

On the relationship between maps and domains in inferotemporal cortex

Michael J. Arcaro and Margaret S. Livingstone 

Abstract | How does the brain encode information about the environment? Decades of research have led to the pervasive notion that the object-processing pathway in primate cortex consists of multiple areas that are each specialized to process different object categories (such as faces, bodies, hands, non-face objects and scenes). The anatomical consistency and modularity of these regions have been interpreted as evidence that these regions are innately specialized. Here, we propose that ventral-stream modules do not represent clusters of circuits that each evolved to process some specific object category particularly important for survival, but instead reflect the effects of experience on a domain-general architecture that evolved to be able to adapt, within a lifetime, to its particular environment. Furthermore, we propose that the mechanisms underlying the development of domains are both evolutionarily old and universal across cortex. Topographic maps are fundamental, governing the development of specializations across systems, providing a framework for brain organization.

The primate ventral visual stream is responsible for our ability to recognize objects. Lesions to this part of the brain can impair object recognition, without affecting the ability to locate objects^{1,2}. Lesions in human ventral visual cortex can result in surprisingly specific deficits in object recognition, such as impairments in the ability to recognize faces, body parts, tools, text or places, without affecting the ability to recognize other object classes^{3,4}. Such specific deficits have led to the idea that different object categories must be represented by anatomically distinct parts of the ventral stream; such clustering of category selectivity has been extensively confirmed using functional MRI (fMRI) in both humans and monkeys^{5–7}. Electrophysiological recordings from monkeys and from humans have revealed that neurons in these regions are indeed selectively responsive to particular object categories^{8,9}.

The processing done by this part of our brain is remarkable: we can immediately recognize a particular familiar face, even though faces differ from each other only in subtle ways and despite the tremendous variation in the activity that any individual

face may produce on our retina. The fluent reading you are currently performing is also an astonishing feat. Despite decades of research, fundamental questions remain regarding how we process faces and more generally recognize objects. Is the clustering into category-specific domains key to these abilities? Does having specific domains for particular categories indicate that our brains evolved specialized circuitry to recognize biologically important object categories such as conspecifics and suitable environments? Or do these specialized domains reflect how we learn to recognize the things we encounter? If these domains are innate, how could such remarkably specific selectivities get ‘wired up’? If their connections are determined by experience, why do virtually all humans have domains for different object categories in stereotyped neuroanatomical locations^{10,11}?

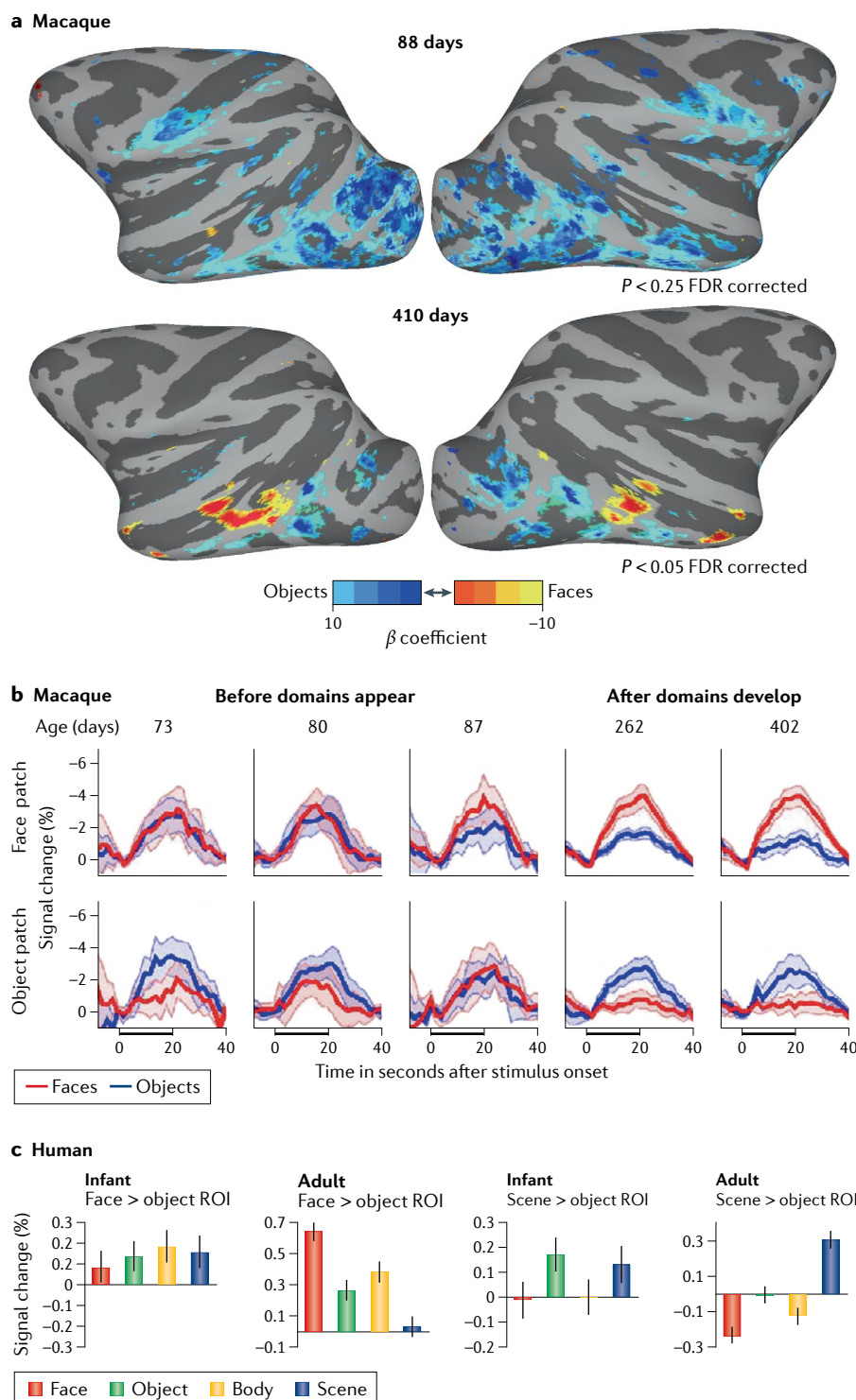
We previously addressed these questions by reviewing the development of face selectivity in inferotemporal cortex (IT)¹²; here, we back farther out and look for still more general principles to explain the localizations of these specializations, their correlations with other topographies and some recently reported cross-modalities.

Asking how such selective circuits get wired up eschews the tempting teleological trap of accepting such circuitry as being ‘for’ what it does. That is, instead of assuming that domains in the ventral stream are ‘for’ what they do, in this Perspective we ask what mechanisms during development could cause different parts to become selective for some highly specific object categories and not others.

Ventral stream development

Innatist versus bottom-up models. There is a long history of attempts to determine whether particular behaviours are innate or learned, starting at least as early as the seventh century BC with the pharaoh Psammetichus I, who had two children raised by a shepherd who was forbidden to speak to them. The goal was to find out what language the children would speak spontaneously. There is an interesting assumption in this experiment: that, even though babies do not speak, language is nevertheless innate and a child needs only to mature for this innate ability to manifest itself. Although most people would now agree that any particular language must be learned, Chomsky influentially proposed that humans do possess innately wired brain circuits that underlie our language ability¹³. It has been similarly influentially argued that the category-selective domains, especially the face domains, are innate, and have evolved to support the uniquely social behaviour of humans and other primates¹⁴; we refer to this as the ‘innatist model’ or ‘top-down model’. We argue here that the burden of proof that lies with this high-level, anthropocentric view is not met. Instead, we aim to show that universal mechanisms of development, common to many species, can account for domain-specific brain regions; we refer to this as the ‘bottom-up model’.

The strongest arguments for the innatist model for ventral-stream domains are that they are found in stereotyped locations in both humans and monkeys and that domains exist for biologically important object categories, such as faces, bodies and places. But humans also have domains for tools and for text, and it is implausible that we evolved a domain for text, given how recently in our evolutionary history literacy has been prevalent. It has been argued that



the text domain (the ‘visual word form area’ (VWFA)) represents recycling of a domain that was previously ‘for’ something else¹⁵. But does it necessarily follow that some high-level visual area is ‘for’ some function, just because in most humans it serves that function? If these high-level visual areas were produced by evolutionary selective pressures to recognize specific object categories, then they should not be co-opted efficiently by some other kind of stimulus, such as text, without retaining selectivity for the stimulus category they evolved to process.

Domain development. Here, we summarize what little is known about the early development of visual-category domains. First, category-selective domains develop in consistent parts of cortex across individuals, but in different regions across species. The object-recognition pathway in monkeys, IT¹⁶, comprises cortex within and ventral to the superior temporal sulcus, areas TEO and TE. In humans, the object-recognition pathway, often referred to as ‘ventral occipital temporal cortex’ (VOT)¹⁷, comprises cortex within and around the fusiform, collateral and lateral occipital sulci. The localization of the ventral visual stream to different sulcal folds across species already indicates that the development of object-selective cortex is not under such rigid constraints as those that do result in localization of functional areas to consistent anatomical markers across primates, such as primary visual cortex to within and along the calcarine sulcus. Second, there is some direct evidence on the development of domains: adult-like face selectivity as measured using fMRI is not present at birth in either monkeys or humans^{18,19} but develops over the first few months after birth, a time rich in face experience for the neonate (FIG. 1).

In humans, the VWFA does not develop until school age and does not develop at all unless an individual has learned to read^{20,21}. Thus, development of a text domain requires text experience. It has been proposed that the stereotyped localization of the VWFA to the left occipitotemporal sulcus is due to its selective connectivity with language areas^{22,23}. But monkeys, which do not have a language area, can develop domains that respond selectively to human symbols if the monkeys are intensively trained as juveniles to recognize symbols^{22–25}; furthermore, symbol use allows these monkeys increased accuracy in judging quantity²⁶. This result suggests that intensive experience alone is sufficient

Fig. 1 | Development of face selectivity in macaque and human infants. Lack of face versus non-face object selectivity in macaque¹⁸ and human¹⁹ infants. **a** | Before approximately 200 days old, macaques do not show face > non-face-object selective regions, and after this age, face selectivity appears and is stable¹⁴⁸, as measured by functional MRI. **b** | Cerebral blood volume signal responses reveal that, before face and non-face-object domains become detectable, monkey inferotemporal cortex is responsive to visual stimuli, but not selective to image category¹⁴⁸. **c** | Regions that are face selective in adult human ventral occipital temporal cortex are not selective for faces over non-face objects in 4–6-month-old human infants, as reflected by percentage differences in blood oxygen level-dependent responses¹⁴. Some selectivity for scenes versus faces was observed in both these studies, but could reflect differences in visual-field stimulation (for example, centre versus periphery) by the stimuli used. FDR, false discovery rate; ROI, region of interest. Parts **a** and **b** adapted from REF.¹⁴⁸, CC BY 4.0 (<https://creativecommons.org/licenses/by/4.0/>). Part **c** adapted from REF.¹⁹, CC BY 4.0 (<https://creativecommons.org/licenses/by/4.0/>).

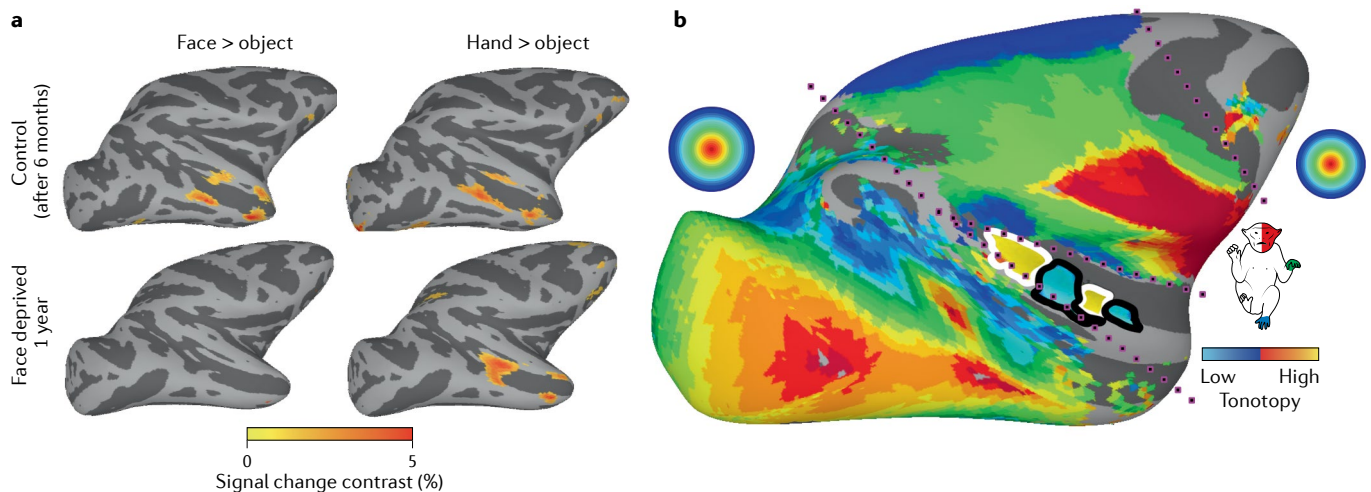


Fig. 2 | Intrinsic and experience-dependent organization in macaque visual cortex. a | Lack of face domains (left) but normal hand domains (right) in face-deprived monkeys (bottom row) as compared with controls (top row)¹⁸. Images show contrasts as determined with functional MRI. **b** | At birth, most of the cortex is made up of maps of the sensory periphery¹². This is a composite image illustrating the coverage of visual²⁷, somatosensory–motor²⁸ and auditory³⁴ maps. Maps of eccentricity in visual space cover occipital, temporal, posterior parietal and frontal eye fields. Maps of the body (face, hands and feet) cover areas within and around the central sulcus. Alternating representations of high (yellow, with white outline) and low (cyan, with black outline) tonotopic frequencies cover parts of the superior temporal gyrus. Part **a** adapted from REF.¹⁸, Springer Nature Limited. Part **b** adapted with permission from REF.²⁸, PNAS, and from REF.³⁴, Springer Nature Limited, using data from REFS^{27,31}.

to produce a category-selective domain that may facilitate category expertise. Conversely, monkeys who are raised for the first year of life without seeing any kind of faces do not develop domains selective for faces over non-face objects, bodies or scenes (although their other domains are normal), indicating that experience is necessary for developing domains¹⁸ (FIG. 2a). The regions in IT of face-deprived monkeys that would be face selective in a control monkey respond more strongly to hands than to faces¹⁸. Thus, monkey studies indicate that extensive early experience of an object category is both necessary and sufficient for developing that category-selective domain, and that, at birth, IT lacks the specialized domain organization typically found in adults.

Proto-architecture

What is present at birth in both human and monkey cortex is topographic maps^{27,28}. Virtually the entire visual system in newborn macaques consists of a series of retinotopic maps, with neighbouring neurons receiving input from neighbouring parts of the retina and therefore adjacent parts of visual space (FIG. 2b). Even higher visual areas in both monkeys and humans that were once considered non-retinotopic, such as human lateral occipital and ventral occipital cortex, and ventral temporal cortex in monkeys, have since been shown to be retinotopically organized in adults^{29–33} as well as in newborn monkeys²⁷. More anteriorly, both somatosensory and motor areas are also already organized at birth into multiple

body maps²⁸. It is likely that in between the multiple visual-field maps and body maps are, as in adults, multiple auditory maps of frequency^{34,35}. Thus, at birth, most of the neocortex consists of orderly spatial maps of the sensory and motor periphery, and what is innate is this machinery to systematically sample and respond to the environment.

In adult humans and macaques, face domains are retinotopically biased to central visual field (that is, face-selective neurons tend to have receptive fields that are located near fixation), and neuronal receptive fields in building and scene domains tend to be located in peripheral parts of the visual map^{36,37}. Furthermore, there is a centre-to-peripheral progression of processing text, faces, tools and scenes³⁸, as well as a gradient in processing objects according to real-world size³⁹, indicating a general retinotopic organizing principle for the localization of different category domains. Both top-down and bottom-up explanations have been offered for this retinotopy–category association^{36,40}. The top-down account posits that this association arises from the specific requirements for the recognition of different object categories; for example, face processing requires central scrutiny, and scene processing requires integration across the visual field^{36,37,40}. Thus, (innately determined) face domains would selectively recruit or retain central-visual-field inputs from antecedent areas, and (innately predetermined) scene domains recruit or retain peripheral inputs. By contrast, the bottom-up model proposes

that, during development, category domain formation and localization are governed by pre-existing retinotopic organization³⁶.

In support of the bottom-up explanation for the retinotopy–category association, we found that in newborn macaques the parts of IT that are destined to become face selective are already centrally biased, and that the parts that are destined to become place domains are already peripherally biased^{27,41}. Thus, retinotopic maps precede category selectivity during development. The innatist interpretation of this fact is that proto-face domains start out with a central-field bias in order to facilitate face processing⁴¹. However, the central field is part of a larger-scale topographic map of visual space that spans object-selective cortex (FIG. 2b). Therefore, the development of the central-visual-field representation must be intrinsically tied to the development of the entire map in IT. It therefore seems implausible that face-domain localization could be predetermined independently of retinotopy as, if this were the case, the entire retinotopic map would have to conform to face-domain localization. More generally, finer-scale architecture must develop after, or at least not before, larger-scale architecture; it cannot be the reverse. That is, the details of a map can fill in only after the layout is established. The bottom-up interpretation of the fact that retinotopy precedes category selectivity would be that the maps, being present at birth, are primary, and there must be something about central vision at that level of the visual cortical

hierarchy that facilitates the emergence of face selectivity there.

Retinotopy carries with it an organization for low-level and mid-level shape features, because acuity varies dramatically with eccentricity, as does receptive-field size. Central-visual-field neurons have tiny receptive fields, whereas peripheral-field neurons have huge fields⁴². Many cells in both cat and monkey visual cortex respond to short contours more strongly than to long contours^{43,44}, and such ‘end-stopped’ cells are prevalent in monkey V2 (REF.⁴⁵) and respond more strongly to curved contours and corners than to extended contours^{43,45}. Therefore, the optimum degree of curvature will scale with eccentricity, with central-visual-field regions preferring high curvature, and peripheral-visual-field regions preferring straighter contours.

Mistaking an elephant for a face

Face domains respond better to round curvy things, such as balls and clocks, and scene domains to rectilinear things⁴⁶. The innatist explanation for an association between category and curvature would be that the selectivity of face domains extends to curvy things because faces have a lot of curvy contours, and so curvy things appear similar to faces, and anything straight looks more like it could be part of a scene than a face. The bottom-up explanation would be the reverse: that, rather than category domains driving similar-shape localization, a systematic variation in shape preference may guide the stereotyped localization of different category domains. Furthermore, when fMRI is used to map responsiveness to curvy things versus straight things, not only does a ‘curvy > straight’ bias characterize

face domains, it also distinguishes central-visual-field regions from peripheral-visual-field regions. This central–peripheral, curvy–straight correlation is present not just in category-selective cortex but also throughout the visual system, including primary visual cortex (V1), in monkeys²⁵ (FIG. 3a) and humans⁴⁷ (FIG. 3b). The fact that both these studies found such a strong correlation between curvature and eccentricity means that whatever causes neurons to prefer curved contours (which we propose is end-stopping) must be prevalent. The association between curvature and eccentricity extends beyond just category-selective regions into early visual areas. By the logic that a large-scale architecture must dominate a finer-scale architecture, category selectivity cannot be the driving force for curvature

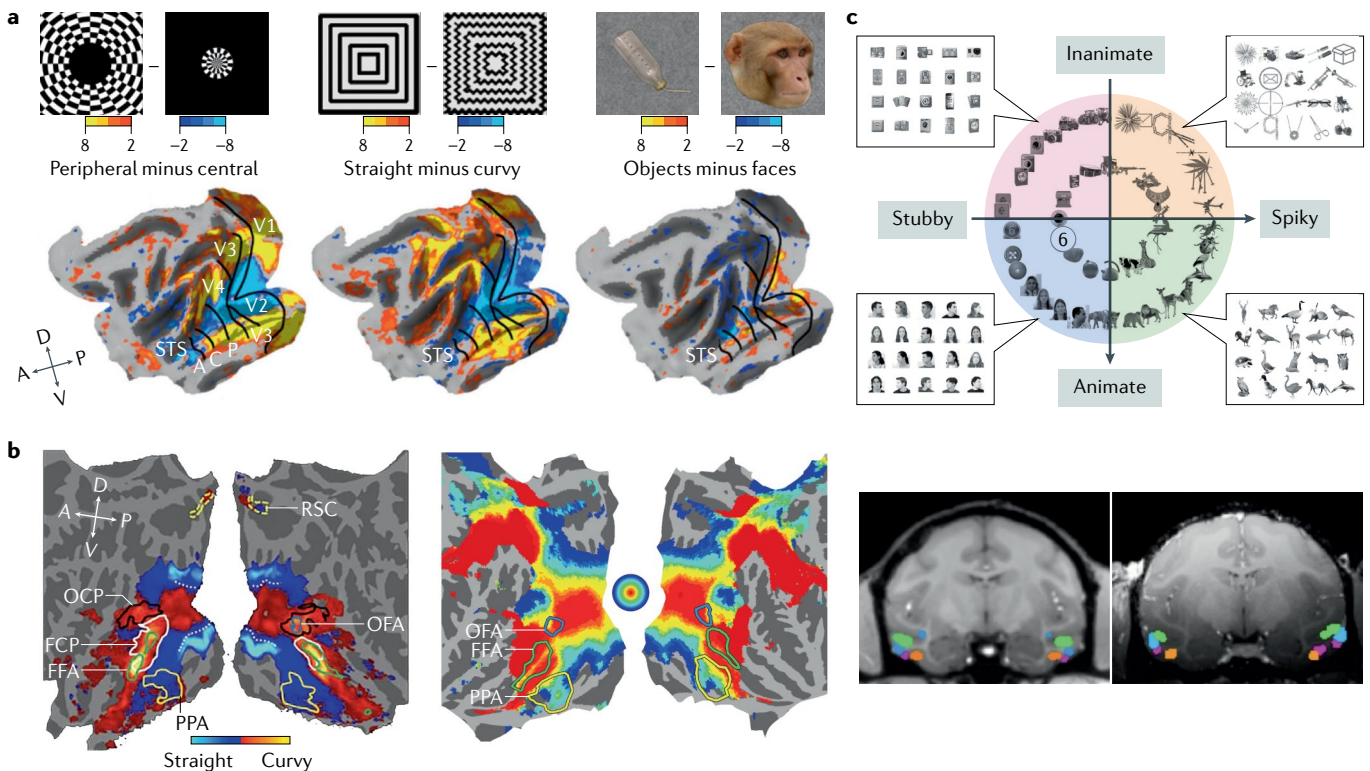


Fig. 3 | Topographic receptive-field tuning. a | Curvature correlates with eccentricity throughout the macaque visual system²⁵. Maps of responses to peripheral minus central-visual-field stimuli (left), straight minus curvy stimuli (centre) and non-face objects minus faces (right), with examples of each stimulus type above each map²⁵. A, C, and P indicate anterior, central and posterior subdivisions of inferotemporal cortex, respectively. **b** | Curvature also correlates with eccentricity in the human visual system⁴⁷. The dashed white line represents the border between central and peripheral visual field in early visual cortex. The cyan–yellow scale correlates with curvilinear values of visual stimuli, such that the areas in red–yellow process curvy features and those in blue–cyan process rectilinear features. The locations of the fusiform face area (FFA; green outline), occipital face area (OFA; blue outline), occipital curvature preference patch (OCP; black outline) and fusiform curvature preference patch (FCP; white outline) are

within areas that preferentially respond to curved features. By contrast, the location of the parahippocampal place area (PPA) is encompassed by the region responding preferentially to rectilinear features. The probabilistic locations of face domains (the OFA and FFA) and the scene domain (the PPA)¹⁴⁹ are superimposed on a human eccentricity map¹⁵⁰ (right). **c** | Gradient of shape selectivity in macaque inferotemporal cortex⁴⁸ that could reflect developmental origins in eccentricity-based low-level shape selectivity gradients. Colours show the correspondence between clustering in image shape space (top) and anatomical space (bottom). A, anterior; D, dorsal; P, posterior; RSC, retrosplenial cortex; STS, superior temporal sulcus; V, ventral; V1, primary visual cortex. Part **a** adapted from REF.²⁵, Springer Nature Limited. Part **b** adapted with permission from REF.⁴⁷, Elsevier (left) and generated using data from REFS^{149,150}. Part **c** adapted from REF.⁴⁸, Springer Nature Limited.

tuning and retinotopy being associated with particular parts of IT; instead, the causality must be the reverse.

But we probably should not think of this eccentricity-derived shape selectivity as just end-stopping, curvature or retinotopy, because at every stage in the visual hierarchy inputs are combined to generate more complex and more abstract receptive-field properties. Such a combinatorial mechanism might result in something like the recently described gradient of shape selectivity in monkey IT⁴⁸ (FIG. 3c), with selectivity for ‘stubby’ things on the lip of the superior temporal sulcus (which maps to central visual field) grading to increasingly spiky things going more peripherally along the gyrus. The stubby-shape regions prefer images with a lot of curved contours (including faces), and spiky domains prefer images containing a lot of straight lines, sticking out. Thus, face and body domains, often postulated as spatially discrete regions, may reflect parts of a larger-scale map. The proportion of face-selective neurons is highest at the centre of a face patch and falls off at the patch edges^{49,50}, consistent with the notion that activity-dependent sorting mechanisms cause neural selectivity to change smoothly along the cortical surface^{51,52}. This parallels the smooth transitions in orientation and retinotopy in early visual cortex⁵³. Indeed, the apparent spatial discreteness of category-selective domains may be a manifestation of the contrasts and thresholds typically used to visualize category-selective regions (see, for example, Fig. 16 in REF.¹²), akin to the limited perspectives of the allegorical blind men interpreting different parts of an elephant. In anatomical tracer studies in IT⁵⁴ and early visual cortex⁵⁵, labelling appears ‘patchy’ (Supplementary Fig. 1): that is, labelled neurons in IT or VOT connected to regions into which a tracer is injected form discrete clusters. Knowing the underlying map organization of early visual cortex makes it clear that these patchy connections reflect connectivity between the same parts of the visual field across maps. We suggest that, rather than demonstrating connectivity between discrete domains, the connections between IT face patches may similarly reflect connections between corresponding parts of maps. The historically minded reader may recall a similar instance when the presence of colour-selective blobs in V1 was originally missed by Hubel and Wiesel and later discovered by Livingstone and Hubel once there were anatomical markers to distinguish them⁵⁶.

We mentioned earlier that one fact that favours the innatist model is that face, body

and scene domains are always found in stereotyped anatomical locations in both humans and monkeys. However, these category domains are distributed along a dorsoventral–mediolateral anatomical trajectory that correlates with an eccentricity gradient, with faces represented more centrally than bodies and scenes represented more peripherally, in both humans and monkeys^{25,36–38,57}. Categories distinguished by animacy and real-world size are also distributed along this same dorsoventral–mediolateral gradient in temporal cortex, with small, animate things represented in central-visual-field parts of the map and large, inanimate things represented peripherally^{39,58}. Recently, these gradients were shown to be accounted for by low-level shape selectivity, rather than lexical category⁵⁹. Together, these studies illustrate an intrinsic link between visual categories and a map of visual space that carries with it biases for scale and shape features such as curvature.

However, although some studies find that category selectivity can be explained by low-level features^{48,59,60}, many have reported that the degree of category selectivity in the ventral visual stream, especially in more anterior areas, cannot be explained by low-level or mid-level features alone^{61–66}. We suggest that this is because postnatal activity-dependent plasticity further sculpts selectivity depending on what is experienced — by shifting and/or narrowing the tuning of neurons towards things that are regularly experienced; that is, experience sculpts a low-level eccentricity-based shape map into a higher-level organization of behaviourally meaningful categories. A gradient from shape-biased selectivity posteriorly to category-biased selectivity more anteriorly would be consistent with later and stronger experience-dependent plasticity going up the visual hierarchy, as has been suggested by the stronger effects of early strabismus (misalignment of the eyes) on areas beyond V1 (REFS^{67,68}). A retinotopic map automatically carries a low-level shape map, which could bias the central-visual-field part of monkey IT or human VOT to become selective to faces, which are made up of many curvy contours. For example, neurons in regions where face patches eventually develop may start out broadly responsive to concentric features, which are characteristic of faces, but also clocks, cookies and doughnuts. Over development, their tuning may refine to respond preferentially to faces, even particular faces, and less to other less frequently experienced or behaviourally important objects with similar features.

Notably, even in adult monkeys, neurons in face patches respond to round things such as balls, cookies and clocks⁶⁹. The presence of topographic maps selective for image scale and curvature in newborn monkeys²⁷ provides further evidence that low-level and mid-level shape biases in IT provide the building blocks for domain development. The refinement of tuning to heavily experienced features (such as faces) throughout development would not break the underlying topographic shape map⁴⁸, but could result in expansion of the parts of the map selective for these particular features, resulting in an apparent growth of selectivity for some visual categories compared with others^{70,71}.

If the visual hierarchy starts out as a series of maps, this means that the brain is wired up at birth to systematically sample the visual environment. To what extent does experience of the environment modify this topography? Certainly postnatal activity is known to affect the circuitry of primary visual, auditory and somatosensory cortices: if one eye is inactive or less active during postnatal development, inputs from the seeing eye come to dominate visual cortex, and inputs from the inactive eye are permanently lost^{72,73}. Similarly, if visual input is filtered such that only a narrow range of orientations is experienced, cortex comes to over-represent the seen orientation^{74–77}. If sounds of a particular frequency range are excessively experienced during postnatal development, that range of frequencies comes to dominate auditory cortex^{78,79}. If a subset of whiskers on a rodent’s muzzle are cut during postnatal development, characteristic cytoarchitectonically distinct ‘barrel fields’ of somatosensory cortex representing those whiskers shrink and are taken over by adjacent whisker barrels⁸⁰. Thus, there is ample evidence that the early postnatal brain does modify its circuitry, at least in primary sensory cortices, to respond preferentially to experienced stimuli, constrained and guided by map topography. Given merely the fact that higher-level sensory areas receive their input from primary sensory areas, it follows that higher sensory areas should also come to be dominated by experienced stimuli. Indeed, there is some evidence that the same kinds of activity-dependent plasticity also hold in cortical regions beyond primary sensory areas^{67,68}. Therefore, IT or VOT is expected to become responsive to the most commonly encountered things in its environment.

Given the importance of activity-dependent plasticity, how young animals look at and experience their visual

world could also bias where on this map category domains develop. From birth, the visual experience of human infants is heavily biased towards seeing faces⁸¹. Both human and monkey infants preferentially look at faces more than other shapes soon after birth^{82–86}. This is often given as evidence that face domains must be innate, but there is no evidence that cortical face domains are more involved in early face-looking behaviour than other cortical or even subcortical structures^{12,83}, and infant monkeys and humans look preferentially at faces before face domains are detectable.

In most infant face-looking studies, schematized faces are compared with scrambled faces, and seldom compared with objects of comparable spatial frequency composition. Indeed, faces turn out to have spatial frequency distributions that are most visible to neonates^{87,88}. Furthermore, small dark features in the upper half of the visual field may also be optimal stimuli for driving infant looking behaviour⁸⁹. However, identifying the intrinsic biases that promote face-looking behaviour is complicated by early reinforcement and learning by the infant. Preferential looking to the parental face increases over the first day after birth, indicating that even a few hours of experience has a behaviourally measurable effect on an infant's face-looking behaviour⁹⁰. Thus, it is far from proven that infants look at faces because they have an innate predisposition to look at faces, rather than because the salient and dynamic features of faces fit the parameters of infant vision better than most other object categories.

One evolutionarily old structure that could support early looking behaviour towards top-heavy, high-contrast moving features is the superior colliculus. Given that the superior colliculus is directly involved in looking behaviour⁹¹, has an upper-field bias⁹² and is relatively mature at birth⁹³, this subcortical structure may be the reason why infants look at faces. Further support for the idea that face looking is not innate is the observation that normal monkeys look preferentially at faces, whereas face-deprived monkeys do not¹⁸. Note that an earlier study by Sugita reported that face-deprived monkeys look at faces more than at non-face objects, but the difference was reported to be not statistically significant⁹⁴. By contrast, Sugita did find a significant effect of postdeprivation experience on behaviour. Altogether, these data indicate that early face-looking behaviour is unlikely to be driven by cortical domains and therefore should not be taken as evidence for innate cortical face networks. However, how

infants look at different things in their environment probably does constrain where on the cortical map those categories become localized, and so looking at faces would bias face-domain formation to central-visual-field locations, and scenes would be biased towards more peripheral parts of the visual field.

Structures without a function?

Although the stereotyped locations of domains in IT or VOT can be explained by postnatal experience acting on a domain-general architecture, several recent studies have reported distinct laminar and connectivity profiles of different domains. In humans, areas containing face and place domains were recently reported to show distinct cytoarchitectonics (that is, different patterns of cell density across the layers of cortex)¹¹, and face domains show selective interconnectivity with other face areas in both humans and monkeys^{95–97}. Furthermore, large-scale connectivity patterns in humans are predictive of the location of face domains⁹⁸, as well as of where the VWFA will appear in children⁹⁹. Are domains thus distinct cortical areas, with characteristic cytoarchitectures, cell-type distributions and intrinsic and extrinsic connectivities?

The anatomical differences between individual domains in IT or VOT might seem to support the idea that they are genetically predetermined. Indeed, face patches in macaques correspond to particular folds along the superior temporal sulcus that form in utero and have distinct laminar organization¹⁰⁰. However, these anatomical features are also present in face-deprived monkeys that lack face patches¹⁰⁰, demonstrating that anatomical features are insufficient to prove the presence of functional specializations and, instead, that the anatomy may correspond to features of the intrinsic topographic architecture, which also correlate with later-emerging domains. Furthermore, although differences in connectivity and circuitry may determine how ventral stream domains perform their highly selective computations, they may not be the reason why a particular domain performs a specific computation; indeed, causality may be the reverse. That is, the reason face patches become face patches may not be because they have predetermined face-specific cytoarchitecture and connectivity; rather, the distinct cytoarchitecture and connectivity of face patches may arise, or refine, because of how these regions are stimulated during development. To unwrap this argument,

we need to look at cortical development in general.

Beyond just IT or VOT, different regions of cortex have long been distinguished on the basis of cytoarchitectonics¹⁰¹. These cytoarchitectonic divisions more often than not are also functionally distinct and have distinct intrinsic and extrinsic connectivities and different distributions of various cell types, molecular profiles and neurotransmitters. This has led to the idea that cells in different parts of the cortex are genetically programmed to express these unique features that allow them to perform computations characteristic of each area. However, individual functional maps do not align cleanly with cytoarchitectonic borders^{11,102}, and different cortical areas are more similar to each other than not: neocortical areas all share the same basic six-layered structure, with the same basic scheme of inputs to layer 4, thalamic outputs from layer 6, other subcortical projections from layer 5 and projections to other cortical areas from layers 2/3 (REFS^{103,104}). Different cortical areas all share the same basic cell types and neurotransmitters, although in different proportions, and they perform similar basic computations of input integration and gain control¹⁰⁴. There are multiple lines of evidence that cortex is multipotential (BOX 1) in the kinds of information it can process, rather than having specific circuitry for different functions.

At birth, inputs to V1 are segregated into ocular dominance domains, thought to form by activity-dependent learning rules originally proposed by Hebb¹⁰⁵ that reinforce correlated inputs and weaken uncorrelated inputs; prenatal waves of activity in the two retinas are independent, and therefore uncorrelated. At birth, V1 is also organized into orientation pinwheels, again thought to be generated by activity-dependent sorting of on and off inputs and lateral connections that encourage neurons to retain inputs similar to those of their neighbours.

Researchers used to argue about the function of ocular dominance columns and orientation columns; what they were 'for'. In a seminal study, Law and Constantine-Paton implanted a third eye on a tadpole, which grew into a three-eyed frog¹⁰⁶. When they injected a tracer into the third eye, they saw ocular dominance columns in an animal that normally never has ocular dominance columns because, in normal frogs, inputs from the two eyes project entirely contralaterally (Supplementary Fig. 2a). This implies that

ocular dominance columns (and probably orientation columns, too) are not ‘for’ anything; rather, they emerge merely as a consequence of a Hebbian clustering rule. This study reminds us that just because a brain region is correlated with a function, it does not necessarily exist because of that function. We propose that the same holds true for the stereotypical clustered architecture of IT or VOT. That is not to say that neurons in these domains are not required for these functions or do not promote a particular behaviour. Rather, the activity-dependent rules that promote clustering may also lead to experience-dependent refinement of neural tuning that supports specific perceptual processes, such as our ability to quickly and accurately recognize objects in our environment. Such simple underlying rules may be evolutionarily old optimizations for information coding in an unspecified environment. Given that neurons are more likely to correlate with, and therefore connect to, nearby neurons, clustering by category facilitates interconnectivity between neurons that code for similar things and therefore facilitates within-category comparisons, and so clustering probably optimizes processing of these categories. This may be especially true for categories such as faces that have a high degree of homogeneity across examples and would thus be naturally predisposed to recruiting similarly tuned neurons within a map of visual features. It remains an open question whether a brain could acquire similar abilities using completely distributed neural ensembles and what behavioural benefits, if any, clustering into domains might provide^{9,107}.

Notably, the spatial scale of category domains in IT or VOT is larger than the ocular dominance and orientation columns in V1 (Supplementary Fig. 2b). Can we attribute the large domains in IT to the same mechanisms that form smaller domains in V1? Nasr and Tootell mapped ocular dominance domains in human V1, and colour and disparity domains in human extrastriate cortex¹⁰⁸. The domains in V2 are coarser than in V1, and coarser in V3 than in V2, and possibly coarser still in V4 (Supplementary Fig. 2c). So, it is likely that domains just get larger going up the visual hierarchy (possibly owing to input convergence at each stage and therefore broader correlation patterns). If so, it would not be too surprising that the huge domains in the ventral stream emerge from the same activity-dependent rules that produce the mesoscale domain organization of V1.

Box 1 | Cortex is cortex

The organization of cortical areas revealed by electrophysiological recordings initially led to the hypothesis that the cortex is tiled by repeating computational units arranged in hierarchies^{151,152}. This idea of a repeating computational unit is consistent with how brains differ in size: across land mammals, brain size scales with body size, over five orders of magnitude, and this is almost entirely accounted for by changes in cortical surface area, not cortical thickness, suggesting that larger brains evolved by increasing the number of these units^{104,153}.

That cortex is pluripotent has been demonstrated directly in several ways. First, experiments in which the optic nerve was routed into somatosensory thalamus in hamsters or auditory thalamus in ferrets revealed visually driven cells in the host cortex with visual properties such as orientation and direction tuning as well as a 2D map of visual space^{154–156}, along with alterations in intrinsic connectivity¹⁵⁷ and corticocortical connectivity¹⁵⁸. Furthermore, the animals could use this misrouted pathway for visually guided behaviour¹⁵⁹. The second line of evidence that developing cortex is pluripotent and acquires its distinguishing features and its connectivity from exogenous cues comes from a series of heterologous transplants of fetal neocortex. The transplanted tissues acquire the cytoarchitectonic features and patterns of connectivity characteristic of the host location^{160,161}. Thus, the inputs, and in particular the pattern of activity in those inputs, to a cortical area determine its cytoarchitectonics, intrinsic connectivity, pattern of gene expression¹⁶² and projections to other areas.

All roads to Rome are topography

We propose that the category domains in IT arise because postnatal visual activity acts on a retinotopic proto-organization by the same kinds of activity-dependent sorting mechanisms that lead to ocular dominance and orientation domains in V1; indeed, these sorting mechanisms may be key to forming the protomaps themselves. Where on this map different categories arise is biased by a low-level shape map based on retinotopic eccentricity and scale and by how different things in the environment are most frequently viewed.

However, in the past few years, there have been several human fMRI studies on blind individuals presented with tactile or auditory stimuli that produce category-selective activations localized to similar parts of ventral temporal cortex as in sighted individuals (Supplementary Table 1). The interpretation of virtually all these studies is that there must be some innate predisposition for these regions to process particular high-level categories, regardless of the modality of the input, and that this bias is driven by top-down influences from still higher areas in the cortex. For example, in studies finding face selectivity in the fusiform sulcus in blind individuals presented with tactile faces¹⁰⁹, or sounds associated with faces¹¹⁰, a social area is proposed to provide top-down input that causes the face area to become selective to faces. Studies show that blind individuals reading Braille (tactile) or hearing auditory input that is associated with text activates a region that in sighted individuals is responsive to reading text^{111,112}. These studies propose that the VWFA is selectively connected to language areas early in development, and that it is this connectivity that predisposes this area to become selective for reading.

However, the idea of an innate reading area seems inconsistent with how recently in human history literacy has become prevalent. There is also a report that the mediolateral gradation of the animate–inanimate category found in sighted individuals is found in blind individuals in response to auditory inputs¹¹³. Furthermore, there is a report that that small manipulable things¹¹⁴ and large immovable things¹¹⁵ presented auditorily to blind individuals map to similar brain locations as they do for images of the same objects in sighted individuals.

It seems like a heavy burden to require top-down influences from higher cortical areas to guide the organization of ventral temporal cortex into domains selective for faces, scenes, body parts, text, big versus small things and animate versus inanimate things. This idea of top-down influences guiding IT or VOT domain formation is especially difficult to accept given that lower areas in the hierarchy preceded higher areas in evolution¹¹⁶, lower areas mature earlier than higher areas^{117,118}, feedforward projections precede feedback projections^{119,120} and most inputs to IT or VOT arise predominantly from other visual areas, not higher areas^{97,121}.

Can a bottom-up explanation account for the localization of these category-selective responses in blind individuals? It has been proposed that cross-modal mapping (after loss of one modality) is generated by a reorganization of either thalamocortical pathways¹²² or corticocortical connectivity between sensory areas¹²³, with connections propagating throughout cortex via intrinsic map organization^{27,28}, or via multimodal intermediate regions such as parietal areas or parts of frontal cortex¹²⁴. Regardless of the specific pathway, we propose that the

explanation for commonalities between original and substituted modalities lies in the global congruence and connectivity of maps that are conserved across evolution. In mice, monkeys and humans, there is shared orientation of the topographic representation of the sensory periphery across modalities (FIG. 4a): in both somatosensory and motor cortex, the face and upper body are mapped ventral in cortex, and the feet and lower body are mapped dorsal in cortex. This is also true in visual cortex: the lower-visual-field maps to dorsal occipital cortex, and the upper-field maps to ventral cortex. Indeed, the orientations of these different sensory maps are established by the same trophic factors, ephrins, early in development¹²⁵.

We propose that topography-preserving long-range projections link up sensory maps, and that the congruency of the topographic maps enforces a regularity by which sensory modalities are interlinked.

Such congruency of sensory maps is apparent in multimodal association cortex, even in the absence of neurological damage. In parietal cortex, there are neurons that receive both somatosensory and visual input and have congruent visual and somatosensory fields¹²⁶ (FIG. 4b). Some visually responsive parietal cells also receive auditory input, with auditory-localization fields aligned with their visual receptive fields^{127,128}. Map congruency is also apparent in the alignment of visual, somatosensory and auditory representations across layers of

the superior colliculus^{129,130}. Such linking of sensory maps provides an infrastructure for the correspondence of information about the environment across modalities. These connections may provide a pathway for inputs from other modalities when the normally dominant modality is damaged or missing.

To explain the localization of cross-modal face selectivity in blind individuals, we invoke the fact that connectivity within IT is along isoeccentricities¹³¹ (FIG. 4c), even in congenitally blind individuals^{132–134}. Isoeccentricity connectivity is broadly aligned anteroposteriorly across IT. We speculate that, projecting that connectivity, the central visual field aligns with low frequencies in the auditory map, and eventually

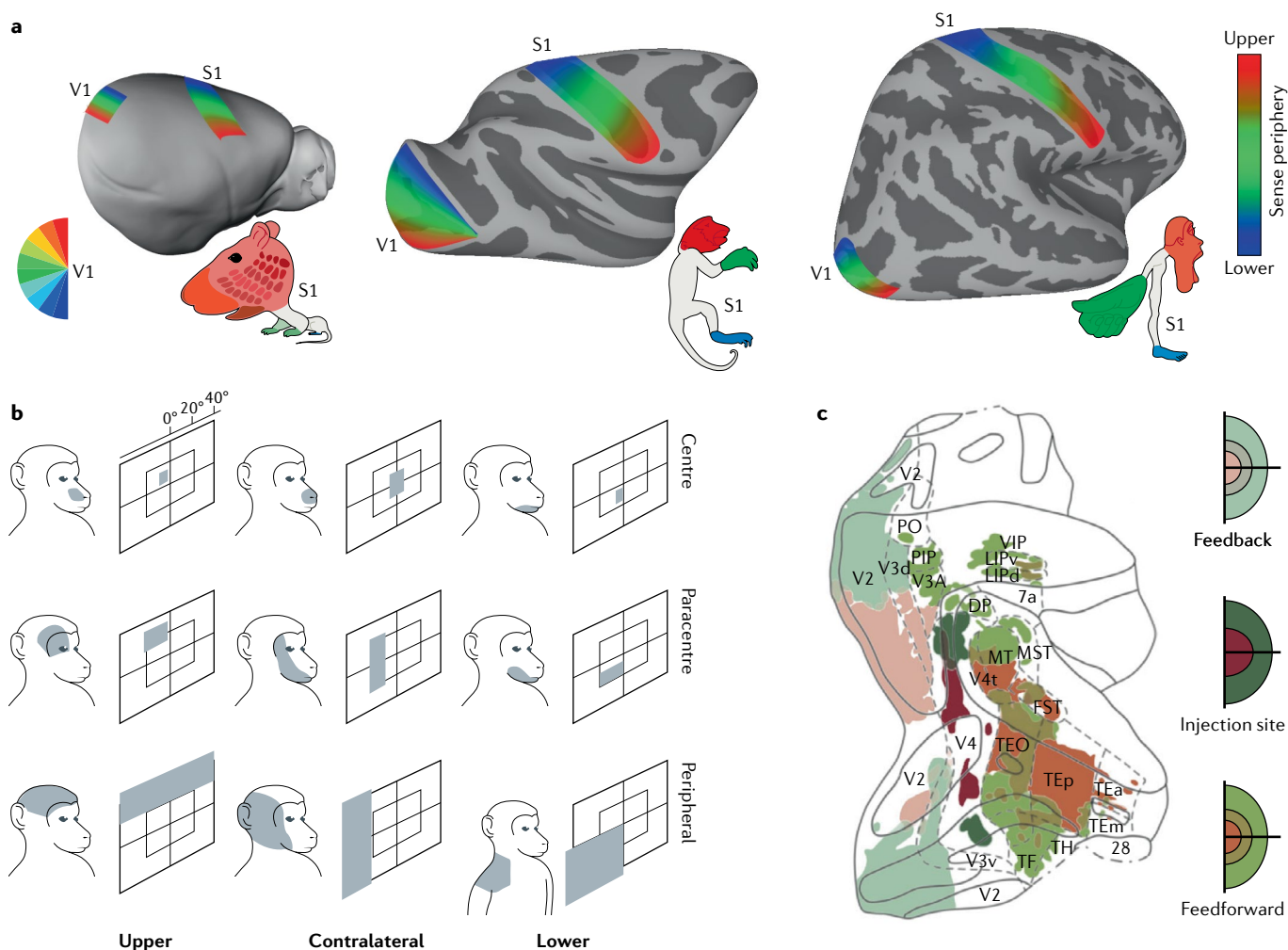


Fig. 4 | Congruence between sensory maps. a | Global congruency of sensory map orientation in mice (left), macaques (middle) and humans (right). **b** | Congruency of visual and somatosensory maps in parietal cortex¹²⁶. Neurons responding to stimulation of central visual space also respond to touch on central parts of the face, whereas neurons responding to peripheral visual space respond to touch on more peripheral body parts. **c** | Connectivity between early visual cortex and inferotemporal cortex is predominantly along isoeccentricities¹³⁴. Neuronal tracer injection sites in

central-visual-field parts of posterior inferotemporal cortex (dark red) are selectively connected to central-visual-field parts of early visual areas and higher visual areas (lighter red). Injection sites in peripheral-visual-field parts of intermediate visual areas (dark green) are selectively connected to peripheral-visual-field parts of both lower and higher visual areas (light green). S1, primary somatosensory cortex; V1, primary visual cortex. Part b adapted with permission from REF.¹²⁶, The American Physiological Society. Part c adapted with permission from REF.¹³⁴, OUP.

with the face part of the somatosensory and motor body maps. It would follow that central-visual-field representations, which are biased by scale and by looking behaviour to become specialized during development for processing faces, would be, early in development, congruent with and interconnected with the low-frequency part of the auditory map (which itself becomes specialized for processing language¹³⁵) and with the face parts of the somatosensory map.

We further postulate that social areas in frontal lobes, rather than preceding and guiding the development of earlier sensory areas, selectively receive inputs from face IT, face body-map areas and vocalization areas owing to topography-preserving constraints, and it is those convergent inputs that define those areas as 'social'. Connectivity with frontal cortex, and indeed across the entire cortex, is topographic^{136,137}. Furthermore, visual domains for faces, scenes, colours and depth may be organized topographically in macaque prefrontal cortex¹³⁸, probably reflecting topographic input from IT cortex^{54,137}. These prefrontal domains are probably responsive to multiple modalities¹³⁹ and suggest a convergence of topographic inputs across sensory modalities (for example, see REFS^{136,140}). Thus, the convergence of inputs from retinotopic maps of early visual areas could lead to topographic organization of experienced categories in mid-level areas such as IT or VOT, and the same principles can lead to still more abstract and multimodal topographies in higher areas such as prefrontal cortex (Supplementary Fig. 3).

The innatist argument that input from social areas drives face domains to respond selectively to faces requires that the social areas have a template for what is social and, in particular, what qualifies as a face. By contrast, the bottom-up model requires only that whatever the infant sees frequently in central visual field will sculpt the selectivity of central IT, and, for most monkey and human infants, that stimulus is faces¹⁴¹. Indeed, newborn humans see faces fully 25% of the time during waking, and mostly very close up⁸¹.

By extension, the activation of the text domain by Braille reading, or the activation of the fusiform gyrus by haptic or auditory face cues, is a small subset of the kinds of cross-modal activations observed in individuals deprived of one sensory modality, or even sometimes in controls^{109,142}. We predict that further studies that incorporate global topographic principles will find that such cross-modal activity shows anatomical specificities based on shared map axes.

Indeed, somatosensory, motor and visual maps are congruent with respect to the environment and are already congruent at birth (FIG. 3) throughout the class of mammals. The fact that newborn precocial mammals, such as lambs, show visual and auditory orienting behaviour at birth¹⁴³ means that visual and motor maps, and visual and auditory maps, at least, are congruent and interconnected at birth. We propose that the apparent category selectivity of cross-modal activations in blind individuals is evidence for map-based congruities, not innate domain organization. Further observations of high-level congruities may reveal how high-level functions such as language develop from low-level topographic functional specializations present in evolutionarily older species.

Conclusions

We have described how the development of category selectivity in temporal cortex begins with the establishment of maps across the entire visual hierarchy. The axes of the primary maps are defined by molecular gradients and refined by activity-dependent synaptic reinforcement and pruning. It is unknown how the multiple higher maps are formed prenatally, but the same activity-dependent sorting mechanisms and waves of activity in peripheral organs and cortical areas^{144–146} would promote smoothly organized topographies. These map axes are congruent across the entire brain, but become less anchored to peripheral inputs, more abstract and more multimodal going up the hierarchy, beyond early sensory areas. The maps within the visual system provide an eccentricity-based curvature organization that provides a proto-shape organization for the topographic organization of categories. During postnatal development, the proto-architecture is modified by daily experience to become selectively responsive to frequently encountered things, biased by low-level features and how these things are typically viewed. Differential cytoarchitectonics may also be acquired as a consequence of patterns of neuronal activity.

Marr¹⁴⁷ said, somewhat teleologically, that we need to understand the goal of a computation before asking how it is neurally implemented; we suggest that asking how a circuit gets wired up may be even more informative. We believe that topographic maps and self-organizing plasticity rules acting both prenatally and postnatally may provide explanations for seemingly complex circuitries, without the necessity to argue that these circuits evolved to do exactly what they do.

Michael J. Arcaro¹ and Margaret S. Livingstone^{1,2}✉

¹University of Pennsylvania, Philadelphia, PA, USA.

²Harvard Medical School, Boston, MA, USA.

✉e-mail: mlivingstone@hms.harvard.edu

<https://doi.org/10.1038/s41583-021-00490-4>

Published online: 03 August 2021

- Underleider, L. G. & Mishkin, M. In *Analysis of Visual Behavior* (eds Ingle, M. A., Goodale, M. I. & Masfield, R. J. W.) 549–586 (MIT Press, 1982).
- Goodale, M. A. & Milner, A. D. Separate visual pathways for perception and action. *Trends Neurosci.* **15**, 20–25 (1992).
- Farah, M. J. *Visual Agnosia* (MIT Press, 2004).
- Konorski, J. Integrative activity of the brain: an interdisciplinary approach (University of Chicago Press, 1967).
- Bell, A. H., Hadj-Bouziane, F., Frihauf, J. B., Tootell, R. B. & Ungerleider, L. G. Object representations in the temporal cortex of monkeys and humans as revealed by functional magnetic resonance imaging. *J. Neurophysiol.* **101**, 688–700 (2009).
- Pinsk, M. A., Desimone, K., Moore, T., Gross, C. G. & Kastner, S. Representations of faces and body parts in macaque temporal cortex: a functional MRI study. *Proc. Natl Acad. Sci. USA* **102**, 6996–7001 (2005).
- Tsao, D. Y., Moeller, S. & Freiwald, W. A. Comparing face patch systems in macaques and humans. *Proc. Natl Acad. Sci. USA* **105**, 19514–19519 (2008).
- Parvizi, J. & Kastner, S. Promises and limitations of human intracranial electroencephalography. *Nat. Neurosci.* **21**, 474–483 (2018).
- Tsao, D. Y. & Livingstone, M. S. Mechanisms of face perception. *Annu. Rev. Neurosci.* **31**, 411–437 (2008).
- Kanwisher, N. Domain specificity in face perception. *Nat. Neurosci.* **3**, 759–763 (2000).
- Weiner, K. S. & Zilles, K. The anatomical and functional specialization of the fusiform gyrus. *Neuropsychologia* **83**, 48–62 (2016).
- Arcaro, M. J., Schade, P. F. & Livingstone, M. S. Universal mechanisms and the development of the face network: what you see is what you get. *Annu. Rev. Vis. Sci.* **5**, 341–372 (2019).
- Chomsky, N. *Knowledge of Language: Its Nature, Origin, and Use* (Praeger, 1986).
- McKone, E., Crookes, K., Jeffery, L. & Dilks, D. D. A critical review of the development of face recognition: experience is less important than previously believed. *Cogn. Neuropsychol.* **29**, 174–212 (2012).
- Dehaene, S. & Cohen, L. Cultural recycling of cortical maps. *Neuron* **56**, 384–398 (2007).
- Desimone, R. & Gross, C. G. Visual areas in the temporal cortex of the macaque. *Brain Res.* **178**, 363–380 (1979).
- Grill-Spector, K. & Malach, R. The human visual cortex. *Annu. Rev. Neurosci.* **27**, 649–677 (2004).
- Arcaro, M. J., Schade, P. F., Vincent, J. L., Ponce, C. R. & Livingstone, M. S. Seeing faces is necessary for face-domain formation. *Nat. Neurosci.* **20**, 1404–1412 (2017).
- Deen, B. et al. Organization of high-level visual cortex in human infants. *Nat. Commun.* **8**, 13995 (2017).
- Brem, S. et al. Brain sensitivity to print emerges when children learn letter-speech sound correspondences. *Proc. Natl Acad. Sci. USA* **107**, 7939–7944 (2010).
- Ben-Shachar, M., Dougherty, R. F., Deutsch, G. K. & Wandell, B. A. The development of cortical sensitivity to visual word forms. *J. Cogn. Neurosci.* **23**, 2387–2399 (2011).
- Bouhali, F. et al. Anatomical connections of the visual word form area. *J. Neurosci.* **34**, 15402–15414 (2014).
- Mahon, B. Z. & Caramazza, A. What drives the organization of object knowledge in the brain? *Trends Cogn. Sci.* **15**, 97–103 (2011).
- Srihasam, K., Mandeville, J. B., Morocz, I. A., Sullivan, K. J. & Livingstone, M. S. Behavioral and anatomical consequences of early versus late symbol training in macaques. *Neuron* **73**, 608–619 (2012).
- Srihasam, K., Vincent, J. L. & Livingstone, M. S. Novel domain formation reveals proto-architecture in inferotemporal cortex. *Nat. Neurosci.* **17**, 1776–1783 (2014).
- Livingstone, M. S., Srihasam, K. & Morocz, I. A. The benefit of symbols: monkeys show linear, human-like, accuracy when using symbols to represent scalar value. *Anim. Cogn.* **13**, 711–719 (2010).

27. Arcaro, M. J. & Livingstone, M. S. A hierarchical, retinotopic proto-organization of the primate visual system at birth. *eLife* **6**, e26196 (2017).
28. Arcaro, M. J., Schade, P. F. & Livingstone, M. S. Body map proto-organization in newborn macaques. *Proc. Natl Acad. Sci. USA* **116**, 24861–24871 (2019).
29. Arsenault, J. T., Janssens, T., Polimeni, J. R., Wald, L. L. & Vanduffel, W. Ultra high-resolution fMRI with an implanted 8-channel array coil reveals detailed retinotopic maps. *Soc. Neurosci. Abstr.* **573**, 63261337 (2012).
30. Kolster, H., Janssens, T., Orban, G. A. & Vanduffel, W. The retinotopic organization of macaque occipitotemporal cortex anterior to V4 and caudovertral to the middle temporal (MT) cluster. *J. Neurosci.* **34**, 10168–10191 (2014).
31. Arcaro, M. J. & Livingstone, M. S. Retinotopic organization of scene areas in macaque inferior temporal cortex. *J. Neurosci.* **37**, 7373–7389 (2017).
32. Kolster, H. et al. Visual field map clusters in macaque extrastriate visual cortex. *J. Neurosci.* **29**, 7031–7039 (2009).
33. Tootell, R. B. & Hadjikhani, N. Where is 'dorsal V4' in human visual cortex? Retinotopic, topographic and functional evidence. *Cereb. Cortex* **11**, 298–311 (2001).
34. Norman-Haignere, S. V., Kanwisher, N., McDermott, J. H. & Conway, B. R. Divergence in the functional organization of human and macaque auditory cortex revealed by fMRI responses to harmonic tones. *Nat. Neurosci.* **22**, 1057–1060 (2019).
35. Petkov, C. I., Kayser, C., Augath, M. & Logothetis, N. K. Functional imaging reveals numerous fields in the monkey auditory cortex. *PLoS Biol.* **4**, e215 (2006).
36. Levy, I., Hasson, U., Avidan, G., Hendler, T. & Malach, R. Center-periphery organization of human object areas. *Nat. Neurosci.* **4**, 533–539 (2001).
37. Lafer-Sousa, R. & Conway, B. R. Parallel, multi-stage processing of colors, faces and shapes in macaque inferior temporal cortex. *Nat. Neurosci.* **16**, 1870–1878 (2013).
38. Hasson, U., Levy, I., Behrmann, M., Hendler, T. & Malach, R. Eccentricity bias as an organizing principle for human high-order object areas. *Neuron* **34**, 479–490 (2002).
39. Konkle, T. & Oliva, A. A real-world size organization of object responses in occipitotemporal cortex. *Neuron* **74**, 1114–1124 (2012).
40. Hasson, U., Harel, M., Levy, I. & Malach, R. Large-scale mirror-symmetry organization of human occipitotemporal object areas. *Neuron* **37**, 1027–1041 (2003).
41. Kamps, F. S., Hendrix, C. L., Brennan, P. A. & Dilks, D. D. Connectivity at the origins of domain specificity in the cortical face and place networks. *Proc. Natl Acad. Sci. USA* **117**, 6163–6169 (2020).
42. Hubel, D. H. & Wiesel, T. N. Uniformity of monkey striate cortex: a parallel relationship between field size, scatter, and magnification factor. *J. Comp. Neurol.* **158**, 295–305 (1974).
43. Hubel, D. H. & Wiesel, T. N. Receptive fields and functional architecture in two nonstriate visual areas (18 and 19) of the cat. *J. Neurophysiol.* **28**, 229–289 (1965).
44. Hubel, D. H. & Wiesel, T. N. Receptive fields and functional architecture of monkey striate cortex. *J. Physiol.* **195**, 215–243 (1968).
45. Hubel, D. H. & Livingstone, M. S. Segregation of form, color, and stereopsis in primate area 18. *J. Neurosci.* **7**, 3378–3415 (1987).
46. Yue, X., Pourladian, I. S., Tootell, R. B. & Ungerleider, L. G. Curvature-processing network in macaque visual cortex. *Proc. Natl Acad. Sci. USA* **111**, E3467–E3475 (2014).
47. Yue, X., Robert, S. & Ungerleider, L. G. Curvature processing in human visual cortical areas. *Neuroimage* **222**, 117295 (2020).
48. Bao, P., She, L., McGill, M. & Tsao, D. Y. A map of object space in primate inferotemporal cortex. *Nature* **583**, 103–108 (2020).
49. Bell, A. H. et al. Relationship between functional magnetic resonance imaging-identified regions and neuronal category selectivity. *J. Neurosci.* **31**, 12229–12240 (2011).
50. Aparicio, P. L., Issa, E. B. & DiCarlo, J. J. Neurophysiological organization of the middle face patch in macaque inferior temporal cortex. *J. Neurosci.* **36**, 12729–12745 (2016).
51. Durbin, R. & Mitchison, G. A dimension reduction framework for understanding cortical maps. *Nature* **343**, 644–647 (1990).
52. Swindale, N. V., Shoham, D., Grinvald, A., Bonhoeffer, T. & Hubner, M. Visual cortex maps are optimized for uniform coverage. *Nat. Neurosci.* **3**, 822–826 (2000).
53. Hubel, D. H. & Wiesel, T. N. Sequence regularity and geometry of orientation columns in the monkey striate cortex. *J. Comp. Neurol.* **158**, 267–293 (1974).
54. Grimaldi, P., Saleem, K. S. & Tsao, D. Anatomical connections of the functionally defined "face patches" in the macaque monkey. *Neuron* **90**, 1325–1342 (2016).
55. Lyon, D. C. & Connolly, J. D. The case for primate V3. *Proc. Biol. Sci.* **279**, 625–633 (2012).
56. Livingstone, M. S. & Hubel, D. H. Anatomy and physiology of a color system in the primate visual cortex. *J. Neurosci.* **4**, 309–356 (1984).
57. Malach, R., Levy, I. & Hasson, U. The topography of high-order human object areas. *Trends Cogn. Sci.* **6**, 176–184 (2002).
58. Konkle, T. & Caramazza, A. Tripartite organization of the ventral stream by animacy and object size. *J. Neurosci.* **33**, 10235–10242 (2013).
59. Long, B., Yu, C.-P. & Konkle, T. Mid-level visual features underlie the high-level categorical organization of the ventral stream. *Proc. Natl Acad. Sci. USA* **115**, E9015–E9024 (2018).
60. Cogan, D. D., Liu, W., Baker, D. H. & Andrews, T. J. Category-selective patterns of neural response in the ventral visual pathway in the absence of categorical information. *Neuroimage* **135**, 107–114 (2016).
61. Peelen, M. V. & Downing, P. E. Category selectivity in human visual cortex: Beyond visual object recognition. *Neuropsychologia* **105**, 177–183 (2017).
62. Bracci, S., Ritchie, J. B. & de Bock, H. O. On the partnership between neural representations of object categories and visual features in the ventral visual pathway. *Neuropsychologia* **105**, 153–164 (2017).
63. Kriegeskorte, N. et al. Matching categorical object representations in inferior temporal cortex of man and monkey. *Neuron* **60**, 1126–1141 (2008).
64. Bracci, S. & Op de Beeck, H. Dissociations and associations between shape and category representations in the two visual pathways. *J. Neurosci.* **36**, 432–444 (2016).
65. Carlson, T. A., Simmons, R. A., Kriegeskorte, N. & Slevc, L. R. The emergence of semantic meaning in the ventral temporal pathway. *J. Cogn. Neurosci.* **26**, 120–131 (2014).
66. Bryan, P. B., Julian, J. B. & Epstein, R. A. Rectilinear edge selectivity is insufficient to explain the category selectivity of the parahippocampal place area. *Front. Hum. Neurosci.* **10**, 137 (2016).
67. Bi, H. et al. Neuronal responses in visual area V2 (V2) of macaque monkeys with strabismic amblyopia. *Cereb. Cortex* **21**, 2033–2045 (2011).
68. Kiorpes, L., Walton, P. J., O'Keefe, L. P., Movshon, J. A. & Lisberger, S. G. Effects of early-onset artificial strabismus on pursuit eye movements and on neuronal responses in area MT of macaque monkeys. *J. Neurosci.* **16**, 6537–6553 (1996).
69. Tsao, D. Y., Freiwald, W. A., Tootell, R. B. & Livingstone, M. S. A cortical region consisting entirely of face-selective cells. *Science* **311**, 670–674 (2006).
70. Golarai, G. et al. Differential development of high-level visual cortex correlates with category-specific recognition memory. *Nat. Neurosci.* **10**, 512–522 (2007).
71. Scherf, K. S., Behrmann, M., Humphreys, K. & Luna, B. Visual category-selectivity for faces, places and objects emerges along different developmental trajectories. *Dev. Sci.* **10**, F15–F30 (2007).
72. Hubel, D. H., Wiesel, T. N. & LeVay, S. Functional architecture of area 17 in normal and monocularly deprived macaque monkeys. *Cold Spring Harb. Symp. Quant. Biol.* **40**, 581–589 (1976).
73. Wiesel, T. N. & Hubel, D. H. Comparison of the effects of unilateral and bilateral eye closure on cortical unit responses in kittens. *J. Neurophysiol.* **28**, 1029–1040 (1965).
74. Blakemore, C. & Cooper, G. F. Development of the brain depends on the visual environment. *Nature* **228**, 477–478 (1970).
75. Sengpiel, F., Stawinski, P. & Bonhoeffer, T. Influence of experience on orientation maps in cat visual cortex. *Nat. Neurosci.* **2**, 727–732 (1999).
76. Stryker, M. P. & Sherk, H. Modification of cortical orientation selectivity in the cat by restricted visual experience: a reexamination. *Science* **190**, 904–906 (1975).
77. Tanaka, S., Ribot, J., Imamura, K. & Tani, T. Orientation-restricted continuous visual exposure induces marked reorganization of orientation maps in early life. *Neuroimage* **30**, 462–477 (2006).
78. Aboitiz, F., Scheibel, A. B., Fisher, R. S. & Zaidel, E. Individual differences in brain asymmetries and fiber composition in the human corpus callosum. *Brain Res.* **598**, 154–161 (1992).
79. Zhang, L. I., Bao, S. & Merzenich, M. M. Persistent and specific influences of early acoustic environments on primary auditory cortex. *Nat. Neurosci.* **4**, 1123–1130 (2001).
80. Van der Loos, H. & Woolsey, T. A. Somatosensory cortex: structural alterations following early injury to sense organs. *Science* **179**, 395–398 (1973).
81. Jayaraman, S., Fausey, C. M. & Smith, L. B. The faces in infant-perspective scenes change over the first year of life. *PLoS ONE* **10**, e0123780 (2015).
82. Goren, C. C., Sarty, M. & Wu, P. Y. Visual following and pattern discrimination of face-like stimuli by newborn infants. *Pediatrics* **56**, 544–549 (1975).
83. Johnson, M. H., Dziurawiec, S., Ellis, H. & Morton, J. Newborns' preferential tracking of face-like stimuli and its subsequent decline. *Cognition* **40**, 1–19 (1991).
84. Kenney, M. D., Mason, W. A. & Hill, S. D. Effects of age, objects, and visual experience on affective responses of rhesus monkeys to strangers. *Dev. Psychol.* **15**, 176–184 (1979).
85. Maurer, D. & Young, R. Newborns' following of natural and distorted arrangements of facial features. *Infant. Behav. Dev.* **6**, 127–131 (1983).
86. Valenza, E., Simion, F., Cassia, V. M. & Umiltà, C. Face preference at birth. *J. Exp. Psychol. Hum. Percept. Perform.* **22**, 892–903 (1996).
87. Banks, M. S. & Salapatek, P. Contrast sensitivity function of the infant visual system. *Vis. Res.* **16**, 867–869 (1976).
88. Kleiner, K. Amplitude and phase spectra as indices of infants' pattern preferences. *Infant. Behav. Dev.* **10**, 49–59 (1987).
89. Simion, F., Valenza, E., Cassia, V. M., Turati, C. & Umiltà, C. Newborns' preference for up-down asymmetrical configurations. *Dev. Sci.* **5**, 427–434 (2002).
90. Bushnell, I. W. R. Mother's face recognition in newborn infants: learning and memory. *Infant. Child. Dev.* **10**, 67–74 (2001).
91. Wurtz, R. H. & Albano, J. E. Visual-motor function of the primate superior colliculus. *Annu. Rev. Neurosci.* **3**, 189–226 (1980).
92. Hafed, Z. M. & Chen, C. Y. Sharper, stronger, faster upper visual field representation in primate superior colliculus. *Curr. Biol.* **26**, 1647–1658 (2016).
93. Wallace, M. T., McHaffie, J. G. & Stein, B. E. Visual response properties and visuotopic representation in the newborn monkey superior colliculus. *J. Neurophysiol.* **78**, 2732–2741 (1997).
94. Sugita, Y. Face perception in monkeys reared with no exposure to faces. *Proc. Natl Acad. Sci. USA* **105**, 394–398 (2008).
95. Davies-Thompson, J. & Andrews, T. J. Intra- and interhemispheric connectivity between face-selective regions in the human brain. *J. Neurophysiol.* **108**, 3087–3095 (2012).
96. Gschwind, M., Pourtois, G., Schwartz, S., Van De Ville, D. & Vuilleumier, P. White-matter connectivity between face-responsive regions in the human brain. *Cereb. Cortex* **22**, 1564–1576 (2012).
97. Moeller, S., Freiwald, W. A. & Tsao, D. Y. Patches with links: a unified system for processing faces in the macaque temporal lobe. *Science* **320**, 1355–1359 (2008).
98. Saygin, Z. M. et al. Anatomical connectivity patterns predict face selectivity in the fusiform gyrus. *Nat. Neurosci.* **15**, 321–327 (2011).
99. Saygin, Z. M. et al. Connectivity precedes function in the development of the visual word form area. *Nat. Neurosci.* **19**, 1250–1255 (2016).
100. Arcaro, M., Mautz, T. & Livingstone, M. S. Anatomical correlates of face patches in macaque inferotemporal cortex. *Proc. Natl Acad. Sci. USA* **117**, 32667–32678 (2020).
101. Brodmann, K. *Localisation in the Cerebral Cortex* (Imperial Press, 1909).
102. Janssens, T., Zhu, Q., Popivanov, I. D. & Vanduffel, W. Probabilistic and single-subject retinotopic maps reveal the topographic organization of face patches in the macaque cortex. *J. Neurosci.* **34**, 10156–10167 (2014).
103. DeFelipe, J., Alonso-Nanclares, L. & Arellano, J. I. Microstructure of the neocortex: comparative aspects. *J. Neurocytol.* **31**, 299–316 (2002).
104. Miller, K. D. Canonical computations of cerebral cortex. *Curr. Opin. Neurobiol.* **37**, 75–84 (2016).
105. Hebb, D. O. *The Organization of Behavior: a Neuropsychological Theory* (Wiley, 1949).

106. Constantine-Paton, M. & Law, M. I. Eye-specific termination bands in tecta of three-eyed frogs. *Science* **202**, 639–641 (1978).
107. Leibo, J. Z., Liao, Q., Anselmi, F. & Poggio, T. The invariance hypothesis implies domain-specific regions in visual cortex. *PLoS Comput. Biol.* **11**, e1004390 (2015).
108. Nasr, S., Polimeni, J. R. & Tootell, R. B. Interdigitated color- and disparity-selective columns within human visual cortical areas V2 and V3. *J. Neurosci.* **36**, 1841–1857 (2016).
109. Murty, N. A. et al. Visual experience is not necessary for the development of face-selectivity in the lateral fusiform gyrus. *Proc. Natl Acad. Sci. USA* **117**, 23011–23020 (2020).
110. van den Hurk, J., Van Baelen, M. & Op de Beeck, H. P. Development of visual category selectivity in ventral visual cortex does not require visual experience. *Proc. Natl Acad. Sci. USA* **114**, E4501–E4510 (2017).
111. Reich, L., Szwed, M., Cohen, L. & Amedi, A. A ventral visual stream reading center independent of visual experience. *Curr. Biol.* **21**, 363–368 (2011).
112. Striem-Amit, E., Cohen, L., Dehaene, S. & Amedi, A. Reading with sounds: sensory substitution selectively activates the visual word form area in the blind. *Neuron* **76**, 640–652 (2012).
113. Mahon, B. Z., Anzellotti, S., Schwarzbach, J., Zampini, M. & Caramazza, A. Category-specific organization in the human brain does not require visual experience. *Neuron* **63**, 397–405 (2009).
114. Mahon, B. Z., Schwarzbach, J. & Caramazza, A. The representation of tools in left parietal cortex is independent of visual experience. *Psychol. Sci.* **21**, 764–771 (2010).
115. He, C. et al. Selectivity for large nonmanipulable objects in scene-selective visual cortex does not require visual experience. *Neuroimage* **79**, 1–9 (2013).
116. Bourne, J. A. & Rosa, M. G. Hierarchical development of the primate visual cortex, as revealed by neurofilament immunoreactivity: early maturation of the middle temporal area (MT). *Cereb. Cortex* **16**, 405–414 (2006).
117. Flechsig, P. E. *Anatomie des Menschlichen Gehirn und Rückenmarks, auf Myelogenetischer Grundlage* (Thieme, 1920).
118. Guillery, R. W. Is postnatal neocortical maturation hierarchical? *Trends Neurosci.* **28**, 512–517 (2005).
119. Burkhalter, A. Development of forward and feedback connections between areas V1 and V2 of human visual cortex. *Cereb. Cortex* **3**, 476–487 (1993).
120. Dong, H., Wang, Q., Valkova, K., Gonchar, Y. & Burkhalter, A. Experience-dependent development of feedforward and feedback circuits between lower and higher areas of mouse visual cortex. *Vis. Res.* **44**, 3389–3400 (2004).
121. Markov, N. T. et al. A weighted and directed interareal connectivity matrix for macaque cerebral cortex. *Cereb. Cortex* **24**, 17–36 (2014).
122. Bavelier, D. & Neville, H. J. Cross-modal plasticity: where and how? *Nat. Rev. Neurosci.* **3**, 443–452 (2002).
123. Bedny, M., Pascual-Leone, A., Dodell-Feder, D., Fedorenko, E. & Saxe, R. Language processing in the occipital cortex of congenitally blind adults. *Proc. Natl Acad. Sci. USA* **108**, 4429–4434 (2011).
124. Fuster, J. M., Bodner, M. & Kroger, J. K. Cross-modal and cross-temporal association in neurons of frontal cortex. *Nature* **405**, 347–351 (2000).
125. Flanagan, J. G. Neural map specification by gradients. *Curr. Opin. Neurobiol.* **16**, 59–66 (2006).
126. Duhamel, J. R., Colby, C. L. & Goldberg, M. E. Ventral intraparietal area of the macaque: congruent visual and somatic response properties. *J. Neurophysiol.* **79**, 126–136 (1998).
127. Andersen, R. A. Multimodal integration for the representation of space in the posterior parietal cortex. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **352**, 1421–1428 (1997).
128. Mullette-Gillman, O. A., Cohen, Y. E. & Groh, J. M. Motor-related signals in the intraparietal cortex encode locations in a hybrid, rather than eye-centered reference frame. *Cereb. Cortex* **19**, 1761–1775 (2009).
129. Wallace, M. T. & Stein, B. E. Sensory and multisensory responses in the newborn monkey superior colliculus. *J. Neurosci.* **21**, 8886–8894 (2001).
130. Triplett, J. W., Phan, A., Yamada, J. & Feldheim, D. A. Alignment of multimodal sensory input in the superior colliculus through a gradient-matching mechanism. *J. Neurosci.* **32**, 5264–5271 (2012).
131. Kravitz, D. J., Saleem, K. S., Baker, C. I., Ungerleider, L. G. & Mishkin, M. The ventral visual pathway: an expanded neural framework for the processing of object quality. *Trends Cogn. Sci.* **17**, 26–49 (2013).
132. Butt, O. H., Benson, N. C., Datta, R. & Aguirre, G. K. The fine-scale functional correlation of striate cortex in sighted and blind people. *J. Neurosci.* **33**, 16209–16219 (2013).
133. Striem-Amit, E. et al. Functional connectivity of visual cortex in the blind follows retinotopic organization principles. *Brain* **138**, 1679–1695 (2015).
134. Ungerleider, L. G., Galkin, T. W., Desimone, R. & Gattass, R. Cortical connections of area V4 in the macaque. *Cereb. Cortex* **18**, 477–499 (2008).
135. Moerel, M., De Martino, F. & Formisano, E. Processing of natural sounds in human auditory cortex: tonotopy, spectral tuning, and relation to voice sensitivity. *J. Neurosci.* **32**, 14205–14216 (2012).
136. O’Rawe, J. F. & Leung, H. C. Topographic mapping as a basic principle of functional organization for visual and prefrontal functional connectivity. *eNeuro* <https://doi.org/10.1523/ENEURO.0532-19.2019> (2020).
137. Yeo, B. T. et al. The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *J. Neurophysiol.* **106**, 1125–1165 (2011).
138. Haile, T. M., Bohon, K. S., Romero, M. C. & Conway, B. R. Visual stimulus-driven functional organization of macaque prefrontal cortex. *Neuroimage* **188**, 427–444 (2019).
139. Tammer-Rosenau, B. J., Dux, P. E., Tombu, M. N., Asplund, C. L. & Marois, R. Amodal processing in human prefrontal cortex. *J. Neurosci.* **33**, 11573–11587 (2013).
140. Janata, P. et al. The cortical topography of tonal structures underlying Western music. *Science* **298**, 2167–2170 (2002).
141. Fausey, C. M., Jayaraman, S. & Smith, L. B. From faces to hands: changing visual input in the first two years. *Cognition* **152**, 101–107 (2016).
142. Eagleman, D. M., Kagan, A. D., Nelson, S. S., Sagaram, D. & Sarma, A. K. A standardized test battery for the study of synesthesia. *J. Neurosci. Methods* **159**, 139–145 (2007).
143. Fox, M. W. A phylogenetic analysis of behavioral neuro-ontogeny in precocial and nonprecocial mammals. *Can. J. Comp. Med. Vet. Sci.* **28**, 197–202 (1964).
144. Jones, T. A., Leake, P. A., Snyder, R. L., Stakhovskaya, O. & Bonham, B. Spontaneous discharge patterns in cochlear spiral ganglion cells before the onset of hearing in cats. *J. Neurophysiol.* **98**, 1898–1908 (2007).
145. Khazipov, R. et al. Early motor activity drives spindle bursts in the developing somatosensory cortex. *Nature* **432**, 758–761 (2004).
146. Meister, M., Wong, R. O., Baylor, D. A. & Shatz, C. J. Synchronous bursts of action potentials in ganglion cells of the developing mammalian retina. *Science* **252**, 939–943 (1991).
147. Marr, D. *Vision: a Computational Investigation into the Human Representation and Processing of Visual Information* (Freeman, 1982).
148. Livingstone, M. S. et al. Development of the macaque face-patch system. *Nat. Commun.* **8**, 14897 (2017).
149. Rosenke, M., van Hoof, R., van den Hurk, J., Grill-Spector, K. & Goebel, R. A probabilistic functional atlas of human occipito-temporal visual cortex. *Cereb. Cortex* **31**, 603–619 (2020).
150. Benson, N. C. & Winawer, J. Individual differences in human retinotopic maps revealed by Bayesian analysis of retinotopic organization. Preprint at *bioRxiv* <https://doi.org/10.1101/325597> (2018).
151. Hubel, D. H. & Wiesel, T. N. Ferrier lecture. Functional architecture of macaque monkey visual cortex. *Proc. R. Soc. Lond. B Biol. Sci.* **198**, 1–59 (1977).
152. Mountcastle, V. B. *Perceptual Neuroscience: the Cerebral Cortex* (Harvard University Press, 1998).
153. Hofman, M. A. On the evolution and geometry of the brain in mammals. *Prog. Neurobiol.* **32**, 137–158 (1989).
154. Metin, C. & Frost, D. O. Visual responses of neurons in somatosensory cortex of hamsters with experimentally induced retinal projections to somatosensory thalamus. *Proc. Natl Acad. Sci. USA* **86**, 357–361 (1989).
155. Roe, A. W., Pallas, S. L., Hahm, J. O. & Sur, M. A map of visual space induced in primary auditory cortex. *Science* **250**, 818–820 (1990).
156. Sur, M., Pallas, S. L. & Roe, A. W. Cross-modal plasticity in cortical development: differentiation and specification of sensory neocortex. *Trends Neurosci.* **13**, 227–233 (1990).
157. Gao, W. J. & Pallas, S. L. Cross-modal reorganization of horizontal connectivity in auditory cortex without altering thalamocortical projections. *J. Neurosci.* **19**, 7940–7950 (1999).
158. Pallas, S. L., Littman, T. & Moore, D. R. Cross-modal reorganization of callosal connectivity without altering thalamocortical projections. *Proc. Natl Acad. Sci. USA* **96**, 8751–8756 (1999).
159. von Melchner, L., Pallas, S. L. & Sur, M. Visual behaviour mediated by retinal projections directed to the auditory pathway. *Nature* **404**, 871–876 (2000).
160. O’Leary, D. D. & Stanfield, B. B. Selective elimination of axons extended by developing cortical neurons is dependent on regional locale: experiments utilizing fetal cortical transplants. *J. Neurosci.* **9**, 2230–2246 (1989).
161. Schlaggar, B. L. & O’Leary, D. D. Potential of visual cortex to develop an array of functional units unique to somatosensory cortex. *Science* **252**, 1556–1560 (1991).
162. Chou, S. J. et al. Geniculocortical input drives genetic distinctions between primary and higher-order visual areas. *Science* **340**, 1239–1242 (2013).

Author contributions

The authors contributed equally to all aspects of the article.

Competing interests

The authors declare no competing interests.

Peer review information

Nature Reviews Neuroscience thanks H.-C. Leung, H. Op de Beeck and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Publisher’s note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Supplementary information

The online version contains supplementary material available at <https://doi.org/10.1038/s41583-021-00490-4>.

© Springer Nature Limited 2021