Smaller Primary Visual Cortex Is Associated with Stronger, but Less Precise Mental Imagery

Johanna Bergmann, Erhan Genç, Axel Kohler, Wolf Singer, and Joel Pearson

Abstract

Despite mental imagery’s ubiquitous role in human perception, cognition and behavior, one standout question remains unanswered: Why does imagery vary so much from one individual to the next? Here, we used a behavioral paradigm that measures the functional impact of a mental image on subsequent conscious perception and related these measures to the anatomy of the early visual cortex estimated by fMRI retinotopic mapping. We observed a negative relationship between primary visual cortex (V1) surface area and sensory imagery strength, but found positive relationships between V1 and imagery precision (spatial location and orientation). Hence, individuals with a smaller V1 tended to have stronger, but less precise imagery. In addition, subjective vividness of imagery was positively related to prefrontal cortex volume, but unrelated to V1 anatomy. Our findings present the first evidence for the importance of the V1 layout in shaping the strength of human imagination.

Key words: early visual cortex, gray matter surface size, individual differences, primary visual cortex, visual imagery

Introduction

Visual imagery enables us to generate an image of an object even in its absence, via retrieving, modifying and recombining sensory information from memory (Kosslyn et al. 2001). The ability to create, sustain and modify images in mind plays a fundamental role in everyday behavior as well as in many mental disorders, where it plays a both a symptomatic and therapeutic role (Foà et al. 1980; Hirsch and Holmes 2007). However, despite imagery’s ubiquity in behavior and mental disorders, little is known as to the cause of strong versus weak imagery. In the 135 years since Sir Francis Galton first documented the large individual variance in self-reported imagery vividness (Galton 1880), empirical research has struggled to explain the cause of such individual variability.

A large cortical network is involved in imagery generation, stretching from prefrontal areas to parietal, temporal, visual and even subcortical areas, with a substantial overlap with the Default Mode Network (Schacter et al. 2007; de Araujo et al. 2012; Ostby et al. 2012; Schlegel et al. 2013). Much of the debate in the past decades, however, has focused on the contribution of early visual areas, in particular the primary visual cortex (V1), with conflicting empirical results (D’Esposito et al. 1997; Chen et al. 1998; Kosslyn et al. 1999; e.g., Cattaneo et al. 2009). These contradictions were reconciled by a meta-analysis that suggested V1 activation might depend on the particular visual feature and degree of spatial resolution in the mental image (Kosslyn and Thompson 2003). In line with this, several studies
have managed to decode mental images from early visual cortex fMRI BOLD response (Albers et al. 2013; Horikawa et al. 2013; Schlegel et al. 2013).

In parallel work, recent brain imaging studies have begun to unfold the functional implications of the large individual variations in V1 surface size, which varies between human brains up to a factor of 4 (Filimonov 1952; Stensaa et al. 1974) and has a strong genetic component (Panizzon et al. 2009; Winkler et al. 2010; Bakken et al. 2012; Chen et al. 2012; Wierenga et al. 2014). As the number of neurons in V1 correlates positively with surface size (Leuba and Krafsur 1994), differences in the macroscopic layout of V1 can be translated into individual differences in the number of neurons in this area. Accordingly, several studies have been able to link individual differences in conscious experience to anatomical variation. A larger V1 has been associated with higher spatial orientation (Song, Schwarzkopf and Rees 2013) and location sensitivity (Song et al. 2015), slower propagation of binocular rivalry-induced perceptual waves ( Genç et al. 2013) and lower susceptibility to contextual illusions (Schwarzko et al. 2011; Schwarzko and Rees 2013). The observation that a larger V1 is generally associated with higher visual precision is in line with the finding that a larger V1 tends to have smaller receptive fields (Harvey and Dumoulin 2011), that is, the size of the portion in the visual field in which stimulation triggers a neuron’s response is smaller. Interestingly, more recent research indicates that individual V1 anatomy might not only be linked to conscious perception but also to properties of higher cognitive functions, such as visual working memory (Bergmann et al. 2014) and attention (Vergheuse et al. 2014): In the studies, individuals with a larger V1 tended to have higher visual working memory storage and greater efficiency in selective spatial attention.

It has been suggested that early visual areas, in particular V1, might also be a crucial stage for the emergence of individual differences in imagery ability (Kosslyn et al. 2001), as it is here where the high-resolution details of a mental image are presumably processed. The aim of the current study was to look at the relationship between imagery ability and early visual cortex anatomy. In other words, if individual differences in early visual cortex anatomy shape conscious visual experience and memory, does the same also apply to visual imagery?

A link between imagery and V1 size seems plausible in light of the recent findings of a relationship between individual V1 anatomy and visual working memory (VWM; Bergmann et al. 2014) and the observation that VWM and imagery have highly similar early visual cortex representations (Albers et al. 2013). In addition, properties of VWM and imagery have been found to correlate with each other (Keogh and Pearson 2011, 2014).

To determine each individual’s imagery, we utilized a perceptual paradigm that measures imagery by quantifying its functional impact on subsequent binocular rivalry perception (Pearson et al. 2008; Pearson 2014). This method can exclude response and decisional bias (Pearson et al. 2008; Chang et al. 2013) and is reliable over time (Rademaker and Pearson 2012). More importantly perhaps, this measure of imagery is dissociable from visual attention, as imagery and binocular rivalry are separated in time and can even be separated by an attentionally demanding filler task (Pearson et al. 2008, 2011). Further, mental imagery generation is susceptible to changes in background luminance, whereas the effects of attention on subsequent rivalry are not. The generation time for imagery to affect subsequent rivalry is ~5 s and increases monotonically with more time, whereas attentional effects are observed at 1 s and remain constant across longer durations (Pearson et al. 2008). We obtained 4 measures of imagery from each participant: imagery strength, orientation- and location-specific imagery precision and subjective imagery vividness. To ensure the reliability of our behavioral measures, the test session was repeated after an average of ~2 weeks with each participant.

To examine the relationship between imagery and the neuroanatomical layout of the early visual cortex, we next estimated the boundaries of V1, V2 and V3 in each participant individually, using standard retinotopic mapping procedures with functional and structural magnetic resonance imaging (MRI).

Our results suggest the need for a clear distinction between the components of imagery that we looked at, with each showing clear, yet different relationships with early visual cortex anatomy: While imagery was stronger in individuals with a smaller V1, a larger V1 was linked to increased imagery precision. In contrast, subjective vividness showed no relationship to early visual cortex anatomy but revealed a significant relationship with prefrontal cortex volume in an exploratory whole-brain analysis.

Methods

Participants

Our sample consisted of 34 right-handed participants between 18 and 36 years old (median: 26 years; 13 males), with normal or corrected-to-normal vision. None of them had a history of psychiatric or neurological disorders. One participant was excluded from data analysis because of a misunderstanding of the task instructions. All participants were paid for participation. Written informed consent was obtained from all participants, and the study was approved by the ethics committee of the Max Planck Society.

Behavioral Experiments

Apparatus

Participants sat in a darkened room with dark walls, wearing red-green anaglyph glasses for the binocular rivalry presentation. Their position was stabilized with a chin rest, and the distance to the screen was 75 cm. The stimuli were presented on a CRT monitor (HP p1230; resolution, 1024 x 768 pixels, refresh rate: 150 Hz; visible screen size: 30° x 22.9°) and controlled by MATLAB R2010a (The MathWorks) using the Psychophysics Toolbox extension (Brainard 1996; Pelli 1997; Kleiner et al. 2007), running on Mac OSX, version 10.7.4.

Stimuli

In all experiments, the background of the screen remained black (<0.1 cd/m²). In both the standard and the orientation-specificity task, the circular Gaussian-windowed Gabor stimuli were presented centrally, spanning a radius of 4.6° around the fixation point in visual angle (thereby covering a diameter of 9.2°), one period subtending a length of 1.2°. The peak luminance starting value was ~0.71 cd/m² for the red grating and ~0.73 cd/m² for the green grating, which was then individually adjusted for each participant to compensate for eye dominance (see the description of the adjustment procedure below).

In the location-specific imagery precision task, the circular Gabor stimuli were presented peripherally in the left hemifield, one in the lower quadrant below the meridian and one in the upper quadrant, at a polar angle of 90° apart from each other and 45° to both sides of the meridian. Each spanned a diameter of 3.4°. Their inner boundary was at a distance of 2.3° to the central fixation point in visual angle their outer boundary at a distance of 5.7°. The faint circle, which was used as a location cue, had a luminance of 0.03 cd/m².
Procedure
Eye dominance adjustments. Individual variability in eye dominance can potentially lead to a perceptual bias for 1 of the 2 eyes. To control for this bias, participants first completed an adaptive testing procedure to carefully balance out the relative strength in luminance and contrast of the 2 rivalry Gabor patterns. The procedure was done for the central stimuli (standard, orientation-specificity paradigm) and the stimulus locations used in the location-specificity task separately, as the perceptual bias tends to vary across the visual field (Carter and Cavanagh 2007; Stanley et al. 2011).

In the adaptive procedure, which has been described previously (Pearson et al. 2008), contrast and then luminance were adjusted separately. Participants were asked to hold fixation and were first presented with the rivalry display (0.75 s) at a given contrast. When the stimuli disappeared from the screen, the participants responded via a button press which of the 2 stimuli had been dominant. They were then shown the dominant stimulus at full contrast over a period of 4 s to achieve adaptation, which weakens the neural response to this stimulus and thereby increases the probability of a switch in perceptual dominance in the succeeding binocular rivalry (Pearson and Clifford 2005). If the same stimulus became dominant 2 times in a row (or if the percept was mixed), the contrast of the dominant stimulus was decreased and the contrast of the suppressed stimulus was increased. This procedure continued until the intervening stimulus induced a perceptual switch in most or all of the rivalry presentations, indicating a balance in perceptual strength. A similar procedure was repeated with luminance, in which the luminance of the green vertical grating was adjusted to that of the red horizontal grating (which was kept constant) for an even finer individual adjustment of the perceptual bias for the 2 rivalry gratings.

Imagery Tasks
Imagery was measured in each participant using 3 tasks: a standard condition to determine imagery strength, and 2 tasks to determine orientation- and location-specific imagery precision (also see Fig. 1).

Standard imagery strength and orientation-specific imagery precision task. At the beginning of each trial, a cue (the letter “R” or “G”) appeared in the center of the screen for 1 s, indicating which of the 2 orthogonal rivalry gratings (the red horizontal or the green vertical Gabor stimulus) should be imagined. The cue was randomized on each trial and each stimulus had to be imagined an equal amount of times in every run. After an imagery interval of 7 s during which the screen remained black (8 s in the orientation-specificity condition), “vividness?” appeared at the center of the screen, and participants were asked to rate on a scale from 1 to 4 how vivid their mental image had been, with 1 representing the lowest and 4 the highest level of vividness. After having given their response via a button press, the rivalry gratings were flashed for 0.75 s, accompanied by a beep. When they disappeared from the screen, participants were instructed to indicate which 1 of the 2 gratings (the red horizontal or the green vertical) had been dominant or whether the percept had been mixed by pressing 1 of 3 keys (“1” = green vertical, “2” = mixed, “3” = red horizontal). Participants were asked to maintain fixation on the central fixation point throughout the course of a trial.

In the “standard imagery strength” condition, the spatial orientation of the presented rivalry stimuli always matched the spatial orientation of the stimuli the participants were asked to imagine (i.e., either horizontal or vertical). Participants completed 100 trials in the standard condition per session. Ten percent of the trials were catch trials with mock-rivalry stimuli, which consisted of a 75/25% mixture of the 2 gratings. In mock trials, participants should give a “mixed” response; if they fail to do so (and report the stimulus as unitary instead), this might be an indicator for a decisional bias. As our analyses showed, the probability of a decisional bias was very low (see Supplementary Results).

In the “orientation-specific imagery precision” task, participants had to imagine either the horizontal or the vertical grating as in the standard condition. However, the orthogonal rivalry gratings were presented in 1 of 5 possible spatial orientations (−45°, −22.5°, 0°, 22.5°, 45°), thereby either matching the imagined spatial orientation (at 0°) or deviating from it by ±45° or ±22.5°. The different rotation angles were randomized on each run and appeared an equal number of times each (30 per rotation angle per session).

Location-specific imagery precision task. Participants were asked to imagine either the red horizontal or the green vertical stimulus in 1 of 2 possible locations in the periphery, which was indicated by a faint circle throughout the imagery phase. The rivalry stimulus were then presented either at this location or at the other location where the stimulus had not been imagined. The location of stimulus presentation was randomized throughout the runs. The number of trials for each imagery location was balanced (120 trials per session in total, 60 per location; in 30 of these, imagery location matched the rivalry location and in the remaining 30, locations were different).

The behavioral test session was repeated after an average of ~2 weeks with each participant. All of the runs were divided into blocks of 33 trials, and participants were asked to take a rest in between.

In one participant, the eye dominance adjustments for the standard and orientation-specificity task had not been done correctly in the first session, resulting in a strong perceptual bias for 1 of the 2 rivalry patterns. Therefore, only the standard and orientation-specificity data sets from the second session were used for later analysis. In another participant, only one orientation-specificity data set was available. All behavioral data were checked for normal distribution using Shapiro–Wilk normality test. In the mean subjective vividness values, we detected a violation of the normal distribution assumption (W(33) = 0.877, P = 0.001). Further exploration using QQ-Plot and Boxplot analysis consistently revealed 2 extreme values of participant S15 and S20 (also see Supplementary Fig. 1D). As the surface-based morphometric analysis with Qdec fits the general linear model at each vertex to explain the behavioral data, which requires normality, we excluded the 2 data points from the analysis; for consistency, the 2 values were also excluded in the correlations between subjective vividness and imagery strength (Supplementary Fig. 1C) and between subjective vividness and early visual cortex surface (shown in Supplementary Fig. 6B). However, we also reevaluated the relationships after a re-inclusion of the 2 data points using nonparametric Spearman rank correlations and using ranks instead of mean subjective vividness values in the Qdec analysis (see Supplementary Results).

Behavioral measures. Imagery strength was defined as the percentage of trials in which the imagined stimulus matched the dominant stimulus in the succeeding rivalry presentation, imagery strength (% primed) = n_p / (n_p + n_n − n_s) × 100, with n_p being the number of trials in which the dominant stimulus was
primed by the imagined stimulus, \( n \) the total number of trials, \( n_m \) the number of mock trials and \( n_s \) the number of trials in which the participants had a mixed percept at rivalry presentation.

Imagery precision was estimated by computing the difference between 2 priming rates: in the location-specificity condition, by subtracting the priming rate at the location where the stimulus

---

**Figure 1.** The imagery task paradigm. (A) The standard imagery strength task. To measure individual imagery strength, we first asked the participants to either imagine the green vertical ("G") or red horizontal ("R") Gabor grating during an interval of 7 s. After the imagery phase, participants made a rating of how vivid their mental image had been by pressing the Buttons 1 ("hardly vivid") to 4 ("highly vivid"). Following this, binocular rivalry (a red horizontal and a green vertical Gabor grating) was presented for 0.75 s. When the stimuli disappeared, the participants indicated which one of the 2 stimuli had been dominant, or whether the percept had been mixed. (B) The orientation-specific imagery precision task. In this condition, participants again imagined either a vertical or horizontal Gabor grating. However, the binocular rivalry stimulus was randomly presented at 5 different orientations, thereby either matching the imagined grating (0°), or deviating from it (by ±45° or ±22.5°). Imagery precision was then estimated from the difference in %primed between the 0° and the ±22.5° condition, and the difference in %primed between the 0° and the ±45° condition. (C) The location-specific imagery precision task. In this condition, participants imagined the horizontal or vertical Gabor grating at 1 of 2 locations in the periphery, in the lower or upper left quadrant, which was cued by a faint circle. Binocular rivalry was then flashed randomly at 1 of these locations, which either matched the location of the imagined stimulus ("primed location") or was different to it ("unprimed location"). In this illustration, the imagery location is in the lower left quadrant, whereas the rivalry pair is flashed in the upper left quadrant (and hence, the "unprimed location"). Imagery precision was then estimated from the difference in %primed between the primed and the unprimed location.
had not been imagined (the “unprimed location”) from the priming rate at the location where the stimulus had been imagined (the “primed location”), Δ% \text{primed}_{\text{loc}} = \% \text{primed}_{\text{primed loc}} - \% \text{primed}_{\text{unprimed loc}}. In the orientation-specificity task, by subtracting the priming rates in the ±22.5° or ±45° orientation condition from the priming rate in the 0° orientation condition, Δ% \text{primed}_{\text{orient}} = \% \text{primed}_{\text{orient}} - \% \text{primed}_{\text{22.5°}}, and Δ% \text{primed}_{\text{orient}} = \% \text{primed}_{\text{45°}}, respectively. To obtain the % \text{primed}_{\text{22.5°}} and the % \text{primed}_{\text{45°}} value, the priming rates of the −22.5° and the 22.5° orientation condition (and of the −45° and the 45° orientation condition, respectively) were averaged.

### Neuroimaging Experiments

#### Apparatus

The neuroimaging experiments were conducted at the Brain Imaging Center Frankfurt am Main, Germany, in a Siemens 3T Trio MRI scanner (Siemens) with an 8-channel head coil and a maximum gradient strength of 40 mT/m. The neuroimaging procedure and data analysis is also described in Genç et al. (2013) and Bergmann et al. (2014).

Anatomical MRI data. T1-weighted anatomical images were acquired from all participants using an MP-RAGE sequence (TR = 2250 ms, TE = 2.6 ms, flip angle: 9°, FoV: 256 mm, resolution = 1 x 1 x 1 mm³).

Functional MRI (fMRI) data. A gradient-recalled echo-planar (EPI) sequence was used for the retinotopic mapping measurements (33 slices, TR = 2000 ms, TE = 30 ms, flip angle = 90°, FoV = 192 mm, slice thickness = 3 mm, gap thickness = 0.3 mm, resolution = 3 x 3 x 3 mm³). The retinotopic mapping stimuli were presented using an MR-compatible goggles system with 2 organic light-emitting-diode displays (MR Vision 2000; Resonance Technology Northridge), with a maximal visual field subtending 24° vertically and 30° horizontally. They were generated with a custom-made program based on the Microsoft DirectX library (Muckli et al. 2005).

Retinotopic mapping procedure. To acquire retinotopic maps of early visual areas V1, V2 and V3, our participants completed 2 functional MRI runs, a polar mapping and an eccentricity mapping run (e.g., Sereno et al. 1995; Wandell et al. 2007). In the polar mapping measurement, which is necessary to estimate the boundaries between the early visual areas, participants viewed a black and white checkerboard wedge that slowly rotated clockwise around the fixation point in front of a gray background. The wedge was 22.5° wide and extended 15° in the periphery. It started at the right horizontal meridian and circled around the fixation point 12 times at a speed of 11.25° in polar angle/volume (2 s), each cycle taking 64 s. During eccentricity mapping, which induces phase-encoded neural activity that allows to match radii of eccentricity on the cortical surface to the corresponding visual angles from the center of gaze, participants viewed a slowly expanding black and white checkerboard ring in front of a gray background, starting with a radius of 1° and increasing up to a radius of 15°. The ring’s contrast reversed at a flickering rate of 4 Hz and the expansion cycle was repeated 7 times, each cycle taking 64 s. The participants’ only task was to maintain fixation on the central fixation point in both runs.

### Neuromaging Data Analysis

FreeSurfer’s surface-based methods were used for individual cortical surface reconstruction from the T1-weighted image of each participant (http://surfer.nmr.mgh.harvard.edu/s/wiki/RecommendedReconstruction; Dale et al. 1999; Fischl, Sereno, and Dale 1999). FSFAST was used for the processing of the functional data and included slice time correction, motion correction and co-registration to the T1-weighted anatomical image. A Fourier transform was applied to each voxel’s fMRI time series to extract the amplitude and phase at stimulation frequency. This phase angle of a voxel is linked to the polar angle or eccentricity that is represented at the voxel’s cortical location. The phase angles were then mapped to different colors, with each color representing a response phase whose intensity is an F-ratio resulting from a division of the squared amplitude of the response at stimulus frequency by the average squared amplitudes at all other frequencies (except for higher harmonics of the stimulus frequency and low-frequency signals). The resulting retinotopic maps were then displayed on the inflated cortical surface maps of the T1-weighted anatomical images. Polar angle phase maps were used to estimate the boundaries between areas V1–V3, whereas the eccentricity phase maps were used to estimate the eccentricity out to 7.2°, which is a typical eccentricity distance used to estimate central visual cortex surface size (Schwarzkopf et al. 2011; Schwarzkopf and Rees 2013; Song, Schwarzkopf, and Rees 2013; Bergmann et al. 2014; Song et al. 2015). Further, it has been shown that brain–behavior relationships are robust across variance in a range of central V1 eccentricities out to 8.5° (Song, Schwarzkopf, and Rees 2013; Song et al. 2015). The estimation of the boundaries was done manually for each individual by 2 independent raters, the first and the second author of this paper, and was done before the behavioral data were analyzed to prevent potential biases in their judgments. FreeSurfer’s Anatomical ROI analysis tool was then used to determine the early visual areas’ anatomical properties. Inter-rater reliability of early visual cortex’ surface size and cortical thickness showed a very high agreement between the 2 raters’ judgments of V1 and V2 (V1 surface: r = 0.898, P < 0.001; V1 thickness: r = 0.971, P < 0.001, V2 surface: r = 0.914, P < 0.001; V2 thickness: r = 0.968, P < 0.001). In V3, the agreement between the 2 raters was slightly lower for surface (r = 0.600, P < 0.001) but stayed very high for thickness (r = 0.952, P < 0.001). For further analyses of individual structure–function relationships, the average of the 2 ratings of the anatomical measures for each participant was used. Additionally, as a control, each individual’s overall brain surface size and cortical thickness was determined, and the Desikan–Killiany atlas implemented in FreeSurfer was used to estimate the anatomical measures of other brain areas.

All anatomical measures obtained from the retinotopic mapping procedure were further checked for normal distribution using Shapiro–Wilk normality test. For the V3 surface area values, we detected a violation of the normal distribution assumption (W(33) = .900, P = .005). Further exploration using QQ-Plot and Boxplot analysis revealed an extreme value of participant S13. We therefore reexamined the correlations between behavior and V3 surface to evaluate the impact of this extreme value on the relationships, which proved to be minor (see Supplementary Results).

Anatomical parcellation of V1. We also used FreeSurfer’s surface-based probabilistic method to estimate V1 boundaries in each participant individually (Hinds et al. 2008). The surface-based prediction method is based on anatomical landmarks and, unlike the fMRI retinotopic mapping approach, aims to measure the area of V1 in its entirety. A highly conservative threshold of 0.8 was used. This corresponds to the probability that a vertex lies within the actual boundary of V1. Again, the anatomical measures of volume, surface and thickness of V1 were determined and related to the behavioral measures.

Volume analysis of the whole cortex. To examine the relationship between the cortical architecture and mean subjective vividness...
and to back-up our results obtained for orientation-specific imagery precision, we ran whole-brain volume analyses using the surface-based analysis approach implemented in FreeSurfer’s Qdec application. For this purpose, the surface, thickness and curvature data of each participant were first smoothed using a full width at half maximum Gaussian filter of 10 mm. As a next step, each individual surface was transformed into a spherical representation and then nonrigidly adjusted to a common-space spherical surface (fsaverage), using the folding patterns as anatomical landmarks to optimize the alignment across participants (Fischl, Sereno, Tootell, et al. 1999). Following this step, volume data of each individual were applied to the common group space, which allowed for comparisons across participants at homologous points of the brain. A general linear model fit to explain the behavioral data was computed vertex-wise using an uncorrected threshold of \( P < 0.01 \). Correction for multiple comparisons was done using a precached Monte Carlo Null-Z simulation with 10,000 iterations and a cluster-wise threshold of \( P < 0.05 \) (Hagler et al. 2006).

**Results**

**Imagery Strength**

To test the strength of imagery, participants were first instructed to imagine 1 of 2 orthogonal Gabor gratings for 7 s. After this interval, they were asked to provide a subjective rating of vividness of the mental image on a scale from 1 to 4 before a rivalry display was flashed for 0.75 s. When the stimulus disappeared, individuals reported the rivalry percept (see Fig. 1A and Methods). Imagery strength was defined as the percentage of trials in which the mental image of the stimulus matched its perceptual dominance on the succeeding rivalry presentation (% primed, see Methods).

Both the behavioral and the anatomical measures showed very high reliability. Behavioral retests after a mean interval of 2 weeks revealed high stability of the imagery strength measure \( r = 0.860, P < 0.001 \), see Supplementary Fig. 1A. Likewise, the blind inter-rater reliability of the neuroanatomical areas was very high (all \( P < 0.001 \), see Methods) and revealed substantial inter-subject variability (also see Supplementary Fig. 2). In addition, a higher subjective vividness rating on each trial predicted a higher percentage of priming (Wilk’s lambda: \( \lambda = 0.511, F_{34, 258} = 7.669, P < 0.001 \), “partial” explained variance \( \epsilon = 0.489 \), see Fig. 2A), replicating previous work on the metacognition of visual imagery (Pearson et al. 2011; Rademaker and Pearson 2012). Interestingly, the mean of the subjective vividness ratings for each participant did not correlate significantly with objective imagery strength \( r = 0.229, P = 0.215 \). However, it too showed good retest reliability \( r = 0.677, P < 0.001 \), see Supplementary Fig. 1B and C).

We found that the retinotopically estimated surface area of V1 was significantly negatively related to imagery strength \( r = -0.376, P = 0.031 \), see Fig. 2B), indicating that individuals with a smaller V1 surface area tended to have stronger imagery. While the correlation with V2 surface area approached significance \( r = -0.340, P = 0.053 \), we observed no relationship with the surface area of V3 \( r = -0.085, P = 0.638 \). There was also no significant relationship between imagery strength and the total surface size of the brain \( r = 0.012, P = 0.948 \) and with the surface size of other brain areas as determined by the gyral-based Desikan–Killiany atlas implemented in FreeSurfer (all \( P > 0.123 \)). Furthermore, we found no relationship with cortical thickness of early visual cortex areas (all \( P > 0.345 \)). Further analyses regarding potential hemispheric asymmetries in the brain–behavior relationship with V1 of the 2 hemispheres (see Supplementary Results). The relationship between V2 surface area and imagery strength approached significance, whereas there was no relationship with V3.

**Figure 2.** Imagery strength and its relationships to subjective vividness and early visual cortex surface size. (A) A higher trial-to-trial vividness rating is associated with a higher probability of subsequent rivalry priming. Before the binocular rivalry presentation in every trial, participants rated the vividness of their mental imagery (1–4, 1 representing the lowest and 4 the highest vividness). Error bars represent ±SEM. (B) Imagery strength is predicted by retinotopically defined V1 surface extensions. The scatterplots show the relationship between early visual cortex surface area and imagery strength as measured by the functional impact of imagery on perception (% primed) in the binocular rivalry task. Each data point indicates the value of one participant. Bivariate correlation coefficients and significance levels are included. Individuals with a smaller V1 tended to have stronger imagery, there were no significant differences for the relationship with imagery strength between V1 of the 2 hemispheres (see Supplementary Results). The relationship between V2 surface area and imagery strength approached significance, whereas there was no relationship with V3.
relationship showed higher correlations for V1 surface size of the left compared with the right hemisphere; however, this difference was nonsignificant (see Supplementary Results) and functional interpretations should be made with caution, as differences in the correlations might also be due to differences in the size of the foveal convergence between left and right hemisphere (Dougherty et al. 2003). In a next step, we looked at the relationship with imagery strength when V1 was estimated anatomically (αV1; Hinds et al. 2008). Unlike the functional retinotopic mapping of V1, the anatomical mapping aims to measure V1 in its entirety. In line with previous studies on the relationship between V1 size and behavior (Schwarzkopf and Rees 2013; Bergmann et al. 2014), we found no relationship between imagery strength and overall αV1 surface size (r = –0.048, P = 0.791). There was also no relationship with overall αV1 cortical thickness (r = 0.016, P = 0.929).

Imagery Precision

The finding that smaller V1 surface area predicted stronger imagery was somewhat surprising, given that larger spatial representations in this cortical area are related to higher visual acuity intra- (Duncan and Boynton 2003) and inter-individually (Song, Schwarzkopf, and Rees 2013) and considering that previous work has proposed that V1 might be essential for enabling the high-resolution visualization of a detailed stimulus (Kosslyn et al. 2001; Kosslyn and Thompson 2003). We therefore examined the relationship between V1 and a different measure of mental imagery—imagery precision, which may be more analogous to perceptual acuity. As visual imagery is known to be specific to a spatial orientation and location (Pearson et al. 2008), we sought to compare these characteristics of imagery with individual V1 surface size.

In an orientation-specific variation of the imagery task (see Fig. 18 and Methods), participants were again instructed to imagine either a horizontal or a vertical Gabor pattern as in the standard task. However, the patterns in the subsequent rivalry presentation now deviated from the imagined stimuli’s spatial orientation by a rotation angle of either ±45° or ±22.5° or matched the spatial orientation of the imagined stimuli (0°). Orientation-specific imagery precision was then defined as the difference in priming rates (Δ%primed) between the 0° and the ±22.5° condition, or as the difference in priming rates between the ±0° and ±45° condition, respectively.

In the location-specific imagery precision task version (see Fig. 1C and Methods), the location of imagery in the visual field either matched the location of the succeeding rivalry presentation or not. Location-specific imagery precision was defined as the difference in priming (Δ %primed) between the conditions when imagery and binocular rivalry occurred at the same location (primed location condition) and when they occurred at different locations (unprimed location condition).

Interestingly, there was no strong linear relationship between imagery precision and imagery strength, that is, precision only increased weakly and nonsignificantly with higher imagery strength (r = 0.210, P = 0.241 for the ±22.5°/0° priming difference; r = 0.333, P = 0.058 for the ±45°/0° priming difference; and r = 0.152, P = 0.400 for the priming difference in the location-specificity condition). Instead, we observed a significant nonlinear quadratic relationship, with intermediate imagers showing the highest degree of imagery precision compared with strong and weak imagers. This effect was particularly strong for the relationship between imagery strength and imagery precision as measured by the ±22.5°/0° priming difference (R = 0.580, F2,30 = 7.591, P = 0.002, see Fig. 3A) but also present in the relationship between imagery strength and imagery precision as measured by the ±45°/0° priming difference (R = 0.529, F2,30 = 5.829, P = 0.007, see Fig. 3B).

When the sample was split into 3 groups based on the degree of imagery strength, the differences in imagery precision were well reflected in the groups’ orientation-tuning functions (Fig. 3C). Orientation tuning was most evident in intermediate imagers. In strong imagers, it was only the larger deviations (±45°) that showed stronger reductions in priming. In contrast, there was no evident imagery orientation tuning in the weak imagery group (see Supplementary Results for a statistical analysis).

In the location-specificity condition, the data also showed a quadratic relationship between imagery strength and precision (R = 0.437, F2,30 = 3.539, P = 0.042), with intermediate imagers tending to show a slightly higher precision compared with strong and weak imagers (see Fig. 3D,E and Supplementary Results).

We then looked at the relationship between imagery precision and early visual cortex anatomy. Of course, a valid estimate of imagery precision is only feasible when individuals are able to produce at least a weak form of visual imagery. We therefore chose a very conservative cutoff for the perceptual bias of >55% primed at the 0° condition (orientation-specificity task) and at the "primed location" condition (location-specificity task) to exclude those participants who showed no signs of imagery strength. In addition, we conducted further analyses with other cutoff points to ensure that statistical significance was stable and largely independent of the precise location of the cutoff point (see Supplementary Fig. 3).

In contrast to imagery strength, we found a positive association between imagery precision and V1 anatomy for both the orientation-specificity and the location-specificity measures. For orientation precision, the relationship with V1 was particularly strong when precision was estimated from priming of smaller spatial orientation differences (Δ %primed between 0° and ±22.5°). Both V1 surface area and V1 cortical thickness predicted imagery precision (V1 surface area: r = 0.409, P = 0.031, see Fig. 4A; V1 thickness: r = 0.418, P = 0.027, N = 28; see Supplementary Fig. 4), whereas there was no correlation with V2 and V3 surface area and thickness (all P > 0.113, see Supplementary Fig. 4). There was also no indication of a correlation with the total brain surface or cortical thickness (r = −0.049, P = 0.806 for total brain surface size, and r = −0.048, P = 0.809 for total brain cortical thickness). A multiple linear regression including V1–V3, total brain surface size and cortical thickness, age and gender as independent variables and imagery precision (Δ %primed between 0° and ±22.5°) as dependent variables further confirmed the unique contributions of V1 surface size and thickness to the prediction of imagery precision. Only V1 surface area was highly significant (beta = 0.661, t27 = 3.075, P = 0.007), and V1 thickness was significant by tendency (beta = 0.463, t27 = 1.978, P = 0.064). All other factors remained nonsignificant (all P > 0.273). To further back up these results, we additionally ran a whole-brain analysis using FreeSurfer’s Qdec application to find out whether any other brain regions might display a significant relationship with orientation-specific imagery precision (see Methods). Using volume, which is the combined measure of vertex area and thickness, we found several significant clusters, for example, in the precuneus and in the occipital pole where central V1 is located. However, only the cluster in the occipital pole of the left hemisphere survived the correction for multiple comparisons, supporting the notion that central V1 is an important stage for individual differences in orientation-specific imagery precision (see Supplementary Fig. 5 and Table 1).
For the priming differences at larger spatial orientation deviations (Δ%primed between 0° and ±45°), the prediction power of V1 anatomy was weaker, with only the composite of V1 surface and thickness, V1 volume, showing a positive relation (V1 volume: \( r = 0.391, P = 0.040 \)), whereas the single measures of V1 surface size (\( r = 0.305, P = 0.115 \)) and cortical thickness (\( r = 0.134, P = 0.495 \)) did not reach statistical significance. There was also a relationship neither with V2 or V3 anatomy (all \( P > 0.182 \)) nor with total brain surface size or cortical thickness (\( r = 0.091, P = 0.645 \) for total brain surface size, and \( r = .245, P = .188 \) for total brain cortical thickness). No clusters survived the correction for multiple comparisons when computing a whole-brain volume analysis with Freesurfer’s Qdec (data not shown).

Similarly, for location-specific imagery precision, the relationship with V1 surface area was significantly positive (\( r = 0.537, P = 0.010, N = 22 \), see Fig. 4B). The relationship with V2 surface area was also significant (\( r = 0.499, P = 0.018 \), see Supplementary Fig. 6), whereas there was no statistically significant relationship with V3 surface (\( r = 0.317, P = 0.150 \)). Furthermore, there was no significant relationship with V1–V3 cortical thickness, but in V2 it approached significance (\( r = .409, P = 0.059 \)). Likewise, no correlation was found with total brain surface size (\( r = -0.181, P = 0.420 \)) and total brain cortical thickness (\( r = 0.014, P = 0.952 \)). In a multiple linear regression analysis including V1, V2, V3, whole-brain surface area, age and gender as independent variables and location-specific imagery precision as dependent variables, V1 surface area was significant by tendency (beta = 0.477, \( t_{21} = 1.91 \),

Figure 3. The relationship between imagery strength and precision in the orientation-specific (left column) and location-specific imagery precision task (right column). (A,B,D) The nonlinear relationship between imagery strength and imagery precision estimated by the priming difference between the 0° and the pooled ±22.5° spatial orientation conditions (A), by the priming difference between the 0° and the pooled ±45° spatial orientation conditions (B), and by the priming difference between primed and unprimed location in the location-specific imagery condition task (D). Each data point reflects the value of one participant, whereas the curved gray line represents the quadratic regression estimate. The dotted vertical lines show the three-way data split based on imagery strength (see C,E). Individuals with intermediate imagery strength tended to have higher imagery precision compared with weak and strong imagers. (C,E) Average imagery-induced priming rates of strong (gray), intermediate (magenta) and weak imagers (blue) as a function of the difference between imagined and rivalry orientation angle (C) and as a function of primed versus unprimed location (E). The data points show each group’s average imagery strength at each orientation angle or location; error bars show ±SEM.
mentioned earlier, we found that across the group, a higher trial-by-trial rating of subjective vividness was linked to a higher likelihood of imagery-induced priming (see Fig. 2A); in contrast, an individual’s average subjective vividness across all trials was unrelated to that same individual’s mean priming rate ($r = 0.229, P = 0.215$; see Supplementary Fig. 1C), which is in line with previous research suggesting that subjective and objective conscious experience may be decorrelated and involve different brain areas (Lau and Passingham 2006; Schwiedrzik et al. 2011), with subjective experience relying on neural activity in left prefrontal cortex.

As imagery vividness also showed good retest reliability ($r = 0.777, P < 0.001$, see Supplementary Fig. 1B), we ran an exploratory whole-brain volume analysis using FreeSurfer’s Qdec application and found a strong positive relationship between an individual’s mean subjective imagery vividness and the volume of the left prefrontal cortex, in an area overlapping superior frontal, lateral orbitofrontal and rostral middle frontal cortices (see Methods, Supplementary Fig. 7 and Table 2). In contrast, mean vividness ratings did not show any strong relationship with early visual cortex surface size (all $P > 0.705$, see Supplementary Fig. 7).

We repeated the Qdec analysis using the subjective vividness data from the orientation-specific and location-specific imagery precision tasks as an additional test for the validity of the Qdec results. Replicating our previous results, we found a significant cluster in the left prefrontal cortex with both data sets, largely convergent with the cluster reported earlier. Here, again, the cluster survived the correction for multiple comparisons using Monte Carlo Null-Z Simulation with a cluster-wise probability threshold of $P < 0.05$ (data not shown).

**Discussion**

Prior exposure to 1 of 2 binocular rivalry patterns can induce either its initial perceptual dominance or its suppression, effects that are known as “flash facilitation” versus “flash suppression” (Wolfé 1984; Brascamp et al. 2007; Pearson et al. 2008). Likewise, prior exposure to a weak perceptual stimulus can reduce subsequent detection thresholds, whereas exposure to stronger stimuli leads to a threshold elevation after effect (Tanaka and Sagi 1998). Evidence suggests a single continuous mechanism, which depends on the visual “energy” of the prior stimulus, that is, its “sensory strength.” According to this, facilitation is more likely if the preceding stimulus is short and/or low contrast, whereas suppression occurs when a prior stimulus is high contrast and/or is shown for a long duration (Brascamp et al. 2007; Pearson and Brascamp 2008; Pearson et al. 2008). Hence, the facilitative effect of a prior stimulus first increases as a function of contrast or presentation duration until it reaches a tipping point at which point the effect reverses and leads to reduced facilitation and increased suppression.

Here, we used a paradigm that exploited this phenomenon to measure the strength of mental imagery. In comparison to perceptual stimuli, the “overall energy” of imagined stimuli is rather weak, and even longer periods of imagery generation lead to stronger priming, not suppression (Pearson et al. 2008). In our study, individual differences in imagery generation for around 7 s gave large but reliable variance (44–100%) in priming rates, suggesting that no individual had a mental image capable of strong suppression, which would present as very low priming rates, for example, a priming rate of ~20%.

The trial-by-trial subjective vividness ratings preceding the binocular rivalry display also lend further support to our assumption that priming rate is a valid measure of imagery strength: A
higher rating of subjective vividness on any individual trial was followed by a higher probability of imagery-induced priming (as indicated by the dominance reports), replicating previous results linking sensory imagery strength estimated by priming to subjective vividness ratings (Pearson et al. 2011; Rademaker and Pearson 2012; see Figure 2A). In addition, prior work has demonstrated that the effects of imagery on subsequent rivalry can be clearly dissociated from visual attention by both the differential attenuation from background luminance and different time courses (Pearson et al. 2008). We therefore conclude that greater imagery-induced neural activation would spread less, resulting in sharper imagery orientation- or location-tuning curves. At very large orientation or location differences, however, the distance between the involved neurons might be too large and the level of interconnectivity too low even in smaller V1 cortices to reveal any variance for the different V1 cortical surface sizes. This could explain why the relationship between orientation precision and V1 drops off at very large differences in orientation space (0° and ±45°).

It is more difficult to disentangle the factors that might be responsible for the relationship between increased imagery strength and reduced V1 surface area. The nonlinear relationship between imagery strength and precision and the opposite signs of their correlations with the anatomy of V1 suggest distinct functional mechanisms that underlie these 2 imagery properties. Naturally, imagery precision is low when individuals show little or no sign of imagery strength. However, among those individuals who are capable of generating mental images, there seems to be a tradeoff between the 2 properties of imagery. Despite the fact that imagery precision and perceptual precision are clearly different characteristics, it seems intuitive that the functional mechanisms that underlie imagery precision are largely overlapping with those of perceptual precision (Duncan and Boynton 2003; Song, Schwarzkopf, and Rees 2013; Song et al. 2015), thereby presenting strong support for the pictorial theories of imagery (Kosslyn et al. 2001; Pearson and Kosslyn 2015). In contrast, imagery strength might be more distinct and strongly dependent on other mechanisms, such as the top-down input of higher-level areas on V1.

A reciprocal relationship has been observed between V1 and the prefrontal areas of the brain, with the volume of these 2 structures anticorrelating (Song et al. 2011). This finding ties in well with interdependencies in the development of these areas over the course of evolution (Pearce et al. 2013). As opposed to reproducing sensory detail, prefrontal areas have an important role in shaping higher cognitive functions (Christoff and Gabrieli 2000; Fletcher and Henson 2001; Ramnani and Owen 2004; Burgess et al. 2007), one of which is the manipulation of visual imagery (Ishai et al. 2000; Mechelli et al. 2004; Schacter et al. 2007; de Araujo et al. 2012; Schlegel et al. 2013; Dentico et al. 2014). Accordingly, studies indicate that content-sensitive top-down connections from the prefrontal cortex regulate category-selective activation in visual areas (Ishai et al. 2000; Mechelli et al. 2004).

In the light of the prefrontally guided interaction during imagery, it is likely that the top-down control from the prefrontal cortex over V1 activity gains even more dominance when V1 is smaller and the bottom-up sensory input is less precise. As a consequence, a stronger top-down influence of high-level areas in manipulating sensory activity might lead to stronger imagery and greater perceptual priming. However, as high-level areas are unlikely to produce or trigger precise sensory details, these images would lack precision. This effect might be further enhanced by the reciprocity of V1 and prefrontal cortex.

There is also another aspect in our location imagery data that lend further support for the speculation that the effect of V1 imagery-induced neural activity spreads to neurons that encode similar spatial orientations and nearby locations in the visual field. This would result in lower imagery precision, that is, a reduced priming difference between 2 orientations or between 2 locations. In contrast, with reduced lateral interconnectivity and a higher number of neurons representing the visual field in larger V1 cortices, the distance between neurons that encode different orientations or locations might be increased, whereas the level of interconnectivity would be decreased. As a consequence, imagery-induced neural activation would spread less, resulting in sharper imagery orientation- or location-tuning curves. At very large orientation or location differences, however, the distance between the involved neurons might be too large and the level of interconnectivity too low even in smaller V1 cortices to reveal any variance for the different V1 cortical surface sizes. This could explain why the relationship between orientation precision and V1 drops off at very large differences in orientation space (0° and ±45°).
surface size on imagery strength is indirect. Due to cortical magnification, there is a larger proportional amount of V1 dedicated to processing input from the center of the visual field. Accordingly, if the negative relationship between imagery strength and V1 size was intrinsically caused by neural processes restricted to V1, we should assume that imagery in the periphery should be stronger compared with central imagery, as—within an individual—less cortical space is allocated to these peripheral parts of the visual field. However, the opposite was the case: Across all participants, peripheral imagery strength (that is, priming at the peripheral locations in the location-specific imagery precision task) was significantly worse than imagery at the center ($t_{24} = 3.773, P = 0.001$). Likewise, the subjective vividness ratings were significantly lower for peripheral imagery compared with central imagery ($t_{24} = 3.525, P = 0.001$). Our results also point in the direction of the previously reported reciprocity in the relationships between prefrontal cortex and V1 (Song et al. 2011); however, the observed relationships here were weaker and outlier affected (see Supplementary Fig. 8).

Another thought-provoking issue requiring further investigation is the potential feature- and category-specificity of the relationship between imagery and visual cortex anatomy. In our experiments, the crucial stimulus features were spatial orientation and location, which are processed in V1. Such low-level features can also be used to decode highly complex mental images like complex paintings from early visual cortex activity, despite the fact that the overall level of activity in these areas is low (Naselaris et al. 2014); in light of these findings, we deem it unlikely that the relationship between V1 and imagery in our experiments only arose due to the simplicity of the stimuli used. However, with mental stimuli becoming more complex mental stimuli (e.g., faces), it should also be expected that the restrictions on mental acuity may not only be set by limitations in early visual cortex processing (V1 in particular, but also succumb to those of feature- or category-specific mid- and higher-level visual areas further upstream.

The relationship between a smaller V1 and increased imagery strength might also be important in the context of several clinical observations. Uncontrollable imagery is known to be increased in psychopathologies such as schizophrenia (Rcman and Landis 1945; Chapman 1967; Oertel et al. 2009; Benson and Park 2013; Matthews et al. 2013) and PTSD (Bryant and Harvey 1996; Shin et al. 1997; Morina et al. 2012), which have also been shown to involve over proportional decreases in V1 volume and neuron number (Dorph-Petersen et al. 2007; Chao et al. 2012). In addition, visual impairments or visual deterioration have been linked to a high susceptibility of developing schizophrenia, whereas this risk is extremely low for individuals with “supernormal” (perfect) vision or the congenitally blind (Landgraf and Osterheider 2013). The underlying mechanism for this relationship has so far remained elusive. Maybe a smaller V1, which seems linked to lower perceptual (Song, Schwarzkopf and Rees 2013; Song et al. 2015) and imagery acuity/precision, but stronger imagery in healthy individuals, could be one of the biological risk factors for the development of schizophrenia.

Taken together, our findings suggest the need to distinguish between at least 3 different properties of mental imagery—imagery strength, precision and subjective vividness. We report the first evidence that the layout of primary visual cortex might be crucial for the strength and precision of human imagination—providing the first mechanistic solution to a long-standing scientific question.

Supplementary Material

Supplementary material can be found at: http://www.cercor.oxfordjournals.org/.

Authors’ Contributions

J.B. and J.P. conceived the study; all authors designed the experiments; J.B. and J.P. programmed the behavioral experiments; J.B. and E.G. conducted the behavioral and fMRI experiments and analyzed the data; J.B. and J.P. wrote the manuscript; and all the authors edited the manuscript.

Funding

This work was supported by Australian NHMRC project grants APP1024800 and APP1046198, ARC grant DP140101560 and a Career Development Fellowship APP1049596 held by J.P. J.B. was supported by an International Postgraduate Research Scholarship and a Brain Sciences UNSW PhD Top Up Scholarship.

Notes

Conflict of Interest: None declared.

References


Landgraf S, Osterheider M. 2013. “To see or not to see: that is the question.” The “Protection-Against-Schizophrenia” (PaSZ) model: evidence from congenital blindness and visuo-cognitive aberrations. Front Psychol. 4:352.


Panizzon MS, Fennema-Notestine C, Eyer LT, Jernigan TL, Prom-Wormley E, Neale M, Jacobson K, Lyons MJ, Grant MD,


