(6, 18) = 4.20, p < 0.001$). To ensure that the oriented placeholders were not explicitly affecting subsequent rivalry, we also ran a passive condition with the same participants. Here all task parameters were the same, with the exception that participants only had to passively fixate (without generating any mental images). There was no reliable effect of the number of placeholder cues on binocular rivalry (Fig. 1B, grey plot; $F(6, 18) = 0.12, p = 0.99$). When directly comparing priming in the imagery and passive conditions there was no main effect of condition ($F(1, 3) = 3.17, p = 0.17$). However, post hoc analysis with a Bonferroni correction for multiple comparisons found that there was significantly more priming at set size one when participants were imagining compared to passive viewing ($p < 0.01$). We also performed a Bayesian analysis of the data and found that there was a moderate Bayes factor ($BF = 8.85$) supporting the hypothesis that imagery was driving our observed capacity functions, whereas in the passive condition there was moderate support for the null hypothesis ($BF = 0.14$), with passive viewing of the cues not leading to any effect on subsequent binocular rivalry. Prior work has demonstrated that the sensory strength, or visual energy, of a prior stimulus (perceptual or imagined) predicts the degree of the effect on subsequent rivalry (Brascamp, Knapen, Kanai, van Ee, & van den Berg, 2007; Pearson, 2014; Pearson & Brascamp, 2008; Pearson et al., 2008). As such, we interpret these data as a possible decline in the sensory strength of visual imagery as a function of set-size. That is, as subjects imagine more images, the images become weaker.

In the first experiment participants were still required to 'remember' the color of each imagined stimulus. To reduce any influence of color memory we ran a second experiment in which we added color to the oriented placeholder cues. This removed the need for participants to hold the color-orientation combination in memory. Again we found that priming decreased as a function of set size ($F(6, 18) = 2.91, p < 0.05$; Fig. 1C), while the passive condition again remained flat ($F(6, 18) = 1.15, p > 0.38$). When comparing priming in the imagery and passive conditions, there was a main effect of condition ($F(1, 3) = 652.92, p < 0.0001$), with participants showing significantly more priming when they were imagining compared to passively viewing the cues. There was a small Bayes Factor ($BF = 2.52$) supporting the hypothesis that imagery was driving our observed capacity limits, whereas in the passive condition there was moderate support for the null hypothesis ($BF = 0.47$).

Is it possible that these decreases in mental image strength as a function of set size are due to demand characteristics? To assess this possibility, we included mock trials in all of our experiments (except the precision experiment due to time constraints). Our mock trials consisted of a fake binocular rivalry stimulus, manufactured from a physical blend of two colored Gabor patches that was presented to both eyes therefore avoiding any real binocular rivalry. If the data was driven by demand or a non-sensory bias to report rivalry dominance in line with the prior cue independently of the actual percept, then these physical blend stimuli should be reported incorrectly as either red or green. We did not observe

any significant priming for mock trials across all participants (mock-trial priming compared to 50%: $t(34) = 1.47, p = 0.15, M = 51.32, SD = 5.34$; Sup. Fig. 1; see Section 2) suggesting that the observed results cannot be explained by demand characteristics.

3.2. Phenomenal capacity limits and metacognition

Next we asked whether the sensory imagery capacity limits were accompanied by a reduction in subjective ratings of vividness, which would suggest good metacognition (awareness of the decline in mental image strength) of internally generated imagery capacity-limits. To assess any changes in phenomenal vividness across set sizes participants performed the same task with the additional requirement that they rate the vividness for the whole imagined array on each individual trial. After imagery generation, participants reported their vividness on a scale of 1–4 (4 = most vivid, 1 = least vivid), with the left hand, before being presented with the single binocular rivalry display at one of the cued locations (Pearson et al., 2011b; Rademaker & Pearson, 2012) and reporting dominance with the right hand.

We found a similar pattern to the priming data, with the 'on-line' reports of imagery vividness also decreasing as a function of set size ($F(4, 48) = 7.31, p < 0.001$; Fig. 1D). Additionally, the decreases in percent primed and mean vividness ratings were very similar with a main effect of set size ($F(4, 48) = 6.61, p < 0.001$) and no significant difference between the two conditions ($F(1, 12) = 0.03, p = 0.87, Fig. 1E$). To assess metacognition (Flavel, 1979) of the representational capacity-limits we grouped the priming scores by the four levels of reported vividness. We found a main effect of the vividness category ($F(3, 36) = 3.20, p < 0.05$), with participants having significantly more priming for items they rated as more vivid. These data suggest that the sensory capacity functions reported above are not epiphenomenal to the subjective sensations of forming mental images, as both appear to have similar attenuation with greater capacity.

3.3. Increased imagery time experiments

Next we tested if the observed capacity function of mental representations might simply be caused by larger arrays requiring more generation time to reach the same level of priming. In a fourth experiment we tested if these capacity functions were specific to the 6 s generation time or generalized to longer imagination periods. Participants again performed the same task, generating the required number of patterns for either 6 or 12 s. We found that there was no clear difference between set-size functions in the 6 or 12 s conditions (Main effect of condition (6 or 12 s): $F(1, 6) = 0.06, p = 0.81, Fig. 1G$).

3.4. Irrelevant visual information (background luminance)

Our data suggest that capacity limits for both the sensory strength and phenomenal vividness of mental visual representations exist, and increases in generation time do not significantly affect these capacity limits. However it may still be possible that attentional limitations (Palmer, 1990) may be contributing to the sensory capacity functions. To assess whether attentional capacity limits might contribute to our results we ran a fifth study in which participants performed the same imagery task with or without a bright luminous background. Previous work has shown that the passive presence of a uniform luminous background during imagery generation attenuates its strength (Keogh & Pearson, 2011, 2014; Sherwood & Pearson, 2010), likely due to afferent visual stimulation's obligatory access to early visual areas, which might interfere with the voluntary formation of visual images. However,

the passive presence of background luminance does not disrupt the effects of visual attention on subsequent rivalry (Pearson et al., 2008). As such, if our results are driven by limits to visual attention spread across the place holders we should not expect concurrent luminance to have any effect on the imagery capacity functions. Fig. 1H shows that the presence of luminance reduced the level of priming across all set sizes, suggesting that sensory capacity functions are likely due to limited image-generation resources and not attention (Main effect of luminance: $F(1,14) = 6.9$, $p < 0.05$). The decrease in priming in the presence of background luminance also provides further evidence that our findings are not due to demand characteristics, as one should not expect uniform, passive luminance to affect non-sensory mechanisms.

These data demonstrate that the ability to intentionally generate visual images in mind, independent of task related memory, has its own severe capacity limits, but what exactly imposes these representational capacity limits on non-retinal imagery? As the imagery array becomes larger, are individuals simply excluding certain stimuli as they breach their capacity limit or do they attempt to create representations of all patterns, but at a lower resolution? One way to approach this question is to probe the effects of featural homogeneity across the stimuli array. If participants only imagine a subset of the images (limited by a rigid number of available 'slots' or 'generators') the degree of featural homogeneity across the array should not affect the capacity functions, whereas if participants have a flexible resource that is able to be spread out and generate multiple representations in mind, one would predict an interaction between capacity functions and featural homogeneity.

To investigate the effect of array homogeneity on the capacity of low-level sensory visual representations we looked at priming as a function of array homogeneity. When participants had to create a representation consisting of all the same features, e.g. green vertical at all locations, (proportion the same = 1; a purely homogeneous array), priming was higher than when participants had to create images of different orientations and colors (Fig. 1I). (Main effect of homogeneity $F(3,81) = 2.76$, $p < 0.05$). These data show an interaction between priming, or imagery strength, and imagery content, e.g. the homogeneity of the information, supporting a dynamic resource model of visual image generation.

3.5. Precision and imagery capacity

If dynamic imagery resources are responsible for the observed capacity limits, the precision of each item's representation should decline as the array size increases due to resources being spread thinly across more and more items. Whereas if imagery generation utilizes a strict rigid 'slot'-like resource, we should expect stable precision of each pattern's representation, as once an individual's number of rigid slots is exceeded no more items can be generated. Such a scenario would result in fewer images being generated, but with no loss in the precision of the items that are generated.

To measure the precision of these mental representations we took advantage of a known property of visual imagery, its selectivity for spatial orientation (Bergmann, Genc, Kohler, Singer, & Pearson, 2015; Pearson et al., 2008). Early physiology research has shown that cells in the visual cortex are selective for orientation (Hubel & Wiesel, 1962, 1968). That is, a given cell in primary visual cortex responds optimally for a certain spatial orientation and will fire less as this orientation moves further from the cell's preferred orientation. This results in a Gaussian shaped 'orientation tuning curve' when the orientation is plotted against cell firing. Previous work by Pearson et al. (2008) has shown that imagery strength (priming rivalry) is similarly orientation specific. This study first demonstrated that when an individual is presented with a weak Gabor pattern of a certain orientation, say vertical,

subsequent rivalry is primed, i.e. they are more likely to see the vertical display, however the further away the images in the binocular rivalry display are tilted from the presented orientation priming drops off. Interestingly, this same orientation specificity was shown to occur with imagined Gabor patterns. We reasoned that the bandwidth for orientation selectivity might serve as a useful measure of a pattern's representational precision. The broader the imagery orientation-tuning curve, the lower the precision of the imagined stimulus, and conversely the narrower the function, the higher the precision.

To obtain psychophysical orientation-tuning functions for mental images, we adapted our capacity procedure by changing the orientation of both rivalry patterns together by 0, 11.5, 22.5, 30 and 37 deg, while always maintaining orthogonality (Fig. 2C). This psychophysical technique has produced orientation-tuning curves for imagery of a single stimulus that closely mirror perceptual orientation-tuning curves (Bergmann et al., 2015; Pearson et al., 2008).

The two mental representation models: dynamic vs. strict rigid mental resources are exemplified in their extremes in Fig. 2A and B. The peak, or amplitude, of the orientation tuning function relates to the strength of the images in mind, while the width, or standard deviation, of the tuning curve provides information about the precision of the images generated in mind. According to a very strict rigid slot model once an individual's fixed capacity is exceeded any new information cannot be represented. This would result in a drop in priming for that particular pattern (e.g. towards 50%), without a corresponding reduction in the bandwidth of the tuning function (see Fig. 2A for hypothetical data).

However, if the source of a participant's sensory representations were a dynamic resource spread across multiple image representations, which interact and compete with each other for neural space, one would expect numerous representations to be held at ever decreasing levels of precision and strength. This would present as changes in both amplitude and bandwidth of the orientation tuning curves (see Fig. 2B for hypothetical data).

Fig. 2D shows orientation tuning curves for our experimental data for set sizes 1–4, data points fit with a difference of Gaussian. Both the amplitude and bandwidth of the tuning curves change sharply from set size 1–2 and stabilize across set sizes 3–4. Critically, the bandwidth of the tuning curves increased as a function of set size, suggesting image precision decreases when multiple patterns are generated.

To further quantify the relation between set size and imagery precision, we plotted precision ($1/SD(\text{set size } n)/1/SD(\text{set size } 1)$), as a function of set size. Again we see a pronounced drop in imagery precision between one and two items with the results nicely fit by a simple power law ($f(x) = 0.968x^{-1.306}$, $R^2 = 0.817$, Fig. 2E). Likewise, all our raw priming data together is fit well by a simple power law ($f(x) = 70.61x^{-0.2096}$, $R^2 = 0.946$, Fig. 2F).

4. Discussion

Together our data suggests that visual imagery has severe capacity limits to the amount of content that can be simultaneously generated. It is unlikely that a strict rigid 'slot'-like model drives these limits, with decreases in both strength and precision emerging with increases in total content. Further, the heterogeneity of the content places further limitations on representational sensory strength suggesting that the type of information being imagined, and not just the number of items, influences capacity limits. It is unlikely that such capacity limitations are driven by limits to attentional resources for two reasons. First, the perceptual capacity limits were accompanied by phenomenal limits, with participants reporting lower levels of imagery vividness with larger

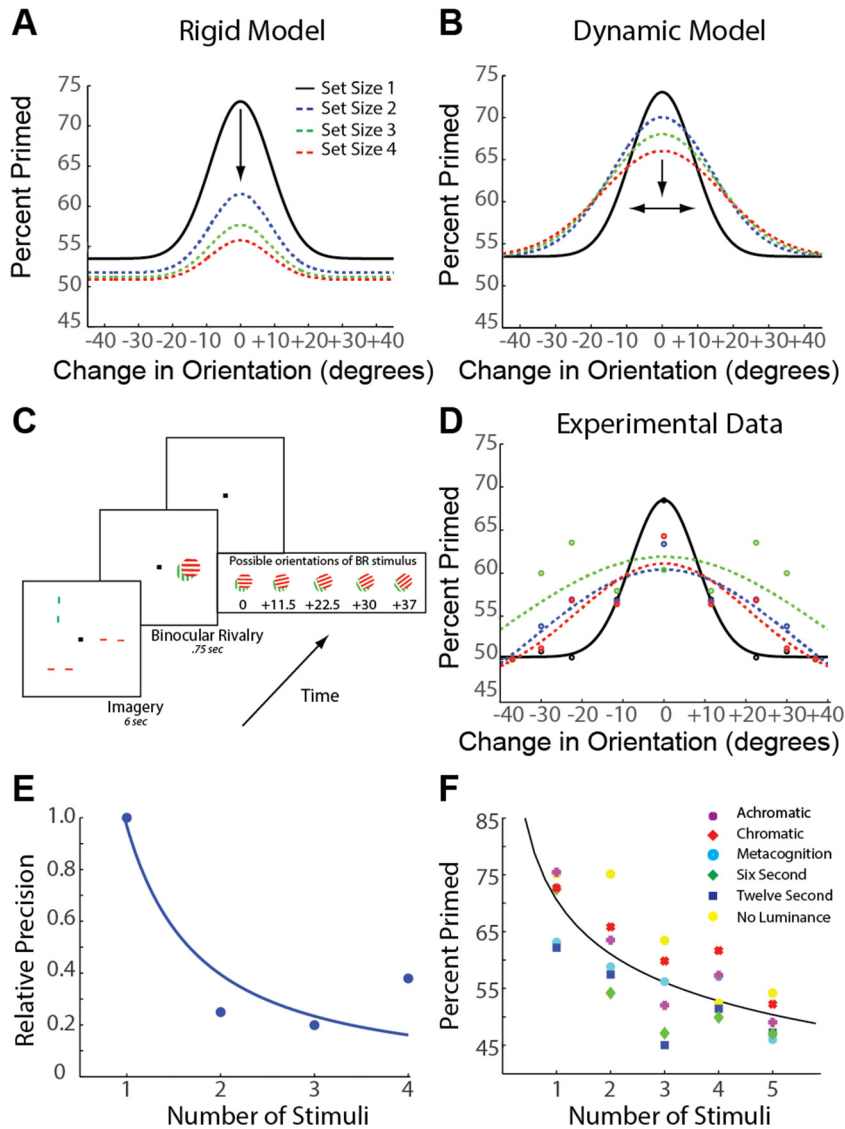


Fig. 2. (A) Theoretical results for the rigid model. If our results are due to a rigid ‘slot’ model we should expect to see decreases in amplitude without a change in the standard deviation (bandwidth). (B) Theoretical results for the dynamic model. If our results are due to dynamic resources we should expect to see decreases in amplitude and standard deviation (bandwidth). (C) Mental image capacity precision experiment timeline. Participants had to imagine one to four Gabor patches for 6 s. After this they were presented with a binocular rivalry display that either matched the imagined items (red horizontal and green vertical) or was tilted 11.5, 22.5, 30 or 37. (D) Data from our participants ($N = 6$). Data shows decreases in amplitude and increases in standard deviation when comparing set sizes 1 through 4. (E) Normalized precision values $(1/SD(\text{set size } n))/1/SD(\text{set size } 1)$ for each set size with a power function fit through the data. (F) Data averaged for all experiments (except mental image precision) across set size 1–5 with a power function fit through the averaged data. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

imagery arrays. Second, prior work has shown that the effects of attention on subsequent rivalry are immune to perturbation from concurrent uniform afferent visual stimulation (Pearson et al., 2008). That is, the presence of background luminance does not affect attention, likely due to attention involving higher-level brain areas that are not ‘contaminated’ by the presence of incoming irrelevant visual information. Here, the perceptual capacity functions were perturbed by concurrent uniform afferent visual stimulation (importantly the luminance was not present during the rivalry presentation), which should not be the case if attention is driving our observed capacity limits, as attention tends to be immune to this form of incoming sensory information.

The brain clearly has finite resources, but where might the specific functional limitations to image generation come from? One recent proposition is that cognitive functions that are contingent on two-dimensional map like representations will be constrained by both the content and limited anatomical space

(Franconeri et al., 2013). Area V1 represents visual information in 2D depictive cluster maps, and there is now much evidence that visual mental representations can involve depictive representations in area V1 (Albers et al., 2013; Cui, Jeter, Yang, Montague, & Eagleman, 2007; Kosslyn, Alpert, et al., 1997; Kosslyn et al., 1993; Naselaris, Olman, Stansbury, Ugurbil, & Gallant, 2015) see (Pearson & Kosslyn, 2015) for summary.

Our findings may have far reaching implications for many areas of cognitive science. With the advent of the cognitive revolution in the 20th century and the rise of computers and artificial intelligence the majority of theories on human thought have purported that thought is amodal in nature, that is, it does not require perceptual systems of the brain (Fodor, 1975; Pylyshyn, 2003). However more recently the hypothesis of grounded cognition has gained much popularity in cognitive science with scientists providing both behavioral and neuroimaging data that appears to support the involvement of modality specific activations during thought

(Amsel, Urbach, & Kutas, 2014; Barsalou, 2008, 2010). These theories suggest that the brain creates mental simulations during most cognitive activities. Mental simulations are described as the re-enactment of the modal states involved during perception and are a central feature for the majority of theories of various types of cognition such as working memory, language comprehension and reading, action intentions, implicit memory and even conceptual knowledge (Barsalou, 2009). These theories however tend to involve complex mental simulations made up of numerous features and items unconstrained by any representational capacity limits. Here we show that the ability to create mental representations, or images, is extremely limited by the content of the representation. These findings help to explain the limitations and errors routinely observed during mental simulations and guide better models and theories of grounded cognition.

5. Conclusions

In conclusion, our studies have demonstrated that visual imagery appears to have its own capacity limits with decreases in strength (Fig. 2F), vividness (Fig. 1E) and precision (Fig. 2D and E) with increases in set size. We propose that these capacity limits are due, in part, to an individual's ability, or lack thereof, to create precise 'maps' of the mental representations in the early visual cortex. Competition between generated content imposed by the inherent anatomical limitations of the two-dimensional representational architecture of V1, likely results in the observed limitations to visual representations in thought.

Author contributions

JP and RK developed the study concept. Testing and data collection was performed by RK. All data analysis was performed by RK. RK and JP wrote the manuscript.

Acknowledgements

JP is supported by Australian NHMRC grants APP1024800, APP1046198 and APP1085404; and a Career Development Fellowship APP1049596; and an ARC discovery project DP140101560.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.cognition.2017.02.004>.

References

- Albers, A. M., Kok, P., Toni, I., Dijkerman, H. C., & de Lange, F. P. (2013). Shared representations for working memory and mental imagery in early visual cortex. *Current Biology*, 23(15), 1427–1431. <http://dx.doi.org/10.1016/j.cub.2013.05.065>.
- Amsel, B. D., Urbach, T. P., & Kutas, M. (2014). Empirically grounding grounded cognition: The case of color. *Neuroimage*, 99, 149–157. <http://dx.doi.org/10.1016/j.neuroimage.2014.05.025>.
- Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology*, 59, 617–645. <http://dx.doi.org/10.1146/annurev.psych.59.103006.093639>.
- Barsalou, L. W. (2009). Simulation, situated conceptualization, and prediction. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 364(1521), 1281–1289. <http://dx.doi.org/10.1098/rstb.2008.0319>.
- Barsalou, L. W. (2010). Grounded cognition: Past, present, and future. *Topics in Cognitive Science*, 2(4), 716–724. <http://dx.doi.org/10.1111/j.1756-8765.2010.01115.x>.
- Bays, P. M., & Husain, M. (2008). Dynamic shifts of limited working memory resources in human vision. *Science*, 321(5890), 851–854. <http://dx.doi.org/10.1126/science.1158023>, pii: 321/5890/851.
- Bergen, B. K., Lindsay, S., Matlock, T., & Narayanan, S. (2007). Spatial and linguistic aspects of visual imagery in sentence comprehension. *Cognitive Science*, 31(5), 733–764. <http://dx.doi.org/10.1080/03640210701530748>.
- Bergmann, J., Genc, E., Kohler, A., Singer, W., & Pearson, J. (2015). Smaller primary visual cortex is associated with stronger, but less precise mental imagery. *Cerebral Cortex*. <http://dx.doi.org/10.1093/cercor/bhv186>.
- Brascamp, J. W., Knapen, T. H., Kanai, R., van Ee, R., & van den Berg, A. V. (2007). Flash suppression and flash facilitation in binocular rivalry. *Journal of Vision*, 7(12), 11–12. <http://dx.doi.org/10.1167/7.12.12>.
- Byrne, P., Becker, S., & Burgess, N. (2007). Remembering the past and imagining the future: A neural model of spatial memory and imagery. *Psychological Review*, 114(2), 340–375. <http://dx.doi.org/10.1037/0033-295X.114.2.340>.
- Chang, S., Lewis, D. E., & Pearson, J. (2013). The functional effects of color perception and color imagery. *Journal of Vision*, 13(10). <http://dx.doi.org/10.1167/13.10.4>.
- Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, 24(1), 87–114. discussion 114–185.
- Cui, X., Jeter, C. B., Yang, D., Montague, P. R., & Eagleman, D. M. (2007). Vividness of mental imagery: Individual variability can be measured objectively. *Vision Research*, 47(4), 474–478. <http://dx.doi.org/10.1016/j.visres.2006.11.013>, pii: S0042-6989(06)00556-6.
- D'Argembeau, A., & Van der Linden, M. (2006). Individual differences in the phenomenology of mental time travel: The effect of vivid visual imagery and emotion regulation strategies. *Consciousness and Cognition*, 15(2), 342–350. <http://dx.doi.org/10.1016/j.concog.2005.09.001>.
- Dobson, M., & Markham, R. (1993). Imagery ability and source monitoring: Implications for eyewitness memory. *British Journal of Psychology*, 84(Pt 1), 111–118.
- Flavel, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist*, 34(10), 6.
- Fodor, J. A. (1975). *The language of thought*. Massachusetts: Harvard University Press.
- Fougnie, D., & Marois, R. (2006). Distinct capacity limits for attention and working memory: Evidence from attentive tracking and visual working memory paradigms. *Psychological Science*, 17(6), 526–534. <http://dx.doi.org/10.1111/j.1467-2820.2006.01739.x>.
- Franconeri, S. L., Alvarez, G. A., & Cavanagh, P. (2013). Flexible cognitive resources: Competitive content maps for attention and memory. *Trends in Cognitive Sciences*, 17(3), 134–141. <http://dx.doi.org/10.1016/j.tics.2013.01.010>.
- Gaesser, B., & Schacter, D. L. (2014). Episodic simulation and episodic memory can increase intentions to help others. *Proceedings of the National Academy of Sciences of the United States of America*, 111(12), 4415–4420. <http://dx.doi.org/10.1073/pnas.1402461111>.
- Ghaem, O., Mellet, E., Crivello, F., Tzourio, N., Mazoyer, B., Berthoz, A., & Denis, M. (1997). Mental navigation along memorized routes activates the hippocampus, precuneus, and insula. *NeuroReport*, 8(3), 739–744.
- Gilbert, D. T., & Wilson, T. D. (2007). Propection: Experiencing the future. *Science*, 317(5843), 1351–1354. <http://dx.doi.org/10.1126/science.1144161>.
- Grimault, S., Nolden, S., Lefebvre, C., Vachon, F., Hyde, K., Peretz, I., ... Jolicœur, P. (2014). Brain activity is related to individual differences in the number of items stored in auditory short-term memory for pitch: Evidence from magnetoencephalography. *Neuroimage*, 94, 96–106. <http://dx.doi.org/10.1016/j.neuroimage.2014.03.020>.
- Halford, G. S., Cowan, N., & Andrews, G. (2007). Separating cognitive capacity from knowledge: A new hypothesis. *Trends in Cognitive Sciences*, 11(6), 236–242. <http://dx.doi.org/10.1016/j.tics.2007.04.001>.
- Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *Journal of Physiology*, 160, 106–154.
- Hubel, D. H., & Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. *Journal of Physiology*, 195(1), 215–243.
- Ishai, A., & Sagi, D. (1995). Common mechanisms of visual imagery and perception. *Science*, 268(5218), 1772–1774.
- Keogh, R., & Pearson, J. (2011). Mental imagery and visual working memory. *PLoS ONE*, 6(12), e29221. <http://dx.doi.org/10.1371/journal.pone.0029221>.
- Keogh, R., & Pearson, J. (2014). The sensory strength of voluntary visual imagery predicts visual working memory capacity. *Journal of Vision*, 14(12). <http://dx.doi.org/10.1167/14.12.7>.
- Kosslyn, S. M. (1999). The role of Area 17 in visual imagery: Convergent evidence from PET and rTMS (vol 284, pg 167, 1999). *Science*, 284(5416), 917.
- Kosslyn, S. M. (2005). Mental images and the brain. *Cognitive Neuropsychology*, 22(3), 333–347. <http://dx.doi.org/10.1080/02643290442000130>, pii: 769488928.
- Kosslyn, S. M., Alpert, N. M., & Thompson, W. L. (1997). Neural systems that underlie visual imagery and visual perception: A PET study. *Journal of Nuclear Medicine*, 38(5), 1205.
- Kosslyn, S. M., Alpert, N. M., Thompson, W. L., Maljkovic, V., Weise, S. B., Chabris, C. F., ... Buonanno, F. S. (1993). Visual mental-imagery activates topographically organized visual-cortex – Pet investigations. *Journal of Cognitive Neuroscience*, 5(3), 263–287.
- 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), 47–60.
- Kosslyn, S. M., Reiser, B. J., Farah, M. J., & Fliegel, S. L. (1983). Generating visual images: Units and relations. *Journal of Experimental Psychology: General*, 112(2), 278–303.
- Kosslyn, S. M., Thompson, W. L., & Alpert, N. M. (1997). Neural systems shared by visual imagery and visual perception: A positron emission tomography study. *Neuroimage*, 6(4), 320–334. <http://dx.doi.org/10.1006/nimg.1997.0295>.

- Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, 390(6657), 279–281. <http://dx.doi.org/10.1038/36846>.
- Matthews, N. L., Collins, K. P., Thakkar, K. N., & Park, S. (2014). Visuospatial imagery and working memory in schizophrenia. *Cognitive Neuropsychiatry*, 19(1), 17–35. <http://dx.doi.org/10.1080/13546805.2013.779577>.
- Miller, G. A. (1956). The magical number seven plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 63(2), 81–97.
- Monsell, S. (2003). Task switching. *Trends in Cognitive Sciences*, 7(3), 134–140.
- Naselaris, T., Olman, C. A., Stansbury, D. E., Ugurbil, K., & Gallant, J. L. (2015). A voxel-wise encoding model for early visual areas decodes mental images of remembered scenes. *Neuroimage*, 105, 215–228. <http://dx.doi.org/10.1016/j.neuroimage.2014.10.018>.
- Neubauer, A. C., & Fink, A. (2009). Intelligence and neural efficiency. *Neuroscience and Biobehavioral Reviews*, 33(7), 1004–1023. <http://dx.doi.org/10.1016/j.neubiorev.2009.04.001>.
- Palmer, J. (1990). Attentional limits on the perception and memory of visual information. *Journal of Experimental Psychology: Human Perception and Performance*, 16(2), 332–350.
- Pearson, J. (2014). New directions in mental-imagery research: The binocular-rivalry technique and decoding fMRI patterns. *Current Directions in Psychological Science*, 23(3), 178–183. <http://dx.doi.org/10.1177/0963721414532287>.
- Pearson, J., & Brascamp, J. (2008). Sensory memory for ambiguous vision. *Trends in Cognitive Sciences*, 12(9), 334–341. <http://dx.doi.org/10.1016/j.tics.2008.05.006>.
- Pearson, J., Clifford, C. W. G., & Tong, F. (2008). The functional impact of mental imagery on conscious perception. *Current Biology*, 18(13), 982–986. <http://dx.doi.org/10.1016/j.cub.2008.05.048>.
- Pearson, J., & Kosslyn, S. M. (2015). The heterogeneity of mental representation: Ending the imagery debate. *Proceedings of the National Academy of Sciences of the United States of America*, 112(33), 10089–10092. <http://dx.doi.org/10.1073/pnas.1504933112>.
- Pearson, J., Naselaris, T., Holmes, E. A., & Kosslyn, S. M. (2015). Mental imagery: Functional mechanisms and clinical applications. *Trends in Cognitive Sciences*.
- Pearson, J., Rademaker, R. L., & Tong, F. (2011a). Evaluating the mind's eye: The metacognition of visual imagery. *Psychological Science*, 22(12), 1535–1542. <http://dx.doi.org/10.1177/0956797611417134>.
- Pearson, J., Rademaker, R. L., & Tong, F. (2011b). Evaluating the mind's eye: The metacognition of visual imagery. *Psychological Science*, 22(12), 1535–1542. <http://dx.doi.org/10.1177/0956797611417134>.
- Pylyshyn, Z. (2003). Return of the mental image: Are there really pictures in the brain? *Trends in Cognitive Sciences*, 7(3), 113–118.
- Rademaker, R. L., & Pearson, J. (2012). Training visual imagery: Improvements of metacognition, but not imagery strength. *Frontiers in Psychology*, 3, 224. <http://dx.doi.org/10.3389/fpsyg.2012.00224>.
- Rogers, R. D., & Monsell, S. (1995). Costs of a predictable switch between simple cognitive tasks. *Journal of Experimental Psychology: General*, 124(2), 24.
- Sack, A. T., van de Ven, V. G., Etschenberg, S., Schatz, D., & Linden, D. E. (2005). Enhanced vividness of mental imagery as a trait marker of schizophrenia? *Schizophrenia Bulletin*, 31(1), 97–104. <http://dx.doi.org/10.1093/schbul/sbi011>.
- Sherwood, R., & Pearson, J. (2010). Closing the mind's eye: Incoming luminance signals disrupt visual imagery. *PLoS ONE*, 5(12), e15217. <http://dx.doi.org/10.1371/journal.pone.0015217>.
- Shine, J. M., Keogh, R., O'Callaghan, C., Muller, A. J., Lewis, S. J., & Pearson, J. (2015). Imagine that: Elevated sensory strength of mental imagery in individuals with Parkinson's disease and visual hallucinations. *Proceedings of the Royal Society B: Biological Sciences*, 282(1798), 20142047. <http://dx.doi.org/10.1098/rspb.2014.2047>.
- Vogel, E. K., & Machizawa, M. G. (2004). Neural activity predicts individual differences in visual working memory capacity. *Nature*, 428(6984), 748–751. <http://dx.doi.org/10.1038/nature02447>. pii.
- Zwaan, R. A., Stanfield, R. A., & Yaxley, R. H. (2002). Language comprehenders mentally represent the shapes of objects. *Psychological Science*, 13(2), 168–171.