

Neural activity in human visual cortex is transformed by learning real world size



Marc N. Coutanche^{a,b,c,*}, Sharon L. Thompson-Schill^d

^a Department of Psychology, University of Pittsburgh, Pittsburgh, PA, 15260, USA

^b Learning Research and Development Center, University of Pittsburgh, Pittsburgh, PA, 15260, USA

^c Brain Institute, University of Pittsburgh, Pittsburgh, PA, 15260, USA

^d Department of Psychology, University of Pennsylvania, Philadelphia, PA, 19104, USA

ARTICLE INFO

Keywords:

Size
Learning
Vision
Concepts
Animacy
Memory

ABSTRACT

The way that our brain processes visual information is directly affected by our experience. Repeated exposure to a visual stimulus triggers experience-dependent plasticity in the visual cortex of many species. Humans also have the unique ability to acquire visual knowledge through instruction. We introduced human participants to the real-world size of previously unfamiliar species, and to the functional motion of novel tools, during a functional magnetic resonance imaging scan. Using machine learning, we compared activity patterns evoked by images of the new items, before and after participants learned the animals' real-world size or tools' motion. We found that, after acquiring size information, participants' visual activity patterns for the new animals became more confusable with activity patterns evoked by similar-sized known animals in early visual cortex, but not in ventral temporal cortex, reflecting an influence of new size knowledge on posterior, but not anterior, components of the ventral stream. In contrast, learning the functional motion of new tools did not lead to an equivalent change in recorded activity. Finally, the time-points marked by evidence of new size information in early visual cortex were more likely to show size information and greater activation in the right angular gyrus, a key hub of semantic knowledge and spatial cognition. Overall, these findings suggest that learning an item's real-world size by instruction influences subsequent activity in visual cortex and in a region that is central to semantic and spatial brain systems.

1. Introduction

The neural activity in a person's visual cortex reflects both the current visual environment and their past experience. Neuronal responses of visual cortex become more selective after a monkey is trained to visually distinguish shapes (Baker et al., 2002; Op de Beeck and Baker, 2010), and repeated visual exposures can increase neural sensitivity in humans (Brants et al., 2016; Harel, 2016; Kourtzi et al., 2005; Sigman et al., 2005). Although these changes can be induced by repeated visual presentations (i.e., experience-dependent plasticity), humans do not require a large number of visual exposures to learn visual properties. A person can instead acquire this knowledge through language. Here, we investigate how activity in visual cortex is changed after humans learn an item's real-world size.

Knowing an item's real-world size is important for correctly judging its distance (which can be determined through the size of the current retinal imprint and knowledge of its actual size). Although few studies

have examined how learning the real-world size of new visual concepts through instruction affects brain systems, some prior studies have examined how real-world size is represented in the ventral stream. Such studies differ in two key dimensions: i) whether they probe how univariate responses vary across areas of cortex, or examine size in multi-voxel patterns; ii) the extent to which they find evidence of size differences in early visual cortex or higher-level ventral temporal (VT) cortex.

Several studies of univariate responses have found that perceiving differently sized man-made objects stimulates different areas of VT cortex, with a medial-lateral organization based on size (Konkle and Caramazza, 2013; Konkle and Oliva, 2012). A number of reasons have been suggested for this large versus small object difference, including variation in items' shape and material properties, their reliance on different parts of the retina (central versus peripheral), and our tendency to interact with small objects compared to using larger objects as landmarks. The importance of this last distinction has been supported by evidence that large objects activate typical scene areas (such as the parahippocampal

* Corresponding author. 3939 O'Hara Street, 646, University of Pittsburgh, Pittsburgh, PA 15260, USA.

E-mail address: marc.coutanche@pitt.edu (M.N. Coutanche).

<https://doi.org/10.1016/j.neuroimage.2018.11.039>

Received 29 August 2018; Received in revised form 8 November 2018; Accepted 22 November 2018

Available online 23 November 2018

1053-8119/© 2018 Elsevier Inc. All rights reserved.

place area) more strongly than do smaller objects (He et al., 2013; Julian et al., 2017). The idea that landmark-potential affects VT activity might also explain why studies examining univariate responses have not found differences in how large versus small animate items are represented in VT cortex. Unlike man-made objects, animate items are not potential landmarks (because they are mobile) and are not typically manipulated.

The above studies' findings come from activity collected while well-known concepts are presented visually. In contrast, several recent studies of non-man-made objects (words and shapes) have found that real-world size can be represented in early visual cortex. A recent study of perceptual versus conceptual properties of concepts presented as words found that their real-world size is reflected in multi-voxel patterns of early visual cortex (e.g., the activity pattern for “camel” was more similar to “cow” than to “goat” in Brodmann Area (BA) 17, after controlling for word length and semantic properties; Borghesani et al., 2016). As the ventral stream progressed anteriorly, real-world size became less influential, so that real-world size was not detectable beyond early visual regions. This supported the authors' framework of a perceptual-to-conceptual gradient in the ventral stream, where real-world size yields to more conceptual dimensions (Borghesani et al., 2016; also see Coutanche et al., 2016). The modulation of early visual cortex by information that is not visually apparent is consistent with other studies showing that early visual regions can be modulated by non-sensory information, such as an object's prototypical color (for grayscale images; Bannert and Bartels, 2013) and meaning (for ambiguous stimuli; Vandenbroucke et al., 2013). Similarly, primary visual cortex has been shown to reflect perceived size (rather than retinal size) in visual illusions (Fang et al., 2008; Murray et al., 2006). In another relevant study, Gabay and colleagues trained participants through extensive exposure to geometric shapes of different sizes, finding that early visual cortex activation was stronger for shapes that had previously been associated with larger sizes (Gabay et al., 2016). Finally, in a recent study of how real-world size is processed in visual cortex, Coutanche and Koch (2018) found that size was represented in early visual cortex beyond taxonomic category. By using animal species that break the typical correlation between real-world size and taxonomic category (e.g., insects that are bigger than birds, and birds that are bigger than mammals), the authors found pattern similarity based on real-world size after accounting for taxonomic and visual differences. Thus, when examining how the brain responds to visually presented items that are neither manipulable nor potential landmarks, real-world size information has been found in early visual cortex (Borghesani et al., 2016; Coutanche and Koch, 2018; Gabay et al., 2016). In some cases, this is accompanied by an absence of size information in VT cortex for these same items (Borghesani et al., 2016; Coutanche and Koch, 2018).

To test the idea that multi-voxel patterns in early visual cortex can be affected by knowledge of real-world size for visually presented concepts, we introduced human participants to images of animals from real, but unfamiliar, species, followed by knowledge about the species' size. We hypothesized that learning the unfamiliar species' real-world sizes would cause a shift in their underlying visual cortex activity patterns to become more similar (i.e., confusable) to known species of a similar size. A similar approach to examining brain changes after learning was recently taken by Bauer and Just (2015), who introduced abstract information about an unfamiliar animal's habitat and diet/eating habits. After learning this information, activity patterns (collected while participants were thinking about the animals) became more similar for pairs of animals with similar (learned) habitats or diets, in relevant regions (Bauer and Just, 2015).

A shift in pattern information can be measured through the ability of a classifier to distinguish patterns generated by the new and size-matched known species. A learning-induced *decrease* in classification accuracy would be consistent with a shift toward the size-matched known animals (i.e., reflecting the new size knowledge). In contrast, if a learning intervention fails to affect activity patterns, there would be no change in classification performance. A third possibility –an increase

in classification accuracy– would indicate an increased distinctiveness of activity patterns, outside the size dimension. For example, increased familiarity with viewpoints of an animal might lead to patterns that are more discriminable from other animals. In this case, the change in activity would not reflect size information (as otherwise, activity patterns for the new and size-matched animals would be more similar, leading to lower classification performance), but instead would reflect greater discriminability. Observing a decrease in classification performance in one region, with an increase in another, can be particularly informative, as the second region's rise in discriminability can rule-out brain-wide noise as being responsible for the first region's discriminability decrease.

How might new size information be maintained in visual cortex activity after learning? The semantic memory network includes several potential hubs that might play a role in modulating visual cortex activity (Lambon Ralph et al., 2017). One hypothesized hub, the anterior temporal lobe (ATL), has been linked to integrating features for known objects (Coutanche and Thompson-Schill, 2015a), making it a possible source of the learned size knowledge. Alternatively, a second hub –the angular gyrus (AG)– has also been linked to semantic integration (Lambon Ralph et al., 2017), in addition to spatial processing (Hirstein et al., 2011; Sack, 2009) “including the spatial analysis of external sensory information and internal mental representations” (Seghier, 2013). Notably, the real-world size of an item has direct spatial implications. This, combined with evidence that the right AG is also critical for perceptual learning (Rosenthal et al., 2009; Seghier, 2013) raises the possibility that the AG could play a role in linking perceptual inputs with size knowledge.

Here, we examine how participants' brains respond to knowledge about animal size because animate items cannot act as reliable landmarks. We chose to compare neural changes for animals to changes in a category that is frequently contrasted with animals, namely, tools (Almeida et al., 2010; Mahon et al., 2007, 2010). Examining neural representations for another type of item allowed us to ask whether any learning-induced changes are specific or could instead result from a general increase or decrease in attention due to changing familiarity. Neuroimaging investigations have suggested that human brain networks respond differently to tools and animals (Mahon et al., 2010) and tools differ from animals in a number of respects – they are manipulable, have a specific function, and do not move on their own. Because large man-made objects can take on neural characteristics associated with landmarks (He et al., 2013; Julian et al., 2017; an extreme example for tools being a crane), we focused on another important property of tools: their functional motion. Like size in animals, functional motion is an important defining property, as is reflected in our need for modifiers during naming (consider “miniature pig” or “swing saw”). Also like size, a functional motion can be learned through instruction, without needing to change the visual appearance of a presented item, allowing us to examine neural changes for a constant visual input. Despite these advantages, it is important to note that an observed change in one dimension and category (e.g., animal size) will not necessarily transfer to other dimensions or categories. With the context of this caveat, we ask whether learning the size of new animals, and functional motion of new tools, will affect activity patterns for humans observing still images.

2. Material and methods

2.1. Participants

Twenty-eight participants were scanned for the study. The data from four participants were removed from analysis because of excessive motion (three) or abnormal behavioral responses (one), leaving 24 analyzed participants (14 females; mean (M) age = 23.2, standard deviation (s.d.) = 5.7). Participants received compensation for their time, and the procedures were approved by the human subjects review board.

2.2. Experimental design

Participants were introduced to two new animal species and two new tools, while their brain activity was recorded over the course of seven functional magnetic resonance imaging (fMRI) scanner runs. During the first three runs (“pre-learning”), participants viewed images of four animal species (two unfamiliar; two familiar) and four tools (two unfamiliar; two familiar) while performing a 1-back task, in which they pressed a button when an image repeated. The unfamiliar items included tapirs, echidnas, pump-drills and wood planes (Fig. 1). A post-study questionnaire confirmed that these items were unfamiliar to participants: none of the 24 participants could identify the tapir or pump-drill; 23 could not identify the plane; 22 could not identify the echidna. Each familiar species was selected based on it having a similar size as one of the unfamiliar species: raccoon for echidna, and sheep for tapir. Each familiar tool had a similar functional motion as one of the unfamiliar tools (e.g., both sliding away from the user): saw for wood plane, and screwdriver for pump-drill. Images of twelve exemplars of each animal or tool, in a variety of viewpoints, were resized to have 500 pixels along their longest side, with each item being flipped to create 24 images. The first three runs each contained eight randomly ordered blocks (one for each animal and tool) separated by 12 s of fixation. Each block contained 24 images (23 unique and one randomly placed repeat) in a random order.

After collecting the (pre-learning) neural data, participants were introduced to information about each new item. At the beginning of this fourth run, to ensure attention, participants were informed that they would be tested on the forthcoming information. Subsequent text then communicated the real-world size and weight of each unfamiliar animal, and the grip and motion used with each unfamiliar tool (Table 1). Each of the four facts was presented for 12 s, followed by 9 s of the participant imagining viewing the animal or using the tool, and then 12 s of fixation. Next, to encourage task engagement, four true/false questions (one per item; two true, two false) were presented for 6 s (Table 2). Correctly responding to the true/false statements required knowing each fact. Participants indicated on a button-box if the displayed information was true. The facts and (a new set of) true/false questions were then presented once more to ensure the knowledge was acquired.

In the last three runs (‘post-learning’), the pre-learning procedure (blocks of images in a 1-back) was repeated with a new random block order. Finally, participants were asked about their pre- and post-study familiarity with each new animal/tool. Participants were shown an image of each new animal or tool and were asked: “On a scale of 1–5, how familiar were you with this animal [tool] before today, where 1 = not at all familiar, 3 = somewhat familiar and 5 = very familiar” and “On a scale

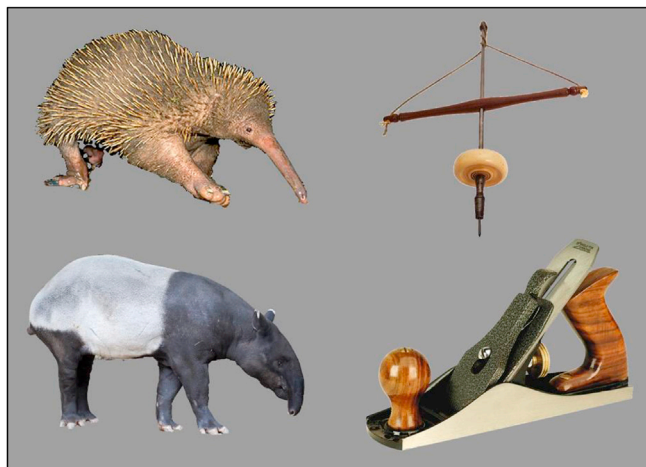


Fig. 1. Example stimuli for the four unfamiliar items. Top left: echidna; bottom left: tapir; top right: pump-drill; bottom right: wood plane.

Table 1

Semantic information communicated to participants for each unfamiliar animal and tool.

Echidna	This animal is between 1 and 1.75 feet long when fully grown. It stands 1 foot or less in height. It weighs between 10 and 13 pounds.
Tapir	This animal is between 6 and 7 feet long when fully grown. It stands 4 feet in height. It weighs between 400 and 800 pounds.
Wood plane	This tool is held by placing one hand on the front and using the other hand to grip the rear handle. To use, push the tool forward against a surface. Bring the tool back to its original position and repeat the action.
Pump drill	This tool is held by gripping the horizontal platform with the dominant hand. To use, push the platform down, causing the tool to spin and the string to become taut. Allow the platform to return to its original position and repeat the action.

Table 2

True/false questions asked as part of the learning phase to verify subjects were attending to the facts.

Echidna	The first animal is approximately the size of a soccer ball. The first animal is too big to easily hide.
Tapir	The second animal is approximately the size of a motorcycle. The second animal is small enough to easily hide.
Wood plane	The first tool is pushed forward across a surface. The first tool is operated with one hand.
Pump drill	The second tool is operated by pushing down. The second tool is operated with two hands.

of 1–5, how familiar do you feel with this animal [tool] now, where 1 = not at all familiar, 3 = somewhat familiar and 5 = very familiar”. Participants were also asked if they knew the name of each item. No participants could name the tapir or the pump-drill. Only one of the 24 analyzed participants could name the wood plane, and two could name the echidna.

2.3. Scanner acquisition

A 3 T Siemens Trio scanner with a 32-channel head coil was used to collect imaging data. A T1-weighted anatomical scan was acquired (TR = 1620 ms, TE = 3.87, TI = 950 ms, 1 mm isotropic voxels), followed by blood oxygen level-dependent echoplanar imaging (TR = 3000 ms, TE = 30 ms, 3 mm isotropic voxels). Seven functional runs were collected. Runs 1–3 (pre-learning) and 5–7 (post-learning) contained 132 TRs. Run 4 (learning) contained 140 TRs.

2.4. Data pre-processing

The collected data were preprocessed using the Analysis of Functional NeuroImages (AFNI) package (Cox, 1996). The first four TRs of each run were removed to allow the signal to reach steady-state magnetization. The functional data were processed using slice time correction, and motion correction to register volumes to the mean functional volume. Low frequency trends were removed with a high-pass filter (0.01 Hz). Voxel activation was scaled to have a mean of 100, and maximum of 200.

2.5. Regions of interest

Early visual cortex was sampled using a 3-voxel-radius sphere (123-voxel volume) placed at each participant’s calcarine sulcus. Prior cytoarchitectural examinations of human primary visual cortex (V1) at autopsy have shown that “the amount of cortical surface included in the calcarine sulcus provides a reasonable indication of V1 area” (Andrews et al., 1997, p. 2862). Additionally, the typical region of MT (V5) –an area linked to visual motion processing– was sampled by placing a 3-voxel-radius sphere at the left and right coordinates associated with visual motion processing in a seminal paper from Zeki and colleagues: at 38x, –62y, 8z and –38x, –74y, 8z (Table 2 in Zeki et al., 1991). These spheres

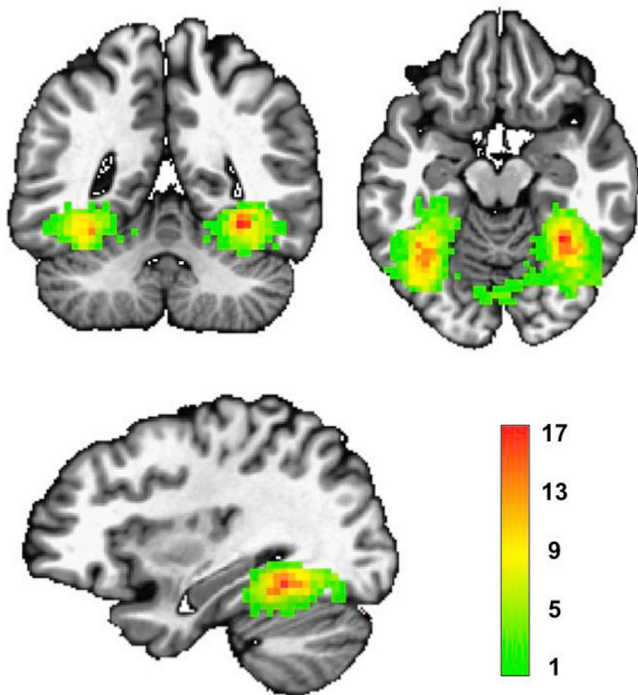


Fig. 2. An overlap map of participants' top 200 VT features. A searchlight was used to classify animals from tools in each participant's VT cortex. The central voxels of the top 200 searchlights were then used as features. The color scale reflects the number of participants with each voxel as a feature. Brain image Talairach coordinates: 32x, -46y, -13z.

were warped into each participant's native space. The left and right AG and ATL (Brodmann Area 38) were selected using AFNI's Talairach Atlas. Each region was then warped to each participant's native space.

To define the VT cortex anterior and posterior boundaries, we employed Talairach y-coordinates of between $y = -20$ and $y = -70$ (as used in Haxby et al., 2001). This definition includes bilateral parahippocampal, inferior temporal, fusiform and lingual gyri, with a mean VT volume of 4925 voxels (s.d. = 444). Because of the region's large size (relative to classified time-points), we avoided overfitting through an orthogonal feature selection. We ran an "animal versus tool" searchlight analysis (3-voxel radius) across the VT area (collapsed across both pre- and post-learning runs) to select searchlights through an orthogonal classification. Accuracy was allocated to the central voxel and the top 200 voxels in each participant were then used as VT features.

2.6. Statistical analysis

The preprocessed functional data were analyzed in MATLAB. The response amplitude of each voxel at each TR was first z-scored within each run. The condition labels were shifted forward in time by two TRs to account for the hemodynamic delay. Machine learning classifiers were trained and tested through a cross-validation procedure across independent runs (leave-one-run-out), ensuring independence between training and testing sets. A Gaussian Naïve Bayes (GNB) classifier was trained on voxel activity patterns at each TR. As well as reporting classification performance, we visualized shifts in neural representations using multidimensional scaling (MDS). Pre- and post-learning confusion matrices were submitted to an MDS analysis. The resulting two MDS plots (of the first two dimensions) were aligned with each other to allow comparison between the pre- and post-learning periods.

We next conducted an examination of how activity in semantic hubs (ATL and AG) covaried with the size information in early visual cortex. First, each participant's (post-shifted) time-points from the post-learning period were categorized based on the success (size discriminability) or

failure (size confusability) at classifying the new animals from the size-matched known animals in early visual cortex (i.e., echidna – raccoon; tapir – sheep). Next, we asked if activity in the ATL and AG differed for time-points that showed V1 size discriminability versus V1 size confusability. We did this by comparing the sets of pre-processed activity patterns (i.e., vectors of voxel responses) associated with these time-points (successfully versus unsuccessfully classified in V1) in terms of their overall activity (i.e., mean across-voxel response) and classification of their multi-voxel patterns.

3. Results

We examined how learning a species' real-world size impacts visually-driven activity in a learner's brain. We presented participants with previously unfamiliar species (Fig. 1) as their neural activity was examined via the blood-oxygen-level-dependent signal, collected through fMRI. Participants viewed images of known and unfamiliar species and tools, before and after being introduced to the animals' real-world size, and to tools' functional motion, to examine changes to visual activity after acquiring this new knowledge. Machine learning classifiers were trained to distinguish the collected neural activity patterns.

3.1. Behavioral performance

Participants' behavioral performance during the in-scan task (1-back) was high in both the pre-learning ($M = 91.7\%$, s.d. = 11.2%) and post-learning ($M = 90.2\%$, s.d. = 10.2%) periods. These pre and post periods did not differ significantly ($t_{19} = -0.79$, $p = 0.44$; from 20 of 24 participants due to a technical issue in four). The learning run included true/false questions about the learned information, to verify that participants were attending to the presented facts. The average accuracy on these true/false questions was high ($M = 78.8\%$, s.d. = 18.8%), particularly for the second set of facts, which occurred at the end of the learning run ($M = 97.2\%$, s.d. = 8.1%). At the end of the experiment, participants rated their familiarity with the new animals as being significantly greater after, compared to before, the study ($t_{23} = 8.66$, $p < 0.001$).

3.2. Discriminability of the new and known matched items in visual cortex

To examine if the new size knowledge influenced activity patterns in visual cortex, we quantified the correspondence between visual patterns for the new species, and patterns for a familiar species with a similar real-world size. We hypothesized that if participants' new size knowledge becomes reflected in neural activity, activity patterns for the new species should become more confusable with the familiar animals of a similar real-world size. We therefore examined decoding performance for the new and familiar size-matched species (tapir with sheep; echidna with raccoon), before and after learning. We compared this with changes to activity patterns for new and familiar tools with similar functional motions (wood plane with saw; pump drill with screwdriver). Classifications were conducted using activity patterns of early visual cortex, marked with a 3-voxel-radius sphere at the calcarine sulcus of each participant (an anatomical marker for the vicinity of V1, as used in prior work; Coutanche et al., 2011). Additionally, we asked the same question of activity patterns in VT cortex. The VT voxels, selected through an orthogonal feature selection (see Methods), are shown in Fig. 2.

We conducted separate 2×2 repeated measures ANOVAs to ask how the learning stage (pre- versus post-learning) and region (early visual cortex versus VT) predicted accuracy at classifying new items from their size- or motion-matched familiar items (adding an animal-size vs. tools-motion contrast through a $2 \times 2 \times 2$ repeated measures ANOVA revealed a significant 3-way interaction; $F_{1,92} = 3.87$, $p = 0.05$). For animal size, the time of data collection (pre- versus post-learning) significantly predicted classification performance ($F_{1,46} = 12.86$, $p < 0.001$). This in turn interacted significantly with region ($F_{1,46} = 14.13$, $p < 0.001$). Specific

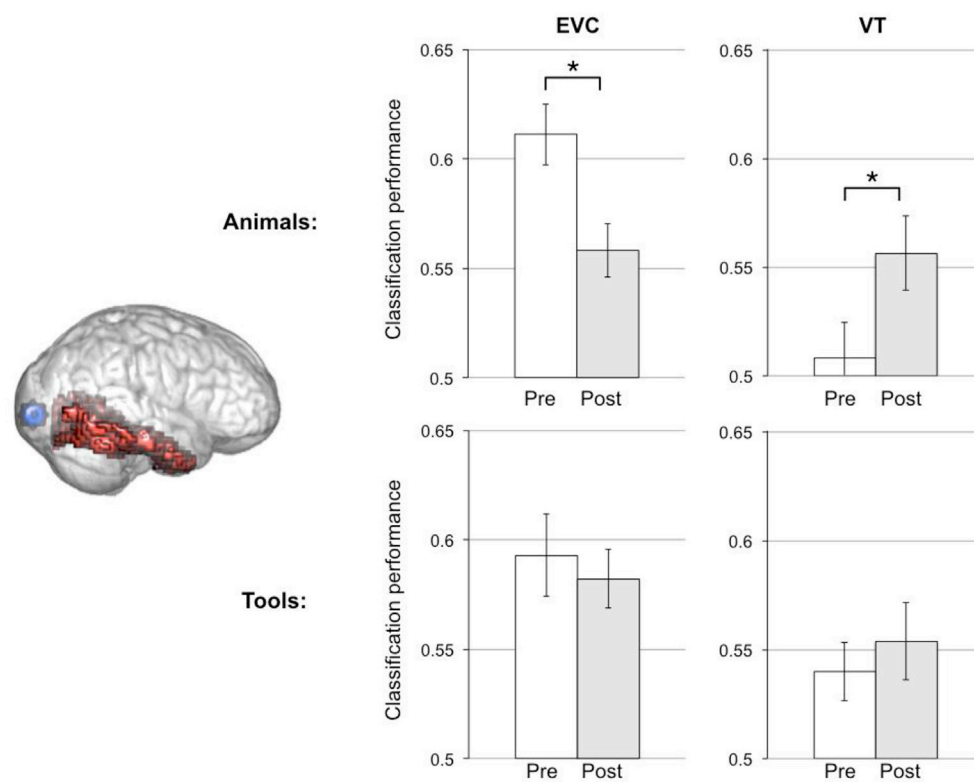


Fig. 3. Decoding new and matched familiar animals and tools before and after learning. Left: Regions-of-interest shown in a transparent brain, with early visual cortex (EVC) depicted in blue and ventral temporal (VT) cortex shown in red. Right top row: Classification performance at discriminating size-matched new and familiar species decreased in EVC and increased in VT cortex after learning. An asterisk indicates a significant difference ($p < 0.05$) in a two-tailed paired t -test. Right bottom row: Classification performance at discriminating the new and motion-matched familiar tools in EVC and VT cortex did not change after learning.

contrasts revealed that activity patterns for the new and size-matched species became more confusable in early visual cortex ($t_{23} = -3.58$, $p = 0.002$) after learning the new animals' size (before: $M = 0.61$, $s.d. = 0.07$; after: $M = 0.56$, $s.d. = 0.06$; Fig. 3; MDS plot in Supplementary Fig. 1). In contrast, in VT cortex—associated with higher-level object processing—the new and known (similarly-sized) species became more discriminable after learning ($t_{23} = 2.13$, $p = 0.04$; before: $M = 0.51$, $s.d. = 0.08$; after: $M = 0.56$, $s.d. = 0.08$; Fig. 3), reflecting a double dissociation between early and later visual regions. The presence of this reverse effect (increased classification accuracy after learning) also suggests the increased confusability in early visual cortex was not due to reduced engagement with the stimuli (or greater general noise), which would also have reduced VT performance. The left and right VT hemispheres did not differ significantly in their respective changes in decoding ($t_{23} = 0.54$, $p = 0.59$). A searchlight procedure was also conducted across the VT area to search for sub-regions that might show learning-induced changes for classifying new and size-matched species. No individual VT searchlights reached significance after correcting for multiple comparisons.

In contrast to the above results, a 2×2 repeated measures ANOVA predicting classification of new and motion-matched tools did not show a significant effect of learning stage (pre versus post; $F_{1,46} = 0.64$, $p = 0.43$) with no interaction by region ($F_{1,46} = 0.78$, $p = 0.38$). Examining this further showed that learning the tools' functional motion did not change pattern confusability in early visual cortex ($t_{23} = -0.54$, $p = 0.60$; before: $M = 0.59$, $s.d. = 0.09$; after: $M = 0.58$, $s.d. = 0.07$; Fig. 3). VT decoding also did not change significantly after learning ($t_{23} = 0.67$, $p = 0.51$; before: $M = 0.54$, $s.d. = 0.07$; after: $M = 0.55$, $s.d. = 0.09$; Fig. 3). Although not the primary focus of this study, we also asked whether activity patterns in MT—a visual motion region—would be affected by the learning instruction (Zeki et al., 1991). The new and motion-matched tools were discriminable both before ($M = 0.57$, $s.d. = 0.08$, $t_{23} = 4.05$, $p < 0.001$) and after ($M = 0.56$, $s.d. = 0.10$, $t_{23} = 3.20$, $p = 0.004$) learning, with no significant change between these stages ($t_{23} = -0.14$, $p = 0.89$).

3.3. Role of the semantic network

How are early visual cortex animal activity patterns modulated by new knowledge? To answer this, we first categorized post-learning time-points (TRs) based on whether each new and size-matched familiar species was confused by the classifier (i.e., tapir confused with sheep; echidna confused with raccoon) or not confused (i.e., discriminated). We then compared each participant's sets of confused versus non-confused time-points in hypothesized semantic hubs (ATL and AG). A hypothesized source of size information should be more active for time-points that have size information (indicated by a classifier confusing similar-sized species in early visual cortex) compared to time-points without size information. Time-points with size information in early visual cortex had greater right AG activation than time-points without size information ($t_{23} = 2.21$, $p = 0.04$). Activation levels did not differ in the left AG ($t_{23} = 1.30$, $p = 0.21$) or ATL (left: $t_{23} = 0.24$, $p = 0.81$; right: $t_{23} = 0.66$, $p = 0.52$). In addition to showing greater activation, the information within activity patterns of the right AG matched the information found in early visual cortex: time-points marked by early visual cortex size-confusion (i.e., the misclassification of size-matched animals) had right AG patterns with more size confusion ($t_{23} = -2.54$, $p = 0.02$) than time-points with early visual cortex size discriminability. This was not apparent in the left AG ($t_{23} = -0.95$, $p = 0.35$) or ATL (left: $t_{23} = -0.93$, $p = 0.36$; right: $t_{23} = -1.29$, $p = 0.21$).

4. Discussion

We have found that introducing participants to information about unfamiliar species' real-world size through instruction led to changes in activity patterns in early visual cortex when these animals were subsequently perceived. After learning the real-world size of two new species, neural activity patterns in early visual cortex become more confusable with visual patterns evoked by similar-sized known animals. In contrast, activity patterns in VT cortex became more discriminable, suggesting the early visual cortex confusability was not due to a reduction in attention,

or greater global noise, which would have also lowered VT classification. In contrast to animal size, learning the functional motion of new tools did not make them more confusable with motion-matched known tools. The presence of size information in early visual cortex co-occurred with stronger activation, and size-information, within the right AG.

Our finding that learning real-world size shifts patterns in early visual cortex to become more confusable with similar-sized known animate items might reflect the start of learning and perceptual processes that eventually lead to well-known concepts evoking early visual cortex patterns that reflect real-world size (Borghesani et al., 2016; Coutanche and Koch, 2018). The ability of instruction to provoke such neural changes might indicate a shortcut that is available to humans through language (for another example, see Bauer and Just, 2015). Specifically, the shift we report mirrors effects observed in early visual cortex after extensive perceptual experience. For example, extensive training with meaningless geometric shapes also leads to changes in early visual cortex based on their learned associated size (Gabay et al., 2016). More broadly, this study is consistent with observations that early visual cortex is modulated by more than sensation from the retina, but also position-invariant stimulus information (Williams et al., 2008), including semantic properties such as prototypical color in grayscale images (Bannert and Bartels, 2013) and perceived meaning (Vandenbroucke et al., 2013).

What computational role might real-world size play in early visual cortex? A lesion or stimulation study is required to determine its necessity for visual recognition, but one speculative possibility is that real-world size information in early visual cortex could help with determining distance to items in the environment. The true distance between an observer and an item can be calculated using the size of its retinal imprint and its real-world size. This real-world size is in turn used for calculating hand movement trajectories, the speed of objects moving in the distance, and so on. Future studies might wish to test such potential roles for real-world size information in early visual cortex by stimulating this area and measuring accuracy or response-time changes during relevant behavioral tasks.

Why did we not find a change in size information in VT cortex after learning? First, the increase in VT decoding that we observed from learning indicates that this was not due to poor signal. The lack of a change in size information in VT is consistent with prior work suggesting that (unlike for man-made objects) VT responses to animals are not spatially organized by size (Konkle and Caramazza, 2013). This past study did not find univariate differences in early visual cortex for differently sized animals, but this might be because examining multi-voxel patterns is required to detect this information (Coutanche, 2013). Indeed, it is notable that two recent studies of real-world size in multi-voxel patterns found real-world size information in early visual cortex, but not in more anterior regions – reflecting a decreasing trend for the representation of size (and greater representation of semantic category) as one proceeds anteriorly (Borghesani et al., 2016; Coutanche and Koch, 2018).

We also introduced participants to the functional motion of new tools. Although the new and familiar tools had similar functional motions, it is important to acknowledge that they differed in other ways, such as the specific grip used. In this study, our intent was to introduce a dimension that might also affect visual attention, but investigators wishing to study how learning a manipulation affects activity patterns might wish to select stimuli with matched grips and physical manipulation. It is also possible that a new motion must be observed (e.g., through a video clip) rather than described through text, to induce activity-pattern changes in areas that are sensitive to visual motion (like MT). A key limitation of our study is that we examined how learning real-world size affects animal patterns, and how learning functional-motion affects tool patterns, which differ in both category and dimension. We chose this combination because it allowed us to examine neural representations for two well studied visual categories (tools and animals), without a confound that with larger sizes, man-made objects can take on properties of landmarks (He et al., 2013; Julian et al., 2017). Unfortunately, this also removed our ability to speak

to the specificity of our effect. For example, the change we observed could be specific to real-world size because this dimension is relevant to interpreting the size of the retinal image, or might be specific to animals because manipulable objects are processed differently in the ventral and dorsal streams (Konkle and Caramazza, 2013; Mahon et al., 2010). Studies in the future might wish to examine the boundary conditions for such learning effects in terms of affected categories and dimensions. For example, a prior study found that associated (but not visually presented) size can change early visual activity for geometric shapes (Gabay et al., 2016).

Future work might also wish to explore how the in-scan task affects the degree to which VT regions are modulated by real-world size – particularly, if size must be explicitly retrieved to see learning-induced change. For example, some studies have instructed participants to imagine objects in their prototypical or atypical size (Konkle and Oliva, 2012) whereas, like this study, others have not (Borghesani et al., 2016; Coutanche and Koch, 2018). An intermediate approach is to instruct participants to think about an item embodying every feature (Bauer and Just, 2015). A related question is how features of the learning procedure might influence resulting neural changes. In our study, we verified learning-engagement by having participants make judgments about each item (Table 2). One possibility is that some questions (e.g., comparing the novel item's size with known items) might be easier than others (e.g., how a novel tool is operated). The role of question difficulty (and perhaps its association with visual imagery) could be a focus of future work that probes how learning interventions can be varied to induce different neural changes (see also Coutanche and Thompson-Schill, 2015b).

Our finding that the right AG was more active for time-points that had size information in early visual cortex is consistent with the AG's joint role in semantic integration and spatial cognition (Seghier, 2013). Analyzing AG multivariate patterns revealed that size confusability in the right AG co-occurred with size confusability in early visual cortex. This finding of inter-region information synchrony (Anzellotti and Coutanche, 2018) is consistent with these regions exchanging size-relevant information after learning. A role for the AG in maintaining recently learned size information in visual cortex integrates these semantic and spatial domains (Seghier, 2013). Our finding of size-relevant activity in the right but not left AG might share a basis with lateralization of spatial tasks that involve coordinate (rather than categorical) spatial relations, which are required for specifying precise distances (Baciu et al., 1999; Kosslyn et al., 1989).

Our use of a learning paradigm to examine the organization of real-world size helps support the idea that knowledge of real-world size influences neural activity in visual cortex, beyond the presence of correlations between size and mid-level perceptual features, such as texture, contours, shape, and other properties (also see Coutanche and Koch, 2018). Although visual features can co-vary with real-world size (Long et al., 2016), the modulation of visual cortex activity by a learning intervention suggests that expectations (based on knowledge) still play a role. Such knowledge might be necessary for size judgments of items that have similar shapes, but differ dramatically in their real-world size (for example, consider a golden retriever adult and puppy).

To conclude, we have found that learning the real-world size of unfamiliar species alters their visual activity patterns to become more similar to size-matched known species. The right AG appears to also play a role in supporting newly learned size knowledge. These findings contribute to the broader idea that more than being a purely bottom-up process, early visual processes can draw on “expectation or hypothesis testing in order to interpret the visual scene” (Gilbert and Li, 2013).

Acknowledgements

This work was supported by a grant awarded to S.L.T-S [R01EY021717] and a Ruth L. Kirschstein National Research Service Award to M.N.C. [F32EY024851] from the National Institutes of Health. The authors declare no competing financial interests.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.neuroimage.2018.11.039>.

References

- Almeida, J., Mahon, B.Z., Caramazza, A., 2010. The role of the dorsal visual processing stream in tool identification. *Psychol. Sci.* 21 (6), 772–778.
- Andrews, T.J., Halpern, S.D., Purves, D., 1997. Correlated size variations in human visual cortex, lateral geniculate nucleus, and optic tract. *J. Neurosci.* 17, 2859–2868.
- Anzellotti, S., Coutanche, M.N., 2018. Beyond functional connectivity: investigating networks of multivariate representations. *Trends Cognit. Sci.* 22, 258–269.
- Baciu, M., Koenig, O., Vernier, M.-P., Bedoin, N., Rubin, C., Segebarth, C., 1999. Categorical and coordinate spatial relations: fMRI evidence for hemispheric specialization. *Neuroreport* 10, 1373.
- Baker, C.I., Behrmann, M., Olson, C.R., 2002. Impact of learning on representation of parts and wholes in monkey inferotemporal cortex. *Nat. Neurosci.* 5, 1210–1216.
- Bannert, M.M., Bartels, A., 2013. Decoding the yellow of a gray banana. *Curr. Biol.* 23, 2268–2272.
- Bauer, A.J., Just, M.A., 2015. Monitoring the growth of the neural representations of new animal concepts. *Hum. Brain Mapp.* 36, 3213–3226.
- Borghesani, V., Pedregosa, F., Buiatti, M., Amadon, A., Eger, E., Piazza, M., 2016. Word meaning in the ventral visual path: a perceptual to conceptual gradient of semantic coding. *Neuroimage* 143, 128–140.
- Brants, M., Bulthé, J., Daniels, N., Wagemans, J., Op de Beek, H.P., 2016. How learning might strengthen existing visual object representations in human object-selective cortex. *Neuroimage* 127, 74–85.
- Coutanche, M.N., 2013. Distinguishing multi-voxel patterns and mean activation: why, how, and what does it tell us? *Cognit. Affect Behav. Neurosci.* 13, 667–673.
- Coutanche, M.N., Koch, G.E., 2018. Creatures great and small: real-world size of animals predicts visual cortex representations beyond taxonomic category. *Neuroimage* 183, 627–634.
- Coutanche, M.N., Thompson-Schill, S.L., 2015a. Creating concepts from converging features in human cortex. *Cerebr. Cortex* 25, 2584–2593.
- Coutanche, M.N., Thompson-Schill, S.L., 2015b. Rapid consolidation of new knowledge in adulthood via fast mapping. *Trends Cognit. Sci.* 19 (9), 486–488.
- Coutanche, M.N., Thompson-Schill, S.L., Schultz, R.T., 2011. Multi-voxel pattern analysis of fMRI data predicts clinical symptom severity. *Neuroimage* 57, 113–123.
- Coutanche, M.N., Solomon, S.H., Thompson-Schill, S.L., 2016. A meta-analysis of fMRI decoding: quantifying influences on human visual population codes. *Neuropsychologia* 82, 134–141.
- Cox, R.W., 1996. AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Comput. Biomed. Res.* 29, 162–173.
- Fang, F., Boyaci, H., Kersten, D., Murray, S.O., 2008. Attention-dependent representation of a size illusion in human V1. *Curr. Biol.* 18, 1707–1712.
- Gabay, S., Kalanithroff, E., Henik, A., Gronau, N., 2016. Conceptual size representation in ventral visual cortex. *Neuropsychologia* 81, 198–206.
- Gilbert, C.D., Li, W., 2013. Top-down influences on visual processing. *Nat. Rev. Neurosci.* 14, 350–363.
- Harel, A., 2016. What is special about expertise? Visual expertise reveals the interactive nature of real-world object recognition. *Neuropsychologia* 83, 88–99.
- Haxby, J.V., Gobbini, M.L., Furey, M.L., Ishai, A., Schouten, J.L., Pietrini, P., 2001. Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science* 293, 2425–2430.
- He, C., Peelen, M.V., Han, Z., Lin, N., Caramazza, A., Bi, Y., 2013. Selectivity for large nonmanipulable objects in scene-selective visual cortex does not require visual experience. *Neuroimage* 79, 1–9.
- Hirnstein, M., Bayer, U., Ellison, A., Hausmann, M., 2011. TMS over the left angular gyrus impairs the ability to discriminate left from right. *Neuropsychologia* 49, 29–33.
- Julian, J.B., Ryan, J., Epstein, R.A., 2017. Coding of object size and object category in human visual cortex. *Cerebr. Cortex* 27, 3095–3109.
- Konkle, T., Caramazza, A., 2013. Tripartite organization of the ventral stream by animacy and object size. *J. Neurosci.* 33, 10235–10242.
- Konkle, T., Oliva, A., 2012. A real-world size organization of object responses in occipitotemporal cortex. *Neuron* 74, 1114–1124.
- Kosslyn, S.M., Koenig, O., Barrett, A., Cave, C.B., Tang, J., Gabrieli, J.D., 1989. Evidence for two types of spatial representations: hemispheric specialization for categorical and coordinate relations. *J. Exp. Psychol. Hum. Percept. Perform.* 15, 723–735.
- Kourtzi, Z., Betts, L.R., Sarkheil, P., Welchman, A.E., 2005. Distributed neural plasticity for shape learning in the human visual cortex. *PLoS Biol.* 3 e204.
- Lambon Ralph, M.A., Jefferies, E., Patterson, K., Rogers, T.T., 2017. The neural and computational bases of semantic cognition. *Nat. Rev. Neurosci.* 18, 42–55.
- Long, B., Konkle, T., Cohen, M.A., Alvarez, G.A., 2016. Mid-level perceptual features distinguish objects of different real-world sizes. *J. Exp. Psychol. Gen.* 145, 95.
- Mahon, B.Z., Milleville, S.C., Negri, G.A.L., Rumiati, R.I., Caramazza, A., Martin, A., 2007. Action-related properties shape object representations in the ventral stream. *Neuron* 55 (3), 507–520.
- Mahon, B.Z., Schwarzbach, J., Caramazza, A., 2010. The representation of tools in left parietal cortex is independent of visual experience. *Psychol. Sci.* 21 (6), 764–771.
- Murray, S.O., Boyaci, H., Kersten, D., 2006. The representation of perceived angular size in human primary visual cortex. *Nat. Neurosci.* 9, 429–434.
- Op de Beek, H.P., Baker, C.I., 2010. The neural basis of visual object learning. *Trends Cognit. Sci.* 14, 22.
- Rosenthal, C.R., Roche-Kelly, E.E., Husain, M., Kennard, C., 2009. Response-dependent contributions of human primary motor cortex and angular gyrus to manual and perceptual sequence learning. *J. Neurosci.* 29, 15115–15125.
- Sack, A.T., 2009. Parietal cortex and spatial cognition. *Behav. Brain Res.* 202, 153–161.
- Seghier, M.L., 2013. The angular gyrus: multiple functions and multiple subdivisions. *Neuroscientist* 19, 43–61.
- Sigman, M., Pan, H., Yang, Y., Stern, E., Silbersweig, D., Gilbert, C.D., 2005. Top-Down reorganization of activity in the visual pathway after learning a shape identification task. *Neuron* 46, 823–835.
- Vandenbroucke, A.R.E., Fahrenfort, J.J., Sligte, I.G., Lamme, V.A.F., 2013. Seeing without knowing: neural signatures of perceptual inference in the absence of report. *J. Cognit. Neurosci.* 26, 955–969.
- Williams, M.A., Baker, C.I., Op de Beek, H.P., Mok Shim, W., Dang, S., Triantafyllou, C., Kanwisher, N., 2008. Feedback of visual object information to foveal retinotopic cortex. *Nat. Neurosci.* 11, 1439–1445.
- Zeki, S., Watson, J.D., Lueck, C.J., Friston, K.J., Kennard, C., Frackowiak, R.S., 1991. A direct demonstration of functional specialization in human visual cortex. *J. Neurosci.* 11 (3), 641–649.