CellPress

Review

Object Domain and Modality in the Ventral Visual Pathway

Yanchao Bi,[1,*] Xiaoying Wang,[1] and Alfonso Caramazza[2,3]

The nature of domain-specific organization in higher-order visual cortex (ventral occipital temporal cortex, VOTC) has been investigated both in the case of visual experience deprivation and of modality of stimulation in sighted individuals. Object domain interacts in an intriguing and revelatory way with visual experience and modality of stimulation: selectivity for artifacts and scene domains is largely immune to visual deprivation and is multi-modal, whereas selectivity for animate items in lateral posterior fusiform gyrus is present only with visual stimulation. This domain-by-modality interaction is not readily accommodated by existing theories of VOTC representation. We conjecture that these effects reflect a distinction between the visual features that characterize different object domains and their interaction with different types of downstream computational systems.

Ventral Visual Cortex: Visual or Multi-Modal?

A core assumption of cognitive science and cognitive neuroscience is that the brain processes information at various levels of representation, progressing from those closely tied to stimulus features to increasingly more general and abstract representations. One of the mysteries in this framework is the transition from modality specific representations – those explicable fully in the language of a given modality – to representations that capture other properties of the object – such as, for example, the possibility that a particular shape is appropriate for a certain type of grip. The conjecture we will articulate here is related to this difficult problem in cognitive science and cognitive neuroscience. In particular, we consider the representational distinctions, or the information encoded in such representations, that might give rise to the well-established domain-level organization in higher-order visual cortex (ventral occipital temporal cortex, VOTC), and the general principles that drive this organization.

The nature of the representations computed in this territory is one of the major topics of investigation in cognitive neuroscience. Various types of visual-level dimensions have been proposed and examined to account for the category-preferring distributions [1–7]. This visual-driven framework has recently been challenged by a wave of studies that reported similar domain preference effects in sighted and congenitally blind individuals, for example, for the animate–inanimate distinction, places, bodies, large objects, and tools [8–13]. A commonly shared contention in these articles, highlighted in a recent review article [14], is that 'These findings provide a consistent demonstration of the supra-modal functional organization of specific task-related cortical networks', marking a shift of sentiment about the VOTC, from being part of the visual cortex to being supra-modal and, at least partly, independent from visual experience. In that framework, 'supra-modal' was defined to be 'brain areas [that] are equally recruited and show overlapping patterns of connectivity, mainly directed toward multisensory brain areas, in both sighted and blind individuals and across different sensory modalities'.

However, this is a one-sided reading of the empirical findings. The literature on the effects of visual deprivation on selectivity for various object categories actually paints an intriguing pattern

Trends

A wave of recent studies has reported similar domain preference effects in ventral occipital temporal cortex (VOTC) in sighted and congenitally blind individuals, leading to the contention that object representation in this region is multi-modal.

However, the effects of visual deprivation on object category selectivity paint an intriguing pattern of heterogeneity: selectivity to spatial navigation stimuli and manipulable artifacts found to be robustly multi-modal, whereas selectivity to animate objects reliably present only for sighted individuals when processing visual stimuli.

We formulate a novel conjecture about the nature of representations in VOTC: representation types are partly driven by the nature of the mapping between object visual properties and other object properties, which differ between animate and inanimate objects.

[1]State Key Laboratory of Cognitive Neuroscience and Learning, and IDG/McGovern Institute for Brain Research, Beijing Normal University, Beijing, 100875, China
[2]Department of Psychology, Harvard University, Cambridge, MA, USA
[3]Center for Mind/Brain Sciences, University of Trento, Rovereto TN, Italy

*Correspondence: ybi@bnu.edu.cn (Y. Bi).

CrossMark

of heterogeneity, with selectivity robustly observed for some categories even in the absence of visual experience, whereas rarely observed, if at all, for some other categories (e.g., [12]). Relatedly, studies that considered the effect of input modality on category selectivity in sighted individuals have similarly found cross-modal selectivity more robustly for some categories than others in VOTC (e.g., [15,16]). Although it has been argued [14,17–20] that interpretation of these results require caution because potential effects obtained with nonvisual input may be affected (contaminated) by visual imagery, it would have to be further explained how imagery might play different roles for different object domains. The different effects of visual deprivation and input modality on category selectivity remain unexplained by existing theories of object representation in VOTC.

The aim here is to bring attention to this intriguing empirical phenomenon regarding the relationship between object domain and modality in VOTC, and propose a novel conjecture to explain this heterogeneity. The conjecture rests on the observation that the contrasting modality effects for different object domains might reflect a distinction between the visual features that characterize those different domains and their interaction with downstream computational systems and, in particular, action systems.

Domain Specificity Effects across Different Modalities Are Different

When sighted individuals view pictures, various clusters in VOTC are more responsive to certain categories of objects, such as faces, bodies, tools, or places. The overall distribution of category preference follows a broad animate versus inanimate distinction, with a further differentiation within the inanimate domain between manipulable and non-manipulable objects. This results in a tripartite organization, from ventral medial regions (parahippocampal and medial fusiform) showing preference for inanimate items broadly related to navigation, including scenes, places, buildings, and large non-manipulable objects, to lateral regions showing a preference for animate items including faces and animals, to more dorsolateral regions showing preference for bodies and small, manipulable objects [21–24] (see also [8,25] for the broad animate–inanimate pattern distinction).

In the following, we summarize the empirical findings regarding these several major object domain effects in VOTC along this tripartite distinction, bringing together the comparison between sighted visual versus nonvisual input and blind nonvisual cases. Studies using fMRI and positron emission tomography (PET) in which category-selective activations in VOTC in sighted or blind participants were investigated through at least one of the following types of stimuli: object names (including generating mental images of object names), object sounds, haptically presented objects, and objects presented through **sensory substitution devices** (see Glossary). We considered only those cases where category selectivity was tested by contrasting the target category to some other type of object category (control category). Experiments where nonvisual modalities were examined but did not yield positive results were included in an attempt to reduce potential file-drawer problems.

The results are presented in Figure 1A and Table 1. They show that the degree of consistency across input modalities and experience groups (blind versus sighted) differs greatly across object domains. Results for each specific domain effect are described in the following sections, beginning with nonvisual experiments in the sighted, followed by experiments in the blind.

Items Related to Spatial Navigation

It is well established that the medial fusiform gyrus/parahippocampal gyrus is more strongly activated when a sighted person sees pictures of scenes, buildings, or large objects relative to other objects [12,24,26,27]. Such selectivity is highly robust across various visual and nonvisual modalities within sighted individuals and across various nonvisual modalities in blind individuals

Glossary

Connectional fingerprints: the unique set of anatomical or functional connections a cortical region owns, which could be measured as the vector of the cortical region's connection strengths with other cerebral regions.

Functional fingerprints: the unique response properties a cortical region exhibits, which could be measured as vector of the region's response strengths to a variety of stimuli or tasks (e.g., object categories).

Multi-modal: for the purpose of this review, multi-modal was used to mean qualitatively similar task-related activation patterns across multiple sensory modalities even if the overall activation strengths were different.

Resting-state functional connectivity: the synchrony between spontaneous temporal fluctuations of brain activity of distinct brain regions in the absence of external stimuli.

Sensory substitution devices: devices that transform stimuli presented in one sensory modality into stimuli of another sensory modality. For example, shape-to-sound sensory substitution devices transform visual images into sounds through isomorphic mapping from plane coordinates and brightness of visual pixels to timing, frequency, and loudness of sounds.

(A) Literature findings about brain regions showing multi-modal category-specific response to animal, tool, human body, human face or scenes and large objects.

Key:
- ● Animal
- ● Human face
- ● Human body and body parts
- ● Scene and large object
- ● Tool

(B) Voxel-wise maps showing between-group or between-modality similarity of connectional or functional fingerprints

Correlation across voxels
$r = 0.56$

Correlation across voxels
$r = 0.71$

| 0.39 | r | 0.95 |
| Between-group connectivity pattern similarity | | |

| −0.44 | r | 0.97 |
| Between-group category activation pattern similarity | | |

| −0.65 | r | 0.98 |
| Between-modality category activation pattern similarity | | |

(C) Functional and connectional properties of characteristic tripartite ROIs

Left anterior medial fusiform gyrus/ parahippocampal gyrus

Left LOTC

Left lateral posterior fusiform gyrus

Key:
—— Sighted visual —— Sighted auditory —— Blind auditory

Trends in Cognitive Sciences

Figure 1. Schematic Summary of Effects of Visual Experience on Category Selectivity in Ventral Occipital Temporal Cortex (VOTC). (A) Brain areas showing category selectivity in blind and sighted nonvisual experiments in VOTC. The studies are reported with the reference numbers in Table 1 and the reference list, with asterisks indicating findings with blind participants. Animal (green): [22]; human face (magenta): [18,19,30–32,34]; human body and body parts (orange): [10*,11*,30,49]; scene and large object (blue): [9*,12*,15,16,18,19,28]; Tool (red): [13*,22,42,44,45,48]. The position of each study is derived from the peak coordinates reported. Only positive findings can be shown here; studies where category selectivity was not observed in nonvisual modalities can be found in Table 1. (B) Comparison of the functional and connectional fingerprints of the VOTC voxels in the blind and sighted groups. The brain maps show voxel-wise similarity of resting-state functional connectivity patterns between blind and sighted groups, of category response patterns between blind and sighted nonvisual experiments, and of category response patterns between sighted visual and nonvisual modalities, respectively. Warmer colors indicate greater similarity, which are primarily found in the left anterior medial fusiform gyrus/parahippocampal gyrus and lateral occipital temporal cortex (LOTC). The lateral posterior fusiform gyrus showed colder colors, indicating less similarity. The scatter plots demonstrate the correspondence between maps across voxels. The contours mark the fusiform gyrus (solid), parahippocampal gyrus (dashed), and inferior temporal gyrus (dash-dot). (C) Properties of the characteristic tripartite VOTC region of interests (ROIs) in blind and sighted individuals. For each ROI, the connectional fingerprint maps show the resting-state functional connectivity strength between each region with the seed ROI (t values); the functional fingerprint map shows the response strengths to the 16 object categories (β values). Panels (B) and (C) are adapted and reprinted with permission from the Society for Neuroscience [51].

Table 1. Summary of the Category Selectivity in VOTC Results in Sighted Nonvisual and Blind Nonvisual Experiments, Organized by the Tripartite Distribution

Anatomical Region	Functional Selectivity	Sighted Nonvisual		Blind	
		Selectivity Observed	Selectivity Not Observed	Selectivity Observed	Selectivity Not Observed
Anterior medial fusiform/ parahippocampal	Scenes and large objects	Auditory word [12] Written word [15,28] Visual imagery [18,19][a] Haptic [9][a] Sound [16][a]		Haptic [9][a] Auditory word [12]	
Lateral occipital temporal cortex (LOTC)	Tools	Auditory word [13,48] Written word [22,48] Sound [42,44], [45][a]	Written word [38][a]	Auditory word [13]	
	Body	Haptic [11,30], [49][a] Visual imagery [49][a]	Auditory word [50] Written word [50] Visual imagery [30][a]	Sensory substitution device [10] Haptic [11]	
Lateral posterior fusiform	Animals	Written word [22]	Auditory word [8,48], [12][a] Written word [38,48] Sound [42–44], [16,45][a] Mental imagery [41]		Auditory word [8], [12][a] Mental imagery [40]
	Faces	Visual imagery [18,19,31][a] Haptic [30], [34] (left fusiform), [32][a]	Written words [15] Visual imagery [30][a] Haptic [31,33][a]		Haptic [31,33][a]

[a]Analyses were carried out using region-of-interest (ROI) analyses or within predefined masks, while the whole-brain analyses were either not performed or yielded no significant results

(i.e., robust across task input and experience), including haptic exploration to Lego scenes (both blind and sighted), imagery generation of buildings, semantic judgment or size judgment of visual or auditory names of places or large non-manipulable objects (both blind and sighted), or listening to sounds associated with landmarks [9,12,15,16,18,19,28].

Animate Items
Lateral posterior fusiform is known to be more strongly activated by pictures of animate items such as faces and animals relative to other objects [22,24,29]. Various studies investigated face-preferring responses in nonvisual tasks in sighted individuals and in blind groups, such as generating visual images in response to auditory cues or descriptions, performing semantic judgment on printed names of famous people, and haptically exploring face masks [15,18,19,30–34]. With sighted individuals performing imagery, haptic, or word tasks, contradictory findings that yield no clear consensus were obtained. For the blind group, congenitally blind participants did not show face selectivity in fusiform gyrus in haptic tasks [31,33], whereas late blind participants did [31].

For non-human animals, using nonvisual object input, selectivity in lateral posterior fusiform was even more fragile. Nearly all papers that included tests of animal selectivity in sighted individuals that used written or spoken object names [8,12,22,35–41] or object sounds [16,42–45] failed to find selectivity for animals relative to other objects.

Items Related to Bodies and Manipulable Objects
The lateral occipital temporal cortex (LOTC) is known to show preference to images of body parts and small manipulable objects, which has been interpreted to reflect the bodily motor

components associated with these stimuli [46,47]. The preference for tools in LOTC was robust in nearly all studies considered, including when sighted subjects responded to printed or auditory object names [13,22,48] and to object sounds [42,44,45]. Similar tool selectivity in LOTC was also observed in congenitally blind individuals [13].

For preference to body parts in LOTC, studies have been carried out testing verbal or haptic stimuli in sighted people and haptic stimuli or sensory substitution device-generated sound stimuli in congenitally blind participants. Haptic exploration of body parts in both sighted and blind individuals, as well as body shape conveyed through sensory substitution devices in blind subjects, consistently activated at least part of LOTC [10,11,30,49]. The results using words are less consistent [30,49,50].

This review shows that the results for sighted and blind participants in experiments that used nonvisual stimuli allow the following empirical generalization: highly convergent results are obtained for sighted and congenitally blind participants; not all object domain selectivity is equally robust across modalities in both subject populations. Although the effects for large objects and scenes (spatial navigation stimuli) and manipulable artifacts show a robust **multi-modal** nature, the effects for animate objects were only observed in sighted individuals and only when processing visual stimuli, with little evidence for domain effects in blind participants or in sighted participants in experiments with nonvisual stimuli. Thus, domain-selective VOTC is neither uniformly multi-modal (amodal) nor uniformly unimodal.

We recently provided direct support for the pattern of object domain by modality and experience effects in VOTC gleaned from the various studies reviewed here [51] (Figure 1B,C). We systematically compared category-related responses and **resting-state functional connectivity** patterns between congenitally blind and sighted individuals across the whole VOTC. We obtained voxel-wise, large-scale, continuous maps of the degree to which **connectional** and **functional 'fingerprints'** of ventral visual cortex depend on visual experience or input. There was close agreement between connectional and functional maps, pointing to a strong interdependence of connectivity and function in VOTC. We observed that although visual input and experience (or the absence thereof) had a pronounced effect on the response and connectivity profiles of early visual cortex, their effects on higher-order 'visual' cortex were not homogeneous. Specifically, the functionality and connectivity of lateral posterior fusiform gyrus was strongly affected, whereas those of the anterior medial (showing strongest responses to navigation-related objects) and posterior lateral parts of ventral visual cortex (showing strongest responses to tools and body-related items) were statistically indistinguishable between blind and sighted individuals.

In short, when the available literature on sighted and congenitally blind individuals is considered from the perspective of modality specificity it reveals a domain-by-modality interaction. Intriguingly, this phenomenon has been in plain sight for some time but it escaped attention most likely because it does not fit easily within current views about the nature of representations computed in VOTC (see Box 1 for details).

Relationships between Visual and Other Object Properties

We formulate a novel conjecture about one of the factors that determines the nature of representations in VOTC: the difference across object domains in terms of modality effects lies in the relationship between visual shape and its functional relevance, understood as the types of computations it triggers downstream and ultimately its connectivity structure. Visual shape strongly constrains the way in which we interact physically with inanimate objects, but much less so and in a far less articulated manner with animate objects. As a consequence, the visual features of artifacts are defined jointly by their visual characteristics and their potential

---

Box 1. How Well Do Existing Theories Work?

Is the phenomenon of domain-by-modality interaction explained by current theories about VOTC representation?

Object Visual Property Hypothesis

Sophisticated vision theories have been developed about how different neural patches in the ventral visual pathway come to prefer specific object categories. The theories emphasize the role of types of visual information that are robustly associated with object categories, such as eccentricity, size, rectilinear, or curvature features [1–3,6,7,61]. Accounts of this type do not explain why selectivity for inanimate domains is strongly independent of visual experience and input, whereas selectivity for animate objects seems to depend on visual experience and input. Specifically, it is unclear how the proposed visual properties (e.g., eccentricity) lead either to the multi-modal or modality-specific nature of these patches of VOTC. These hypotheses of the factors that determine domain specificity are not incompatible with the observed modality heterogeneity of VOTC, but neither do they provide an explanation for the observed phenomenon.

Connectivity–Constraint Hypothesis

It has been proposed that domain specificity in VOTC is driven in large part by structural and/or functional connections with other brain regions that process nonvisual properties of the corresponding categories [54,62–65]. However, this hypothesis, too, is silent on why domain-preferring regions in VOTC are differentially affected by visual experience and input type.

The interaction of domain by modality showing visual specificity for animate objects in VOTC is reminiscent of explanations of category-specific semantic deficits for living things, which proposed that such deficits reflect damage to a visual semantic system, which is especially important for living categories [4,37,66–69]. Whatever the merits of this hypothesis (see [70,71] for critical analysis), the observation that visual properties are relatively more salient for living versus non-living things does not explain why domain-preferring regions in VOTC differ in terms of whether they respond only to visual object stimuli or to multiple modalities and types of stimuli (e.g., tactile input, spoken words).

---

action value (e.g., the handle of a cup), making such representations accessible through different modalities and, hence, multi-modal.

This conjecture is based on the assumption that, particularly for inanimate objects, shape properties severely constrain their motor and function representations; that is, how we interact with them and what they are used for. For example, a thin longish shape (blade) could afford the motor act and function of cutting; a flattish solid shape could afford the motor act and function of pounding. Small artifacts can be manipulated, whereas large artifacts generally indicate a fixed location [52,53] and involve whole-body movements such as approaching or going around them. Because of this, 'visual' representations of inanimate objects are typically parsed to reflect those visual characteristics that are relevant for physical interaction. In other words, at this level of visual representation the information encoded is not arbitrary object parts but shape properties that map naturally onto action-based systems [54,55]. Such articulated correspondence across modalities makes the VOTC representation (and probably other related modality representations) directly addressable from different input systems and hence multi-modal, at least in this sense.

For animate items the picture is rather different. The relationship between visual shape properties (at least at the level of size or shape) and how we interact motorically with an animal is much less articulated. An elongated animal could either be a harmless, even beneficial friend (e.g., earthworm) or a dangerous enemy (venomous snake). Two spiders with similar shape may elicit different types of actions – one can be benign and 'useful' for humans, yet another can be deadly and requires opposite reactions. The visual shape similarity between a wolf and a dog is greater than that between a dog and a cat, yet very often the latter two elicit a more similar reaction (petting instead of fleeing) from human observers than the former two. Item size is not indicative of how to react either. Both small animals (leech or black widow spider) and large animals (e.g., tiger) may elicit a same 'flee' action, just as both small animals (rabbit) and large animals (deer) may elicit a same 'catch-for-food' act. That is, the relationship is more 'holistic' and arbitrary – for example, potential danger is not associated with any particular shape feature. Owing to the lack of articulated correspondence between visual shape properties and properties

Box 2. Anatomical Properties of Subregions in VOTC

What do we know about the anatomical properties of subregions in higher-order VOTC that might be related to our functional proposal?

For the fusiform gyrus, a rostrocaudal shift from unimodal to polymodal cortex within the fusiform gyrus in humans has been proposed [37]. A series of recent studies (see review in [29]) highlighted a stable macroanatomical structure, mid-fusiform sulcus (MFS), that consistently marks various anatomical property and functional topological changes. Anatomically, the lateral and medial (posterior) fusiform presented distinct properties in terms of cytoarchitectonics, receptor architectonics, myelination, and white-matter connectivity [72–76]: the region medial to MFS has a columnar organization, whereas the one lateral to MFS is not columnar and has a higher cell density [72,73,76]. Also aligned with MFS are the transitions in the receptor organization and the measures of tissue contrast associated with myelin gradients in VOTC [74,75]. Differential long-range white-matter connections to VOTC were observed, constraining functional topology associated with functional selectivity to faces and scenes [62,77].

The anatomical distinction between lateral and medial fusiform has been suggested to be in line with the different functional preferences and eccentricity [29]. Our analysis here of the functional profile of VOTC, showing a modality-by-object domain interaction, raises new specific questions regarding the relationship between the underlying neural architecture with the more concretely defined functional properties. More specifically, how are the anatomical properties of lateral posterior fusiform (e.g., higher cell density without columnar structure) related to animate shape processing, and how are those of the medial fusiform (columnar cell organization) and LOTC related to the mapping between scene layout/navigation and shape/manipulation, respectively? Crucially, in light of the analysis here, how is the former domain shaped by visual experience and how are the latter two domains independent of it?

of other modalities, the visual representations are more 'isolated' from other inputs for animals. As a consequence, fusiform 'feature' representations do not interact directly with other sensory–motor systems, making them relatively inaccessible from other input modalities, and hence their more strictly visual representational format.

Intriguingly, for some categories selectivity is found in multiple subregions within this tripartite distribution. For instance, body parts not only induce selective responses in LOTC [56–58] but also in lateral posterior fusiform [57–60]. The different roles played by these two regions in body part recognition remain to be worked out. However, according to our conjecture, these multiple subregions represent different information types, which interact with different downstream processing systems. The representation of body parts in LOTC is appropriately formatted for interaction with the motor system, whereas that in lateral posterior fusiform is appropriate for further processing within the visual processing system itself. This view predicts that body selectivity in lateral posterior fusiform gyrus is only found for visual stimuli, whereas body selectivity in LOTC is relatively more immune to modality changes. The unimodal/multi-modal functional specialization in domain-preferring VOTC may be related to recent findings about the anatomical heterogeneity within this region, which show that the lateral and medial fusiform have distinct anatomical properties (Box 2).

Concluding Remarks

We began this review article by drawing attention to an empirical phenomenon regarding the effect of visual experience (and stimulus input) on object representations in higher-order visual cortex, showing that contrary to recent claims that this territory is multi-modal, there is a clear animate/inanimate dimension along which effects of modality differ. This empirical pattern motivates a novel conjecture about the nature of representations in VOTC: representation types are partly driven by the nature of the mapping between object visual properties and other object properties, such as how we interact motorically with them, and generates a line of predictions that could be empirically tested (see Outstanding Questions).

Outstanding Questions

The conjecture presented here emphasizes a particular characteristic of domain properties in shaping the organization of VOTC: the nature of the mapping between visual and other object properties. This generates a line of empirical predictions and questions that are yet to be tested:

The correlation between visual properties (e.g., shape) and other types of properties (e.g., motor/manipulation, navigation, function) should be significantly higher in artifacts than in animate objects. If similarity spaces were generated for various types of features, the correlation across visual feature similarity space and specific nonvisual feature similarity space would be greater for artifacts than for animate items, and the cross-feature space correlation computed would be predictive of how 'multi-modal' an object category's representation is in VOTC.

The brain connectivity (structural and/or functional) pattern between VOTC and other modality streams (e.g., motor) can be decoded from visual features (e.g., curvature, shape, size) for artifacts but not necessarily for animate things.

An important question concerns the relationship between the representations in VOTC and conceptual/semantic systems. If artifact representations in VOTC are multi-modal, can they be considered to be conceptual/semantic in nature, whereas the modality-specific representations of animate things do not reflect directly conceptual/semantic properties? Although there are hints for this distinction, a systematic analysis is lacking.

References

1. Levy, I. *et al.* (2001) Center–periphery organization of human object areas. *Nat. Neurosci.* 4, 533–539

2. Arcaro, M.J. *et al.* (2009) Retinotopic organization of human ventral visual cortex. *J. Neurosci.* 29, 10638–10652

3. Nasr, S. *et al.* (2014) Thinking outside the box: rectilinear shapes selectively activate scene-selective cortex. *J. Neurosci.* 34, 6721–6735

4. Rogers, T.T. *et al.* (2005) Fusiform activation to animals is driven by the process, not the stimulus. *J. Cogn. Neurosci.* 17, 434–445

5. Gauthier, I. *et al.* (1999) Activation of the middle fusiform 'face area' increases with expertise in recognizing novel objects. *Nat. Neurosci.* 2, 568–573

6. Hasson, U. *et al.* (2002) Eccentricity bias as an organizing principle for human high-order object areas. *Neuron* 34, 479–490

7. Konkle, T. and Oliva, A. (2012) A real-world size organization of object responses in occipitotemporal cortex. *Neuron* 74, 1114–1124

8. Mahon, B.Z. *et al.* (2009) Category-specific organization in the human brain does not require visual experience. *Neuron* 63, 397–405

9. Wolbers, T. *et al.* (2011) Modality-independent coding of spatial layout in the human brain. *Curr. Biol.* 21, 984–989

10. Striem-Amit, E. and Amedi, A. (2014) Visual cortex extrastriate body-selective area activation in congenitally blind people 'Seeing' by using sounds. *Curr. Biol.* 24, 687–692

11. Kitada, R. *et al.* (2014) The brain network underlying the recognition of hand gestures in the blind: the supramodal role of the extrastriate body area. *J. Neurosci.* 34, 10096–10108

12. He, C. *et al.* (2013) Selectivity for large nonmanipulable objects in scene-selective visual cortex does not require visual experience. *Neuroimage* 79, 1–9

13. Peelen, M.V. *et al.* (2013) Tool selectivity in left occipitotemporal cortex develops without vision. *J. Cogn. Neurosci.* 25, 1225–1234

14. Ricciardi, E. *et al.* (2013) Mind the blind brain to understand the sighted one! Is there a supramodal cortical functional architecture?. *Neurosci. Biobehav. Rev.* 41, 64–77

15. Fairhall, S.L. *et al.* (2014) Person- and place-selective neural substrates for entity-specific semantic access. *Cereb. Cortex* 24, 1687–1696

16. Adam, R. and Noppeney, U. (2010) Prior auditory information shapes visual category-selectivity in ventral occipito-temporal cortex. *Neuroimage* 52, 1592–1602

17. Ricciardi, E. and Pietrini, P. (2011) New light from the dark: what blindness can teach us about brain function. *Curr. Opin. Neurol.* 24, 357–363

18. Ishai, A. *et al.* (2000) Distributed neural systems for the generation of visual images. *Neuron* 28, 979–990

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

20. Ishai, A. (2010) Seeing faces and objects with the 'mind's eye'. *Arch. Ital. Biol.* 148, 1–9

21. Konkle, T. and Caramazza, A. (2013) Tripartite organization of the ventral stream by animacy and object size. *J. Neurosci.* 33, 10235–10242

22. Chao, L.L. *et al.* (1999) Attribute-based neural substrates in temporal cortex for perceiving and knowing about objects. *Nat. Neurosci.* 2, 913–919

23. Martin, A. (2007) The representation of object concepts in the brain. *Annu. Rev. Psychol.* 58, 25–45

24. Kanwisher, N. (2010) Functional specificity in the human brain: a window into the functional architecture of the mind. *Proc. Natl. Acad. Sci. U.S.A.* 107, 11163–11170

25. Kriegeskorte, N. *et al.* (2008) Matching categorical object representations in inferior temporal cortex of man and monkey. *Neuron* 60, 1126–1141

26. Epstein, R. and Kanwisher, N. (1998) A cortical representation of the local visual environment. *Nature* 392, 598–601

27. Epstein, R. (2005) The cortical basis of visual scene processing. *Vis. Cogn.* 12, 954–978

28. Fairhall, S.L. and Caramazza, A. (2013) Category-selective neural substrates for person- and place-related concepts. *Cortex* 49, 2748–2757

29. Grill-Spector, K. and Weiner, K.S. (2014) The functional architecture of the ventral temporal cortex and its role in categorization. *Nat. Rev. Neurosci.* 15, 536–548

30. Kitada, R. *et al.* (2009) Functional specialization and convergence in the occipito-temporal cortex supporting haptic and visual identification of human faces and body parts: an fMRI study. *J. Cogn. Neurosci.* 21, 2027–2045

31. Goyal, M.S. *et al.* (2006) Tactile perception recruits functionally related visual areas in the late-blind. *Neuroreport* 17, 1381–1384

32. Kitada, R. *et al.* (2010) Brain networks involved in haptic and visual identification of facial expressions of emotion: an fMRI study. *Neuroimage* 49, 1677–1689

33. Kitada, R. *et al.* (2013) Early visual experience and the recognition of basic facial expressions: involvement of the middle temporal and inferior frontal gyri during haptic identification by the early blind. *Front. Hum. Neurosci.* 7, 7

34. Kilgour, A.R. *et al.* (2005) Haptic face identification activates ventral occipital and temporal areas: an fMRI study. *Brain Cogn.* 59, 246–257

35. Devlin, J.T. *et al.* (2005) Category-related activation for written words in the posterior fusiform is task specific. *Neuropsychologia* 43, 69–74

36. Devlin, J.T. *et al.* (2002) Anatomic constraints on cognitive theories of category specificity. *Neuroimage* 15, 675–685

37. Price, C.J. *et al.* (2003) How is the fusiform gyrus related to category-specificity? *Cogn. Neuropsychol.* 20, 561–574

38. Devlin, J.T. *et al.* (2002) Is there an anatomical basis for category-specificity? Semantic memory studies in PET and fMRI. *Neuropsychologia* 40, 54–75

39. Rogers, T.T. *et al.* (2006) Anterior temporal cortex and semantic memory: reconciling findings from neuropsychology and functional imaging. *Cogn. Affect. Behav. Neurosci.* 6, 201–213

40. Lambert, S. *et al.* (2004) Blindness and brain plasticity: contribution of mental imagery? An fMRI study. *Cogn. Brain Res.* 20, 1–11

41. Lambert, S. *et al.* (2002) Neural substrates of animal mental imagery: calcarine sulcus and dorsal pathway involvement – an fMRI study. *Brain Res.* 924, 176–183

42. Doehrmann, O. *et al.* (2008) Probing category selectivity for environmental sounds in the human auditory brain. *Neuropsychologia* 46, 2776–2786

43. Engel, L.R. *et al.* (2009) Different categories of living and non-living sound-sources activate distinct cortical networks. *Neuroimage* 47, 1778–1791

44. Lewis, J.W. *et al.* (2005) Distinct cortical pathways for processing tool versus animal sounds. *J. Neurosci.* 25, 5148–5158

45. Tranel, D. *et al.* (2005) Naming the same entities from visual or from auditory stimulation engages similar regions of left inferotemporal cortices. *J. Cogn. Neurosci.* 17, 1293–1305

46. Bracci, S. and Peelen, M.V. (2013) Body and object effectors: the organization of object representations in high-level visual cortex reflects body–object interactions. *J. Neurosci.* 33, 18247–18258

47. Bracci, S. *et al.* (2012) Closely overlapping responses to tools and hands in left lateral occipitotemporal cortex. *J. Neurophysiol.* 107, 1443–1456

48. Noppeney, U. *et al.* (2006) Two distinct neural mechanisms for category-selective responses. *Cereb. Cortex* 16, 437–445

49. Costantini, M. *et al.* (2011) Haptic perception and body representation in lateral and medial occipito-temporal cortices. *Neuropsychologia* 49, 821–829

50. Le Clec'H, G. *et al.* (2000) Distinct cortical areas for names of numbers and body parts independent of language and input modality. *Neuroimage* 12, 381–391

51. Wang, X. *et al.* (2015) How visual is the visual cortex? Comparing connectional and functional fingerprints between congenitally blind and sighted individuals. *J. Neurosci.* 35, 12545–12559

52. Mullally, S.L. and Maguire, E.A. (2011) A new role for the parahippocampal cortex in representing space. *J. Neurosci.* 31, 7441–7449

53. Troiani, V. *et al.* (2014) Multiple object properties drive scene-selective regions. *Cereb. Cortex* 24, 883–897

54. Mahon, B.Z. and Caramazza, A. (2011) What drives the organization of object knowledge in the brain? *Trends Cogn. Sci.* 15, 97–103

55. Bracci, S. *et al.* (2015) Representational similarity of body parts in human occipitotemporal cortex. *J. Neurosci.* 35, 12977–12985

56. Downing, P.E. *et al.* (2001) A cortical area selective for visual processing of the human body. *Science* 293, 2470–2473

57. Taylor, J.C. *et al.* (2007) Functional MRI analysis of body and body part representations in the extrastriate and fusiform body areas. *J. Neurophysiol.* 98, 1626–1633

58. Peelen, M.V. and Downing, P.E. (2007) The neural basis of visual body perception. *Nat. Rev. Neurosci.* 8, 636–648

59. Peelen, M.V. and Downing, P.E. (2005) Selectivity for the human body in the fusiform gyrus. *J. Neurophysiol.* 93, 603–608

60. Schwarzlose, R.F. *et al.* (2005) Separate face and body selectivity on the fusiform gyrus. *J. Neurosci.* 25, 11055–11059

61. Srihasam, K. *et al.* (2014) Novel domain formation reveals proto-architecture in inferotemporal cortex. *Nat. Neurosci.* 17, 1776–1783

62. Saygin, Z.M. *et al.* (2012) Anatomical connectivity patterns predict face selectivity in the fusiform gyrus. *Nat. Neurosci.* 15, 321–327

63. Striem-Amit, E. *et al.* (2012) Reading with sounds: sensory substitution selectively activates the visual word form area in the blind. *Neuron* 76, 640–652

64. Gomez, J. *et al.* (2015) Functionally defined white matter reveals segregated pathways in human ventral temporal cortex associated with category-specific processing. *Neuron* 85, 216–228

65. Hannagan, T. *et al.* (2015) Origins of the specialization for letters and numbers in ventral occipitotemporal cortex. *Trends Cogn. Sci.* 19, 374–382

66. Warrington, E.K. and Shallice, T. (1984) Category specific semantic impairments. *Brain* 107, 829–853

67. Warrington, E.K. and McCarthy, R. (1983) Category specific access dysphasia. *Brain* 106, 859–878

68. Cree, G.S. and McRae, K. (2003) Analyzing the factors underlying the structure and computation of the meaning of chipmunk, cherry, chisel, cheese, and cello (and many other such concrete nouns). *J. Exp. Psychol. Gen.* 132, 163–201

69. Farah, M.J. and McClelland, J.L. (1991) A computational model of semantic memory impairment: modality specificity and emergent category specificity. *J. Exp. Psychol. Gen.* 120, 339–357

70. Caramazza, A. and Shelton, J.R. (1998) Domain-specific knowledge systems in the brain the animate–inanimate distinction. *J. Cogn. Neurosci.* 10, 1–34

71. Caramazza, A. and Mahon, B.Z. (2003) The organization of conceptual knowledge: the evidence from category-specific semantic deficits. *Trends Cogn. Sci.* 7, 354–361

72. Weiner, K.S. *et al.* (2014) The mid-fusiform sulcus: a landmark identifying both cytoarchitectonic and functional divisions of human ventral temporal cortex. *Neuroimage* 84, 453–465

73. Caspers, J. *et al.* (2013) Cytoarchitectonical analysis and probabilistic mapping of two extrastriate areas of the human posterior fusiform gyrus. *Brain Struct. Funct.* 218, 511–526

74. Glasser, M.F. and Van Essen, D.C. (2011) Mapping human cortical areas in vivo based on myelin content as revealed by T1- and T2-weighted MRI. *J. Neurosci.* 31, 11597–11616

75. Van Essen, D.C. *et al.* (2013) The WU-Minn Human Connectome Project: an overview. *Neuroimage* 80, 62–79

76. Caspers, J. *et al.* (2015) Receptor architecture of visual areas in the face and word-form recognition region of the posterior fusiform gyrus. *Brain Struct. Funct.* 220, 205–219

77. Osher, D.E. *et al.* (2015) Structural connectivity fingerprints predict cortical selectivity for multiple visual categories across cortex. *Cereb. Cortex* Published online January 26, 2015. http://dx.doi.org/10.1093/cercor/bhu303