

The limited contribution of early visual cortex in visual working memory for surface roughness
(Abbreviated title: fMRI decoding surface roughness in VWM)

Munendo Fujimichi^{1,2}, Hiroki Yamamoto¹, and Jun Saiki¹

¹Graduate School of Human and Environmental Studies, Kyoto University, Yoshida-Nihonmatsucho, Sakyo-ku,
Kyoto 606-8501, Japan

²Japan Society for the Promotion of Science, Tokyo, Japan

Corresponding author email address: fujimichi@cv.jinkan.kyoto-u.ac.jp

Abstract

Are visual representations in the human early visual cortex necessary for visual working memory (VWM)? Previous studies suggest that VWM is underpinned by distributed representations across several brain regions, including the early visual cortex. Notably, in these studies, participants had to memorize images under consistent visual conditions. However, in our daily lives, we must retain the essential visual properties of objects despite changes in illumination or viewpoint. The role of brain regions—particularly the early visual cortices—in these situations remains unclear. The present study investigated whether the early visual cortex was essential for achieving stable VWM. Focusing on VWM for object surface properties, we conducted fMRI experiments while male and female participants performed a delayed roughness discrimination task in which sample and probe spheres were presented under varying illumination. By applying multi-voxel pattern analysis to brain activity in regions of interest, we found that the ventral visual cortex and intraparietal sulcus were involved in roughness VWM under changing illumination conditions. In contrast, VWM was not supported as robustly by the early visual cortex. These findings show that visual representations in the early visual cortex alone are insufficient for the robust roughness VWM representation required during changes in illumination.

Keywords: roughness, visual working memory, fMRI, early visual cortex

Declarations

Funding: This work was supported by JSPS KAKENHI Grant Number JP16H01672 and JP18J22547

Conflict of interest statement: The authors have no conflicts of interest, financial or otherwise

Availability of data and material: Not applicable

Code availability: Not applicable

Authors' contributions: M.F., H.Y., and J.S. designed the experiment and M.F. carried out the experiment. M.F. analyzed fMRI data with H.Y. and wrote the manuscript with support from H.Y. and J.S.

Acknowledgements

We thank Hiroyuki Tsuda and Hiroyuki Yamashiro for their coding support. This study was conducted using the MRI scanner and related facilities of Kokoro Research Center, Kyoto University. This work was supported by JSPS KAKENHI Grant Number JP16H01672 and JP18J22547.

37 **Note**

38 The final publication is available at <https://doi.org/10.1007/s00221>
39 [-020-05881-0](https://doi.org/10.1007/s00221)

Introduction

Visual working memory (VWM) plays an important role in cognition. Functional brain imaging studies employing univariate analysis have revealed that brain activation in the intraparietal sulcus (IPS) reflected maintenance of spatial information in visual short-term memory (VSTM; Harrison et al. 2010; Todd & Marois 2004, 2005; Ungerleider et al. 1998; Xu & Chun 2006). Studies using multi-voxel pattern analysis (MVPA) have shown that multiple brain regions contribute to VWM. In particular, the early visual cortex may play a role in orientation (Harrison & Tong 2009), contrast (Xing et al. 2013), and color pattern elements (Christophel et al. 2012) of VWM. In these studies, participants were asked to memorize an image under set visual conditions; however, our daily lives necessitate retention of the essential properties of objects against changes despite changes in visual conditions, such as illumination and viewpoint. The role of specific brain regions, especially the early visual cortices, under changing visual conditions remains unclear.

To address this issue, we focused on object surface properties, such as glossiness and roughness, that humans can perceive despite changes in the illumination of visual images (Fleming et al. 2003; for a review, see Komatsu & Goda 2018; Tsuda & Saiki 2018). The ventral visual pathway has a vital role in material perception, whereas the early visual cortex has a limited role. The early visual cortices can compute low-level image features, such as luminance skewness, to estimate object surface qualities (Cant et al. 2009; Motoyoshi et al. 2007; Sun et al. 2015; Wada et al. 2014). Although computation alone is insufficient to achieve material perception, this process may be a necessary component (Hiramatsu et al. 2011). Regarding VWM for illumination-invariant material properties, illumination-invariant ventral representation may need to be accompanied by representation in the early visual cortex to retain stability (Christophel et al. 2017). Unlike material perception with online visual input, VWM without visual input may require interaction between the ventral visual pathway and early visual cortex to maintain a stable representation. Alternatively, only illumination-invariant surface properties may be maintained in the ventral visual pathway, and representation in the early visual cortex may be discarded to avoid confusion due to changes in visual images between memorized and test stimuli.

Numerous studies without visual interference during maintenance of visual features have reported memory representations in the early visual cortex. A recent study showed that visual distractors during VWM maintenance disrupted representations in the early visual cortex (Bettencourt & Xu 2016). However, this finding alone is insufficient to conclude that the early visual cortex is not involved in retaining stable VWM representations. For instance, there may be apparent reduction in decoding performance due to increased neuronal response to task-irrelevant visual distractors. Thus, it remains unclear whether the early visual cortex is involved in maintaining VWM that is robust against changes in illumination and viewpoint. It is important to examine whether anticipation of visual image changes in a test stimulus are sufficient to discard the representation within the early

visual cortex to elucidate the nature of distributed VWM representations.

To address these questions, we focused on the surface property of roughness. We conducted an experiment using a delayed roughness discrimination task in which sample and probe objects were rendered in different illuminations. To cope with the different illumination, we hypothesized that the visual system would establish an illumination-invariant roughness representation for VWM in lieu of an image-based representation in the early visual cortex. We applied MVPA to brain activity specific to VWM from the ventral visual pathway, early visual cortex, and IPS during a delay period. If illumination-invariant representation of roughness alone can be selectively maintained, we expected that the ventral visual cortex and IPS, but not early visual cortex, would be involved in roughness VWM. Alternatively, if roughness VWM requires a distributed representation of image-based and illumination-invariant information, we expected that the early visual cortex in addition to the ventral visual cortex and IPS would be involved.

Materials and Methods

Participants

Twenty-four undergraduate and graduate students with no history of psychiatric or neurological diseases participated in a visual working memory experiment, a functional localization scan, and a retinotopy scan (8 women; mean age = 22.04 years). All participants had normal color vision, confirmed by an Ishihara test, as well as normal or corrected-to-normal vision acuity with non-magnetic glasses. Participants viewed stimuli monocularly with the left eye while vision in the right eye was obstructed by inserting a black mask in the right lens of the glasses. Monocular viewing was employed to prevent double vision caused by a head coil shielding a major portion of the binocular visual field. After receiving a detailed description of the study, all participants provided written informed consent. The experiment was conducted in accordance with the guidelines of the Ethical Committee of the Graduate School of Human and Environmental Studies at Kyoto University. All participants received a book coupon (5000 Japanese yen) for their participation.

Visual working memory experiment

Stimuli

Computer-generated three-dimensional spheres, similar to those used in a previous behavioral study (Tsuda & Saiki 2018), were used as sample and probe stimuli. The surface appearance of the sample spheres was set at one of two different levels of roughness with jitters (Fig. 1). The sharpness parameter (α) for each level of roughness was 0.016667 (± 0.001667) and 0.080000 (± 0.001667), respectively while glossiness (specular reflectance parameter; ρ_s) was maintained at 0.040617. The α and ρ_s parameters were controlled using RADIANCE (Ward 1994). For probe spheres, three different levels of roughness were rendered in a similar way; the levels of roughness were 0.003333

(± 0.001667), 0.040000 (± 0.001667), and 0.098333 (± 0.001667), respectively, with glossiness maintained at 0.040617. All spheres were rendered in two different real-world illuminations derived from Debevec (1998): ‘Kitchen’ and ‘Overcast Breezeway’.

Procedure

In the VWM experiment, participants performed a delayed roughness discrimination task. A sample of the trial design is shown in Figure 2. Each trial began with two samples ($10.3^\circ \times 10.3^\circ$) presented sequentially for 500 ms each. A black cross ($1.2^\circ \times 1.2^\circ$) was displayed for 500 ms between the two samples. The sample stimuli consisted of one rough sphere and one smooth sphere. Following the two samples, a numerical cue (1 or 2; $1.9^\circ \times 3.4^\circ$) was displayed, indicating which sample should be memorized. After an 11-s delay, a probe sphere was presented and participants were asked to indicate within 2 s whether the probe sphere had higher or lower roughness than the memorized sphere. The illumination contexts of the sample spheres and probe sphere were always different to prevent memorization of the visual image instead of the material properties. Between trials, a black cross was presented for 0–10 s. Each run (144 s) consisted of four trials, plus an initial 14-s black cross presentation that was excluded from analysis. Within each run, the order of rough and smooth samples and the samples sphere to be memorized were counterbalanced. All participants performed 10 runs.

Functional magnetic resonance imaging (fMRI) data acquisition

This study was conducted using the MRI scanner and related facilities at the Kokoro Research Center, Kyoto University. Data were collected on a 3 T Siemens scanner (3.0 T MAGNETOM Verio). Functional data were obtained with T2*-weighted imaging (repetition time (TR) = 2000 ms; field of view = 224×224 mm; voxel size = $3.5 \times 3.5 \times 3$ mm; slice number = 34 slices). T1-weighted anatomical images were also acquired for each participant.

fMRI data analysis

Brain surface reconstruction, preprocessing, and multi-voxel pattern analysis

All fMRI data were analyzed with the supercomputer at the Academic Center for Computing and Media Studies, Kyoto University. We reconstructed the cortical surfaces for each participant from the T1-weighted image volumes. We used the recon-all script in FreeSurfer-6.0.0 (surfer.nmr.mgh.harvard.edu/) for gray and white matter segmentation, surface smoothing, and inflation. Preprocessing of fMRI data was performed using Analysis of Functional NeuroImages (AFNI; Cox 1996) and Surface Mapping with AFNI (SUMA; Saad et al. 2004).

Accuracy of MVPA for roughness was determined using the following six steps: (1) The fMRI data were corrected for slice timing and head motion with realignment to the first volume. (2) The

functional images were registered to the cortical surfaces for each participant. (3) Voxel-based preprocessing, linear trend removal, and percent signal change conversion were applied to the fMRI data. (4) A deconvolution technique (Dale 1999) was applied to the fMRI data, estimating the time course for each rough or smooth memory condition. (5) The deconvoluted signals were sampled within each region of interest (ROI) as described below. (6) MVPA was performed for the signals at each TR using the libsvm-3.20 script (Chang & Lin 2011) in MATLAB 9.1.0 (MathWorks). The leave-one-out cross-validation procedure was used to estimate decoding performance. Learning runs were trained for each TR to create a decoder that could predict the memorized roughness condition based on brain activity patterns. The decoder was then used to predict whether the brain activity pattern in each test run belonged to memorization of the rough or smooth sphere. This procedure was repeated until all runs had been tested, and the resultant roughness classification performance was defined as the MVPA accuracy.

Regions of interest

Following the VWM experiment, we performed two types of scans on the left eye of each participant to localize the visual area. The first scan (functional localization scan) localized the higher-order visual areas using natural images in one of the following categories: faces, scenes, objects, and color. In addition, scrambled faces, scrambled scenes, scrambled objects, and gray images were used. These images were collected from open source material databases and websites, including Japanese movie information (faces), Japanese real estate companies (scenes), and car dealers and supermarkets (objects). A block design was used for functional localization scans (Fig. 3). Each run consisted of eight blocks: a face block, a scrambled face block, a scene block, a scrambled scene block, an object block, a scrambled object block, a color block, and a gray block. Within each 12-s block, including rest duration (4 s), 16 collages of natural images belonging to a single block condition were presented sequentially for 0.5 s each. Participants pressed a button when each block started. These blocks were conducted in a random order at various intervals between 0–10 s. Intervals were determined pseudo-randomly using optseq2—a FreeSurfer tool for scheduling events (surfer.nmr.mgh.harvard.edu/optseq/). A black cross ($0.9^\circ \times 0.9^\circ$) was presented during the rest intervals. Each run (374 s total) consisted of 24 blocks, plus an initial 14-s black cross presentation that was excluded from analysis. All participants performed three runs.

Based on the functional localization scan, we defined freehand ROIs in the left and right ventral visual cortices (Fig. 4a) and left and right IPSs (Fig. 4b) for each participant individually. We selected the left and right ventral visual cortices because they are involved in face-, scene-, object-, and color-processing. To identify these areas, we contrasted the brain activities in response to face vs scrambled face blocks, scene vs scrambled scene blocks, object vs scrambled object blocks, and color vs gray blocks using a standard general linear model (GLM) approach in AFNI (afni_proc

function). The fMRI time series of each block condition was modeled by convolving a 16-s boxcar function with a canonical hemodynamic response function. We computed GLM β coefficients of standard AFNI parameters for each voxel, and estimated the contributions of each category (faces, scenes, objects, and color) to the blood oxygen level dependent responses of the voxels in the ventral visual cortex. For the mapping of the results of this comparison, the voxel-wise p -value threshold was set to .001 for faces, scenes, and objects and .01 for color because color selectivity has been shown to be weaker than the selectivity in other categories (Lafer-Sousa et al. 2016). We carefully defined ROIs as the statistically significant (not adjusted) voxel clusters in the visual anatomical brain regions for each participant individually. The average ROI sizes for all participants were as follows: left/right face area = 264/236 voxels, left/right scene area = 142/128 voxels, left/right object area = 205/139 voxels, and left/right color area = 110/113 voxels. These corresponded to a cluster threshold level of $\alpha \sim .10$ estimated using the 3dClustSim function in AFNI. We defined the left and right face processing related IPSs as the IPS because of the interplay between the IPS and face processing. The parietal cortex, including the IPS, is connected to the posterolateral aspect of the fusiform gyrus (FG) (pFus-faces) via the ventral occipital fasciculus (for a review, see Grill-Spector et al. 2018; Takemura et al. 2016). Moreover, the IPS has a vital role in spatial representations (Colby & Goldberg 1999; Fitzgerald et al. 2012; Goodale & Milner 1992), and the spatial relationships between the eyes, nose, and mouth are important for face perception (e.g., Tsao & Livingstone 2008). Hoffman and Haxby (2000) revealed that both the FG and the IPS responded to faces more than to control stimuli (scrambled pictures) in a face perception task. Thus, we established the IPS ROIs by comparing the brain activity in response to faces vs scrambled faces in the functional localization task using a similar procedure as that used for identifying the ventral visual cortices ($p < .05$, $\alpha \sim .10$). The average size of the left and right IPS ROI across all participants was 110 and 113 voxels, respectively. The results of MVPA for the ventral visual cortex and IPS were calculated by averaging the accuracy of the left and right ventral visual cortices and left and right IPSs because there were no significant hemispheric differences in MVPA accuracy.

The second type of scan was a retinotopy scan. Three 6-min retinotopy scans were performed to identify low-level and mid-level visual areas (Fig. 5). We used a standard phase-encoding method in which the retinotopic stimulus consisted of the natural images (same as those used in the functional localization scan) displayed within a wedge- or ring-shaped aperture traversing the visual field. The details of the procedure have been described in previous publications (e.g., Yamamoto et al. 2012; Yamashiro et al. 2014). Unfortunately, in the present study, we could not collect reliable retinotopic mapping data for most of the participants, which is likely attributed to the small number of the scans and/or participant fatigue; due to the time limits for use of the MRI scanner, we could only perform three retinotopy scans after the main experiment when the participants were tired. Since the data were mostly insufficient in size and quality, clear retinotopy maps were only obtained for two

participants from which we identified V1, V2, V3, and V3A/B. Based on AFNI's Talairach Atlas, V1, V2, and V3 were defined as the regions along the calcarine sulcus (Brodmann's area (BA) 17), the cuneus and lingual gyrus (BA 18) surrounding V1, and the cuneus and lingual gyrus surrounding V2, respectively. We also defined V3A/B, considering V3A as the area near the transverse occipital sulcus and V3B as the area lateral relative to V3A (for a review, see Wandell et al. 2007). The average ROI sizes across participants were as follows: left/right V1 = 193/289 voxels, left/right V2 = 129/160 voxels, left/right V3 = 127/152 voxels, left/right V3A/B = 174/136 voxels. The MVPA accuracy for the early visual cortices was calculated by averaging the left and right hemispheres because no significant hemispheric differences were found.

Results

Behavioral results

In the VWM experiment, participants performed the delayed roughness discrimination task in which they indicated whether the probe sphere had higher or lower roughness than the memorized sample sphere. Data from one participant who did not understand the task were removed from the analysis. Overall, participants performed well in this experiment: the mean accuracy was 68.04%, which exceeded chance ($t(22) = 8.51, p < .001, r = .88$).

fMRI MVPA results

We analyzed data from all participants except for the removed one for the fMRI analysis. First, we investigated the topographic relationship between the categorical processing visual areas in the ventral visual pathway and roughness VWM. Specifically, we localized face-, scene-, object-, and color-related areas in the ventral visual pathway and IPS. In the ventral visual pathway, the face-related area was localized near the FG, in keeping with literature reporting that this region is involved in face perception (McCarthy et al. 1997). The scene-related area corresponded to the parahippocampal place area where scene information is believed to be processed (e.g., Epstein & Kanwisher 1998). The object-related region appeared to correspond with the lateral occipital complex (Malach et al. 1995), and the color-related region was identified between the face- and scene-related areas (Lafer-Sousa et al. 2016). For each these areas, we averaged the MVPA accuracy from 8–12 s following the start of each trial (Fig. 6). We found that the MVPA accuracies were above chance in the face ROI of the ventral visual cortex and the IPS ROI. The face-related ventral visual cortex had 54.86% MVPA accuracy, which significantly higher than chance ($t(22) = 2.18, p < .05, r = .42$). In contrast, the MVPA accuracies of the scene-, object-, and color-related ventral visual cortices were not significantly different from chance (51.38%, $t(22) = 0.45, p = .66, r = .10$; 53.15%, $t(22) = 0.88, p = .39, r = .18$; 50.51%, $t(22) = 0.18, p = .86, r = .04$). The IPS showed 54.64% accuracy, which exceeded chance ($t(22) = 2.28, p < .05, r = .44$). No statistical differences in

MVPA accuracies were found between the left and right hemispheres for all regions (face-related ventral visual cortex: $t(22) = 1.77, p = .09, r = .35$; scene-related ventral visual cortex: $t(22) = 0.40, p = .69, r = .09$; object-related ventral visual cortex: $t(22) = 0.02, p = .98, r = .01$; color-related ventral visual cortex: $t(22) = 0.04, p = .97, r = .01$; and IPS: $t(22) = 0.94, p = .36, r = .20$).

Second, we examined whether the early visual cortex was involved in roughness VWM. We anatomically defined V1, V2, V3, and V3A/B and calculated the MVPA accuracy for each area during the delay period (Fig. 7). We found that the MVPA accuracies were not significantly different from chance in all ROIs (V1: 51.74%, $t(22) = 0.67, p = .51, r = .14$; V2: 51.23%, $t(22) = 0.47, p = .64, r = .10$; V3: 52.24%, $t(22) = 0.78, p = .44, r = .16$; V3A/B: 52.97%, $t(22) = 1.04, p = .31, r = .22$). There were also no statistical differences in MVPA accuracies between the left and right hemispheres in all regions (V1: $t(22) = 0.74, p = .47, r = .16$; V2: $t(22) = 1.56, p = .13, r = .03$; V3: $t(22) = 1.10, p = .28, r = .23$; V3A/B: $t(22) = 0.64, p = .53, r = .13$).

Discussion

In this study, we examined whether visual image information represented in the early visual cortex was involved in the maintenance of VWM representations of illumination-invariant visual properties. If the maintenance of VWM representations of roughness is independent of visual properties that substantially change appearance due to a change in illumination, visual image information in the early visual cortex should not be required. Alternatively, if visual image information is involved in maintaining stable illumination-invariant VWM representations of roughness (Ester et al. 2016; Christophel et al. 2017), these representations should include image-based information in the early visual cortex. To address this question, we focused on a single object surface property (i.e., roughness) and performed a delayed roughness discrimination task in which sample and probe stimuli differed in illumination. By applying MVPA to brain activity in the ventral visual cortex, IPS, and early visual cortex, we found that the ventral visual cortex and IPS, but not the early visual cortex, contributed to roughness VWM. This result suggests that a stable roughness VWM representation can be established without visual images features from the early visual cortex.

Comparison with previous studies on VWM

Our findings are not consistent with those of some neuroimaging studies on VWM. Previous studies have reported that VWM stability increases with distributed functions over multiple regions (e.g., Christophel et al. 2017). Although some studies have shown that the IPS has an important role in the maintenance of VSTM (e.g., Todd & Marois 2004), more recent studies have identified the early visual cortex as another potential region involved in VWM maintenance. For example, in the delayed orientation discrimination task used in Harrison and Tong (2009), participants had to memorize the orientation of a sample Gabor patch and judge whether the test Gabor was rotated clockwise or

counterclockwise compared with the memorized sample. Christophel et al. (2012) used more complex stimuli (i.e., color patterns) and conducted a delayed similarity detection task. In that task, participants had to indicate which of two probe color patterns was similar to the memorized sample color pattern.

Two factors may have contributed to the discrepancies in findings between our study and other published studies. First, differences in visual features can influence whether the early visual cortex is involved in VWM. While previous studies on VWM used a Gabor patch, we focused on an object surface property. Although the early visual cortex is involved in maintenance of simple visual features, complex features are retained in high-level areas. Second, differences in visual images between the memory phase and the discrimination phase in a delayed discrimination task can influence VWM maintenance in the early visual cortex. In past studies, participants had to retain a cued sample visual image in their VWM. Then, in the discrimination phase, participants had to judge the probe images based on the sample visual image maintained in VWM. However, in day-to-day life we often encounter visual images in illuminations and viewpoints that are different than those in which the image was first seen. In these conditions, visual input from the novel illumination or viewpoint of a previously-encountered stimulus can interfere with low-level representation in VWM, and the robustness of VWM may be reduced. Therefore, maintenance of low-level information in the early visual cortex may disrupt robust VWM representations. In these two factors, illumination-invariant visual features and changes in visual images between the sample and probe stimuli are not mutually exclusive, and either factor alone may disrupt VWM representation in the early visual cortex. This is a matter that needs to be examined in a future study.

The results of our study expand on previous studies investigating viewpoint in scene recognition (e.g., Xiao et al. 2010) that suggest that high-level areas such as the IPS show a viewpoint-invariant response. This study by Xiao et al. (2010) was inconclusive regarding the role of the early visual areas in scene recognition because they only used univariate analysis. In general, univariate analysis is not sufficiently sensitive to evaluate representation in a brain region, as evidenced by failures of univariate analysis in locating VWM representation in the early visual cortex (cf. Harrison & Tong 2009). In this respect, our findings with MVPA provide more convincing evidence that involvement of the early visual cortex is not essential for retaining a stable VWM representation.

Our results could have also been affected by top-down processes in the human brain. Bettencourt and Xu (2016) conducted a delayed orientation discrimination task, similar to Harrison and Tong (2009), except that visual distractors were presented during the delay period. When visual distractors were presented in a random trial, the orientation of a memorized Gabor patch was retained in the early visual cortex and IPS. In contrast, when the presence of visual distractors was predictable, only the IPS was involved in retaining the VWM representation. We built upon this finding in that the early visual cortex was not related to VWM when the participants could predict not only the presence of distractors but also environmental differences between a memorized sample stimulus and a probe

stimulus. Indeed, distribution of VWM representation in the human brain is altered depending on task demand (Lee et al. 2013; for a review, see Lee & Baker 2016). A future study should examine the top-down effect by manipulating the predictability of illumination change. If our results are based on stimulus features alone, the early visual cortex should not be related to VWM irrespective of the predictability of illumination changes. Alternatively, if a top-down factor influences the involvement of the early visual cortex in VWM, the early visual cortex should contribute to VWM only when illumination change is unpredictable.

More recently, Rademaker et al. (2019) employed a delayed-match-to-sample (DMTS) task and showed that memorized orientation was not impeded by noise distractors. Conversely, behavioral and memory reconstruction performance decreased in the presence of visual distractors, in contrast to the findings of Bettencourt and Xu (2016). Rademaker et al. (2019) attributed this discrepancy to difference in protocols, suggesting that the use of a single memory sample and continuous report design rendered VWM maintenance more sensitive to visual distractors during behavioral performance than that in the study by Bettencourt and Xu (2016). To account for distractor-resistant VWM, Rademaker et al. (2019) proposed that the early visual cortex may represent the input during VWM maintenance because different neuronal populations process sensory and memory representations; however, this theory predicts that VWM representation in the early visual cortex should have been observed in our study. This incongruity may be explained by differences in the levels of visual features used in each study (i.e., low-level Gabor patches in Rademaker et al. (2019) and high-level roughness renderings in the present study). Future research is needed to examine whether roughness VWM information can be interpreted from brain activity in a DMTS task for roughness.

Comparison with previous studies on material perception

We showed that the ventral visual cortex, but not the early visual cortex, is involved in illumination-invariant roughness representation in VWM. Although recent neuroimaging studies have revealed that the ventral visual cortex plays an important role in material perception (Sun et al. 2015; Wada et al. 2014), the early visual cortex was also reported to be involved because material information was perceived in a series of processes from low- to high-level areas. By examining the role of the ventral visual cortex during maintenance in a VWM task, we were able to distinguish the role of the ventral visual cortex from that of the early visual cortex in VWM for the object surface feature of roughness in different illumination conditions.

Our results are consistent with past findings on material perception (Hiramatsu et al. 2011; Sun et al. 2015; Wada et al. 2014). Material properties are encoded in various brain regions, from the early visual cortex to the ventral visual cortex. Although low-level visual features are needed for material perception, they are not sufficient on their own to represent material properties (Hiramatsu et al. 2011). Once material perception is established, visual information in the early visual cortex may no

longer be needed. Fleming et al. (2003) demonstrated that humans could match object glossiness and roughness between two spheres, even when these stimuli were rendered under different real-world illuminations. Moreover, Tsuda and Saiki (2018) revealed that the constancy of glossiness and roughness against illumination was also observed in VWM. These findings indicate that we can extract surface properties even when low-level features such as color, contrast, and highlights of the reflected illumination image (e.g., Chadwick & Kentridge 2015; Marlow et al. 2011) are lost because of illumination changes.

A recent neuroimaging study showed that the ventral visual cortex was involved in VSTM for material category. Otsuka and Saiki (2019) conducted image-based and material category-based change detection tasks and found differences in brain activation in the ventral visual cortex between these tasks. However, no previous studies have investigated brain processes of VWM for specific material properties (e.g., roughness) that are involved with the material categorization (Fleming et al. 2013). The main reason for the lack of the research regarding VWM for material properties lies in the difficulty of manipulating the stimuli (Fleming 2014; for a review, see Komatsu & Goda 2018). In this study, we manipulated a roughness parameter accurately using computer graphics and revealed that the ventral visual cortex was involved in VWM for roughness. Our finding expands on that of Otsuka and Saiki (2019) in that the ventral visual cortex is related to VWM/VSTM for both material properties and material category.

Conclusion

It has been shown that distributed brain regions, including the early visual cortex, form and maintain stable VWM representations. However, to achieve a stable VWM representation for illumination-invariant object properties, illumination-invariant representation may be formed without the need for information from the early visual cortex to distinguish changes between sample and probe stimuli. We performed a delayed roughness discrimination task and found that the ventral visual cortex and IPS, but not the early visual cortex, were involved in roughness VWM. This finding suggests that a stable VWM representation is not always accompanied by visual features from the early visual cortex during maintenance, even in the absence of distractors.

References

- Bettencourt KC, Xu Y (2016) Decoding the content of visual short-term memory under distraction in occipital and parietal areas. *Nat Neurosci* 19:150–157. <https://doi.org/10.1038/nn.4174>
- Cant JS, Arnott SR, Goodale MA (2009) fMR-adaptation reveals separate processing regions for the perception of form and texture in the human ventral stream. *Exp Brain Res* 192:391–405. <https://doi.org/10.1007/s00221-008-1573-8>
- Chadwick AC, Kentridge RW (2015) The perception of gloss: a review. *Vision Res* 109:221–235.

<https://doi.org/10.1016/j.visres.2014.10.026>

Chang CC, Lin CJ (2011) LIBSVM: a library for support vector machines. *ACM Trans Intell Syst Technol* 2:1–27

Christophel TB, Hebart MN, Haynes JD (2012) Decoding the contents of visual short-term memory from human visual and parietal cortex. *J Neurosci* 32:12983–12989. <https://doi.org/10.1523/JNEUROSCI.0184-12.2012>

Christophel TB, Klink PC, Spitzer B, Roelfsema PR, Haynes JD (2017) The distributed nature of working memory. *Trends Cogn Sci* 21:111–124. <https://doi.org/10.1016/j.tics.2016.12.007>

Colby CL, Goldberg ME (1999) Space and attention in parietal cortex. *Annu Rev Neurosci* 22:319–349. <https://doi.org/10.1146/annurev.neuro.22.1.319>

Cox RW (1996) AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Comput Biomed Res* 29:162–173. <https://doi.org/10.1006/cbmr.1996.0014>

Dale AM (1999) Optimal experimental design for event-related fMRI. *Hum Brain Mapp* 8:109–114

Debevec PE (1998) Rendering synthetic objects into real scenes: Bridging traditional and image-based graphics with global illumination and high dynamic range photography. *Proc SIGGRAPH*:189–198. <https://doi.org/10.1145/280814.280864>

Epstein R, Kanwisher N (1998) A cortical representation of the local visual environment. *Nature* 392:598–601. <https://doi.org/10.1038/33402>

Ester EF, Rademaker RL, Sprague TC (2016) How do visual and parietal cortex contribute to visual short-term memory? *eNeuro* 3: ENEURO.0041-16.2016. <https://doi.org/10.1523/ENEURO.0041-16.2016>

Fitzgerald JK, Swaminathan SK, Freedman DJ (2012) Visual categorization and the parietal cortex. *Front Integr Neurosci* 6:18. <https://doi.org/10.3389/fnint.2012.00018>

Fleming RW (2014) Visual perception of materials and their properties. *Vision Res* 94:62–75. <https://doi.org/10.1016/j.visres.2013.11.004>

Fleming RW, Dror RO, Adelson EH (2003) Real-world illumination and the perception of surface reflectance properties. *J Vis* 3:347–368. <https://doi.org/10.1167/3.5.3>

Fleming RW, Wiebel C, Gegenfurtner K (2013) Perceptual qualities and material classes. *J Vis* 13:9. <https://doi.org/10.1167/13.8.9>

Goodale MA, Milner AD (1992) Separate visual pathways for perception and action. *Trends Neurosci* 15:20–25. [https://doi.org/10.1016/0166-2236\(92\)90344-8](https://doi.org/10.1016/0166-2236(92)90344-8)

Grill-Spector K, Weiner KS, Gomez J, Stigliani A, Natu VS (2018) The functional neuroanatomy of face perception: from brain measurements to deep neural networks. *Interface Focus* 8:20180013. <https://doi.org/10.1098/rsfs.2018.0013>

Harrison A, Jolicoeur P, Marois R (2010) “What” and “Where” in the intraparietal sulcus: an fMRI study of object identity and location in visual short-term memory. *Cereb Cortex* 20:2478–2485.

<https://doi.org/10.1093/cercor/bhp314>

Harrison SA, Tong F (2009) Decoding reveals the contents of visual working memory in early visual areas. *Nature* 458:632–635. <https://doi.org/10.1038/nature07832>

Haynes JD, Rees G (2005) Predicting the orientation of invisible stimuli from activity in human primary visual cortex. *Nat Neurosci* 8:686–691. <https://doi.org/10.1038/nn1445>

Hiramatsu C, Goda N, Komatsu H (2011) Transformation from image-based to perceptual representation of materials along the human ventral visual pathway. *Neuroimage* 57:482–494. <https://doi.org/10.1016/j.neuroimage.2011.04.056>

Hoffman EA, Haxby JV (2000) Distinct representations of eye gaze and identity in the distributed human neural system for face perception. *Nat Neurosci* 3:80–84. <https://doi.org/10.1038/71152>

Just MA, Carpenter PA, Maguire M, Diwadkar V, McMains S (2001) Mental rotation of objects retrieved from memory: A functional MRI study of spatial processing. *J Exp Psychol Gen* 130:493–504. <https://doi.org/10.1037/0096-3445.130.3.493>

Komatsu H, Goda N (2018) Neural mechanisms of material perception: quest on shitsukan. *Neuroscience* 392:329–347. <https://doi.org/10.1016/j.neuroscience.2018.09.001>

Lafer-Sousa R, Conway BR, Kanwisher NG (2016) Color-biased regions of the ventral visual pathway lie between face- and place-selective regions in humans, as in macaques. *J Neurosci* 36:1682–1697. <https://doi.org/10.1523/JNEUROSCI.3164-15.2016>

Lee SH, Baker CI (2016) Multi-voxel decoding and the topography of maintained information during visual working memory. *Front Syst Neurosci* 10:2. <https://doi.org/10.3389/fnsys.2016.00002>

Lee SH, Kravitz DJ, Baker CI (2013) Goal-dependent dissociation of visual and prefrontal cortices during working memory. *Nat Neurosci* 16:997–999. <https://doi.org/10.1038/nn.3452>

Malach R, Reppas JB, Benson RR, Kwong KK, Jiang H, Kennedy WA, Ledden PJ, Brady TJ, Rosen BR, Tootell RB (1995) Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. *Proc Natl Acad Sci USA* 92:8135–8139. <https://doi.org/10.1073/pnas.92.18.8135>

Marlow P, Kim J, Anderson BL (2011) The role of brightness and orientation congruence in the perception of surface gloss. *J Vis* 11:16. <https://doi.org/10.1167/11.9.16>

McCarthy G, Puce A, Gore JC, Allison T (1997) Face-specific processing in the human fusiform gyrus. *J Cogn Neurosci* 9:605–610. <https://doi.org/10.1162/jocn.1997.9.5.605>

Motoyoshi I, Nishida S, Sharan L, Adelson EH (2007) Image statistics and the perception of surface qualities. *Nature* 447:206–209. <https://doi.org/10.1038/nature05724>

Otsuka S, Saiki J (2019) Neural correlates of visual short-term memory for objects with material categories. *Heliyon* 5:e03032. <https://doi.org/10.1016/j.heliyon.2019.e03032>

Rademaker RL, Chunharas C, Serences JT (2019) Coexisting representations of sensory and mnemonic information in human visual cortex. *Nat Neurosci* 22:1336–1344. <https://doi.org/10.1038/s41593-019-0500-0>

g/10.1038/s41593-019-0428-x

- 473 Saad ZS, Reynolds RC, Argall B, Japee S, Cox RW (2004) SUMA: an interface for surface-based
474 intra- and inter-subject analysis with AFNI. *Proc IEEE Int Symp Biomed Imaging* 2:1510–1513.
475 <https://doi.org/10.1109/ISBI.2004.1398837>
- 476 Sun HC, Ban H, Di Luca M, Welchman AE (2015) fMRI evidence for the processing of surface gloss
477 in human visual cortex. *Vision Res* 109:149–157. <https://doi.org/10.1016/j.visres.2014.11.012>
- 478 Takemura H, Rokem A, Winawer J, Yeatman JD, Wandell BA, Pestilli F (2016) A major human white
479 matter pathway between dorsal and ventral visual cortex. *Cereb Cortex* 26:2205–2214.
480 <https://doi.org/10.1093/cercor/bhv064>
- 481 Todd JJ, Marois R (2004) Capacity limit of visual short-term memory in human posterior parietal
482 cortex. *Nature* 428:751–754. <https://doi.org/10.1038/nature02466>
- 483 Todd JJ, Marois R (2005) Posterior parietal cortex activity predicts individual differences in vis
484 ual short-term memory capacity. *Cogn Affect Behav Neurosci* 5:144–155. <https://doi.org/10.3758/cabn.5.2.144>
- 485
- 486 Tsao DY, Livingstone MS (2008) Mechanisms of face perception. *Annu Rev Neurosci* 31:411–437.
487 <https://doi.org/10.1146/annurev.neuro.30.051606.094238>
- 488 Tsuda H, Saiki J (2018) Constancy of visual working memory of glossiness under real-world
489 illuminations. *J Vis* 18:14. <https://doi.org/10.1167/18.8.14>
- 490 Ungerleider LG, Courtney SM, Haxby JV (1998) A neural system for human visual working memory.
491 *Proc Natl Acad Sci USA* 95:883–890. <https://doi.org/10.1073/pnas.95.3.883>
- 492 Wada A, Sakano Y, Ando H (2014) Human cortical areas involved in perception of surface glossiness.
493 *Neuroimage* 98:243–257. <https://doi.org/10.1016/j.neuroimage.2014.05.001>
- 494 Wandell BA, Dumoulin SO, Brewer AA (2007) Visual field maps in human cortex. *Neuron* 56:366–
495 383. <https://doi.org/10.1016/j.neuron.2007.10.012>
- 496 Ward GJ (1994) The RADIANCE lighting stimulation and rendering system. *Comput Graphics*
497 18:459–472. <https://doi.org/10.1145/192161.192286>
- 498 Xiao C, McNamara TP, Qin S, Mou W (2010) Neural mechanisms of recognizing scene con
499 figurations from multiple viewpoints. *Brain Res* 1363:107–116. <https://doi.org/10.1016/j.br>
500 [ainres.2010.09.068](https://doi.org/10.1016/j.brainres.2010.09.068)
- 501 Xing Y, Ledgeway T, McGraw PV, Schluppeck D (2013) Decoding working memory of stimulu
502 s contrast in early visual cortex. *J Neurosci* 33:10301–10311. <https://doi.org/10.1523/JNEU>
503 [ROSCI.3754-12.2013](https://doi.org/10.1523/JNEUROSCI.3754-12.2013)
- 504 Xu Y, Chun MM (2006) Dissociable neural mechanisms supporting visual short-term memory for
505 objects. *Nature* 440:91–95. <https://doi.org/10.1038/nature04262>
- 506 Yamamoto H, Fukunaga M, Takahashi S, Mano H, Tanaka C, Umeda M, Ejima Y (2012) In
507 consistency and uncertainty of the human visual area loci following surface-based registr

508 ation: probability and entropy maps. Hum Brain Mapp 33:121–129. [https://doi.org/10.100](https://doi.org/10.1002/hbm.21200)
509 2/hbm.21200
510 Yamashiro H, Yamamoto H, Mano H, Umeda M, Higuchi T, Saiki J (2014) Activity in early visual
511 areas predict interindividual differences in binocular rivalry dynamics. J Neurophysiol
512 111:1190–1202. <https://doi.org/10.1152/jn.00509.2013>

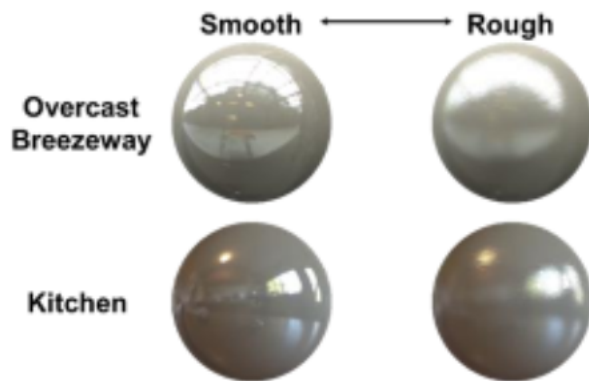


Fig. 1 Computer-generated three-dimensional spheres used as sample stimuli. The surface appearance of the sample spheres was set at one of two levels of roughness with jitters.

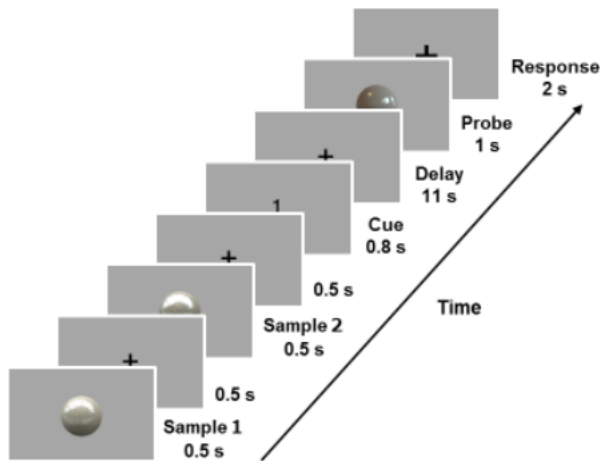


Fig. 2 Sample trial of the delayed roughness discrimination task. The two samples spheres were of different roughness. After an 11-s delay following the memorization cue, a probe sphere was presented, and participants were asked to indicate whether the probe had higher or lower roughness than the memorized sample sphere.

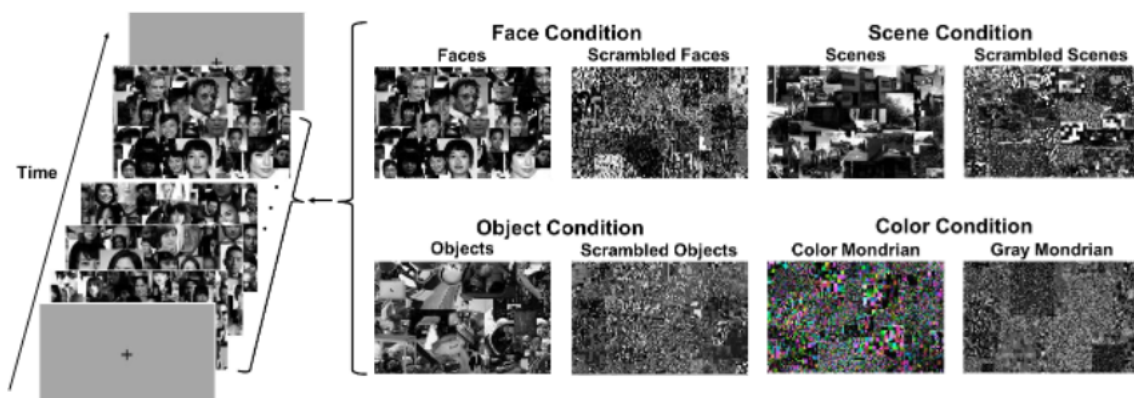


Fig. 3 Sample block of the functional localization scan. Within each block, collages of natural images belonging to a single block condition were presented for sequentially 0.5 s each.

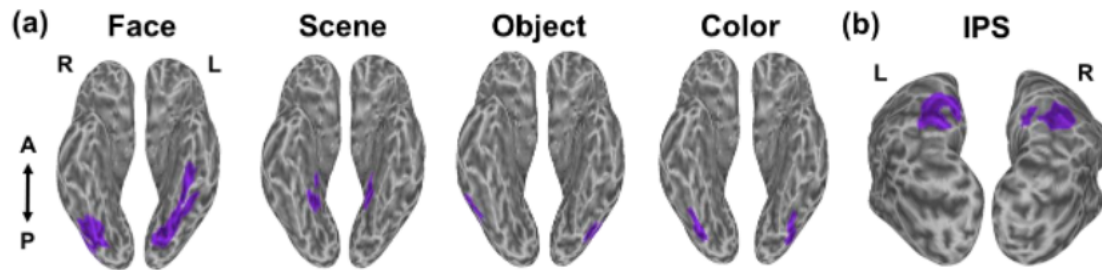


Fig. 4 Representative example of left and right extruded brain surfaces. We defined regions of interest based on the functional localization scans. **a** Left and right ventral visual cortices were identified, which are related to face-, scene-, object-, and color-processing. **b** Left and right intraparietal sulci (IPSSs) were localized based on the face condition in the functional localization scans.

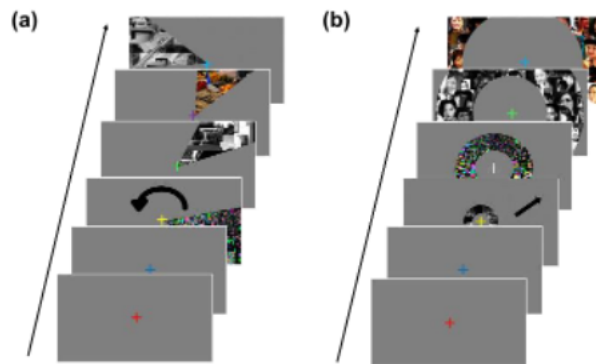


Fig. 5 Sample retinotopy scans. **a** A counterclockwise polar angle wedge run (CCW run) during which a 45° wedge rotated counterclockwise around the entire screen at 12 Hz for 60 s. The CCW run was conducted twice. **b** An expanding ring run (EXP run) during which a ring expanded across the entire screen for 60 s. The EXP run once was conducted once.

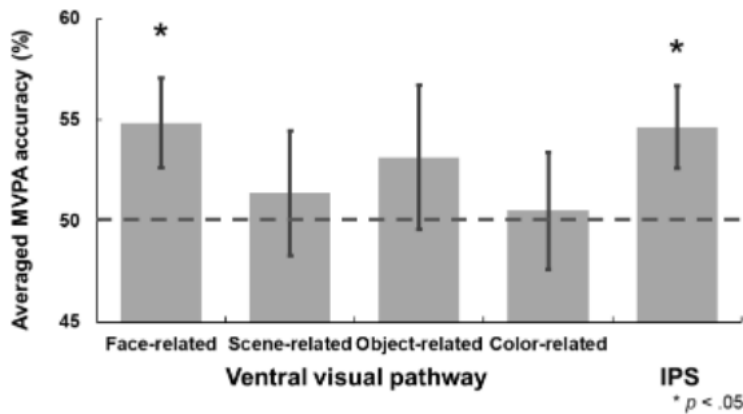


Fig. 6 Averaged multi-voxel pattern analysis (MVPA) accuracy for the dissociation of roughness (rough or smooth) during the period from 8–12 s following the start of each trial. The gray dotted line signifies the chance level of accuracy (50%) and the error bars show standard error. (IPS: intraparietal sulcus).

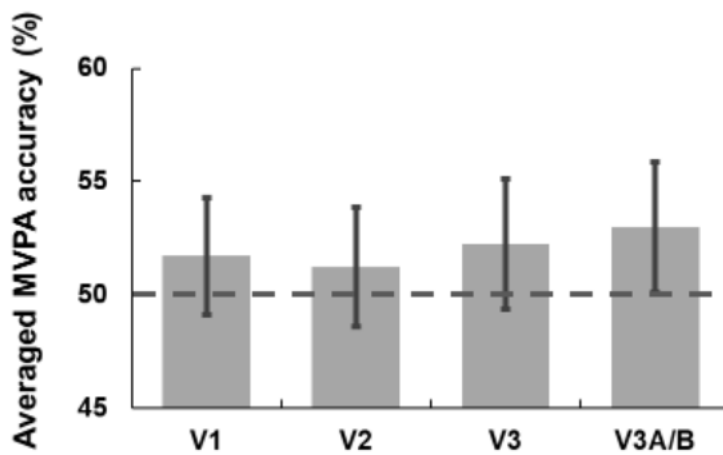


Fig. 7 Averaged multi-voxel pattern analysis (MVPA) accuracy in the early visual cortex during the delay period. The gray dotted line signifies the chance level of accuracy (50%) and the error bars show standard error.