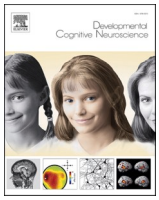
Developmental Cognitive Neuroscience 58 (2022) 101179

Contents lists available at ScienceDirect

Developmental Cognitive Neuroscience

journal homepage: www.elsevier.com/locate/dcn



Anatomical correlates of category-selective visual regions have distinctive signatures of connectivity in neonates

Laura Cabral [a,*], Leire Zubiaurre-Elorza [b], Conor J. Wild [c], Annika Linke [d], Rhodri Cusack [e]

[a] Department of Radiology, University of Pittsburgh, Pittsburgh 15224, PA, USA
[b] Department of Psychology, Faculty of Health Sciences, University of Deusto, Bilbao 48007, Spain
[c] Brain and Mind Institute, Western Interdisciplinary Research Building, Western University, London, Ontario N6A 3K7, Canada
[d] Brain Development Imaging Laboratories, San Diego State University, San Diego 92120, CA, USA
[e] Trinity College Institute of Neuroscience, Trinity College Dublin, College Green, Dublin 2, Ireland

ARTICLE INFO

Keywords:
Infant
Vision
Neuroimaging
Diffusion MRI
Tractography
Classification

ABSTRACT

The ventral visual stream is shaped during development by innate proto-organization within the visual system, such as the strong input from the fovea to the fusiform face area. In adults, category-selective regions have distinct signatures of connectivity to brain regions beyond the visual system, likely reflecting cross-modal and motoric associations. We tested if this long-range connectivity is part of the innate proto-organization, or if it develops with postnatal experience, by using diffusion-weighted imaging to characterize the connectivity of anatomical correlates of category-selective regions in neonates (N = 445), 1–9 month old infants (N = 11), and adults (N = 14). Using the HCP data we identified face- and place- selective regions and a third intermediate region with a distinct profile of selectivity. Using linear classifiers, these regions were found to have distinctive connectivity at birth, to other regions in the visual system and to those outside of it. The results support an extended proto-organization that includes long-range connectivity that shapes, and is shaped by, experience-dependent development.

1. Introduction

The development of the ventral visual stream depends upon both its innate architecture and on visual experience. The role of innate architecture is apparent in the presence of category-selective regions (Kanwisher and Yovel, 2006; Epstein, 2008; Malach et al., 1995; Kanwisher et al., 1997; Epstein and Kanwisher, 1998) at stereotypical locations even in people who have had no visual experience (Mahon et al., 2009) and in the genetic contributions to recognition performance (McKone et al., 2012). But, experience is also essential for typical development (Arcaro et al., 2017; Gomez et al., 2019), and category-selective regions can develop even for modern socio-cultural artifacts like writing (Cohen et al., 2002; Lerma-Usabiaga et al., 2018; Dehaene and Cohen, 2011).

Current evidence suggests an innate proto-organization within the ventral visual stream that shapes experience-dependent development (Arcaro and Livingstone, 2021), with mechanisms that may be universal across cortex. The fusiform face area (FFA) (Kanwisher et al., 1997), for example, receives stronger visual input from the fovea and lower spatial frequencies (Arcaro and Livingstone, 2017), leading to the acquisition of representations that are particularly suited for processing faces. Recently, it was found that this proto-organization is present at 6–27 days, as a face-selective region had stronger functional connectivity with foveal V1 than peripheral V1, while the reverse was seen for a scene region (Kamps et al., 2020). A partial proto-organization has also been found for the language-oriented visual word form area (VWFA), where one-week old neonates had adult-like functional connectivity to language areas but not object selective regions, suggesting that initial connectivity could drive further specialization (Li et al., 2020).

While there is evidence for early proto-organization for regions in the ventral visual stream, less is known about the connections from the ventral stream to other areas of the brain. By adulthood, different category-selective regions have distinct connectivity patterns (Osher et al., 2016; Saygin and Kanwisher, 2014). Such long-range connections are thought to encode the cross-modal, motoric and affective associations characteristic of rich semantic categories (Patterson et al., 2007; Huth et al., 2016; Binder et al., 2009; Peelen and Downing, 2017; Mahon and Caramazza, 2011; Martin, 2016). As concrete examples, seeing a

* Correspondence to: UPMC Children's Hospital of Pittsburgh, Department of Radiology, University of Pittsburgh, Pittsburgh 15224, PA, USA.
E-mail address: cabrall@upmc.edu (L. Cabral).

https://doi.org/10.1016/j.dcn.2022.101179
Received 21 July 2021; Received in revised form 15 November 2022; Accepted 21 November 2022
Available online 24 November 2022

silent video of a dog barking evokes the representation of its sound in auditory cortex (Meyer et al., 2010), and for tools and objects, category representations in the ventral stream are integrated with action representations in the dorsal stream (Almeida et al., 2013; Coccia et al., 2004; Goodale and Milner, 1992). The importance of long-range connections in semantics was recently demonstrated by multivariate decoding of white matter pathways in brain-injured patients with semantic deficits (Fang et al., 2018).

At present, it is unclear whether this long-range connectivity develops with experience or is innate. It could reflect the statistical learning of associations. For example, face selectivity might mature in a bottom-up way through biased input (Arcaro et al., 2017; Arcaro and Livingstone, 2017; Kamps et al., 2020; Grill-Spector and Weiner, 2014) and the auditory system might be maturing from its own proto-organization, perhaps due to selectivity to temporal modulations, to develop voice-selective regions (Belin et al., 2000). Subsequently, as faces co-occur with voices, there might be Hebbian strengthening of the long-range connections between face and voice selective regions. An alternative model is that the distinct long-range category-selective connectivity is innate. According to this model, a congenital connection between the eventual face- and voice- selective regions could shape the development of the unimodal regions.

To differentiate these models, we used neuroimaging to investigate the maturity of long-range structural connectivity of representative category-selective regions in neonates to 9-months-old infants. Most of the previous work in this area has used functional connectivity (Kamps et al., 2020) and less is known about early structural connectivity. To measure structural connectivity, we used diffusion-weighted imaging and tractography. We extracted the characteristic signatures of connectivity of two category-selective regions and a complementary third region in the ventral stream. The regions were independently defined from the Human Connectome Project, in adults using linear classifiers, and then tested for generalization in the anatomically corresponding regions in three groups of infants.

2. Experiment 1

Our goal was to characterize the brain-wide pattern of connectivity of category-selective regions in the ventral visual stream. These included a face and place selective region and a third region, with multivariate selectivity, located between the two. This allowed us to determine if connectivity was adult-like early in life, and if it matured across the first 9 months.

As the infants were scanned while sleeping, we could not use fMRI localizers to define visual category-selectivity, although to measure individual differences this would have been the ideal method. We therefore derived population-level estimates of category-selective regions in a large group of adults from the Human Connectome Project (HCP), by summarizing visually-evoked fMRI activity in each of the regions from the multimodal parcellation defined in that cohort (Glasser et al., 2016). To measure both face and place selectivity, we selected the region in the ventral visual stream that had the strongest *category>average* contrast in the HCP working memory task. We used the *category>average* contrasts to identify the most likely region to have a preferential response to the target category. Further, we choose a medial region in the ventral stream, between the face and place region, that had the strongest response to the *tool>average* contrast. This region was intended to assist in measuring the medial to lateral gradient of selectivity in the ventral stream, including object size, but this was not selective for a single visual category. This medial region is referred to as the ventromedial visual area 3 (VMV3), as named by Glasser et al. (2016), throughout the paper.

We then performed additional statistics using data from all the subjects from the HCP 1200, and confirmed that the region we choose to represent selectivity responded more to its chosen category than the other two regions chosen to represent the other categories. Statistics for this analysis can be found in the methods section below. However, as would be expected, each category did activate other regions too, as can be seen in Supplementary fig. 1, which we return to in the discussion.

Signatures of structural connectivity were then extracted from the two category-selective regions, VMV3, and the other ventral regions using a voxel-wise approach, allowing us to train a classifier on connectivity from the adult data and test for adult-like signatures in 1–9 months old infants.

2.1. Method

2.1.1. Participants

Diffusion-weighted MRI was available from 10 infants. One infant was scanned twice, but the scans were two months apart, and they were treated as separate participants, for a total of 11 datasets. Approval was provided by Western University's Health Sciences Research Ethics Board. Informed consent was obtained for all participants. Details are included in the Supplementary methods (SM).

2.1.2. Data Acquisition

Diffusion-weighted MRI was acquired using a 3T Siemens Prisma Scanner at the Center for Functional and Metabolic Mapping of Western University. Using a 20-channel head coil, the Minnesota multiband sequence was used (128 directions, 2 mm isotropic, no gap between slices, $b = 1500$ mm s^{-2}, multiband acceleration 4, monopolar diffusion encoding gradients, time of acquisition: 9 min and 18 s). Infants wore Mini-Muff sound protection (Natus, 7 dB attenuation) and ear defenders (29 dB attenuation), and adults wore standard ear protection. Infants were scanned during natural sleep.

2.1.3. Preprocessing

Data were analyzed with a pipeline built using automatic analysis (*aa*) software (Cusack et al., 2014), FSL (Smith et al., 2004), and custom Matlab (R2016a). *aa* organized files and combined images with opposite-phase encoding polarities using FSL's TOPUP. Eddy corrected for distortion and movement. Description of modules and FSL calls can be found in the Supplementary methods.

To map individual brains into standard (MNI) space for the infants and adults, the normalization procedure typical for tract-based spatial-statistics (Smith et al., 2006) was run. Normalization using a structural (T1 or T2) image was not possible for infants, as the structural images were of poor quality due to motion. Visualization of the registration is in Fig. 1.

2.1.4. Human connectome project (HCP): Glasser et al. (2016)

In order to identify seed and target regions, the parcellation from the HCP was used, which segments the brain into 180 distinct regions in each hemisphere, based on structural, functional, and diffusion data. The 14 regions that overlapped with our fMRI field-of-view, seven in each hemisphere, defined as the ventral visual stream (VVS) were chosen as seed regions (region 4, supplementary neuroanatomical results) (Glasser et al., 2016). These regions were named by Glasser et al. (2016) as V8, the Ventromedial Visual Complex, the PIT Complex, the Fusiform Face Complex, and Ventro-Medial Visual areas 1, 2 and 3. All voxels that were part of these 14 regions were used as seeds and were excluded from tractography targets. The other 346 regions from the parcellation served as the target regions for the tractography. These seed and target regions were projected from cortical surface into volumetric MNI space and then the normalization parameters were used to project into each participant's individual diffusion data space for tractography.

To select the regions in the VVS which were most responsive to faces and places, as well as to determine if there was selectivity for an object-like tool region, the functional MRI localizers from the HCP project were used. The working memory *category-average* contrasts were used to select regions. Together these regions were the fusiform complex, the ventromedial visual area 2, and ventromedial visual area 3, respectively. The regions can be seen in Fig. 2a. To test if they were reliably selective

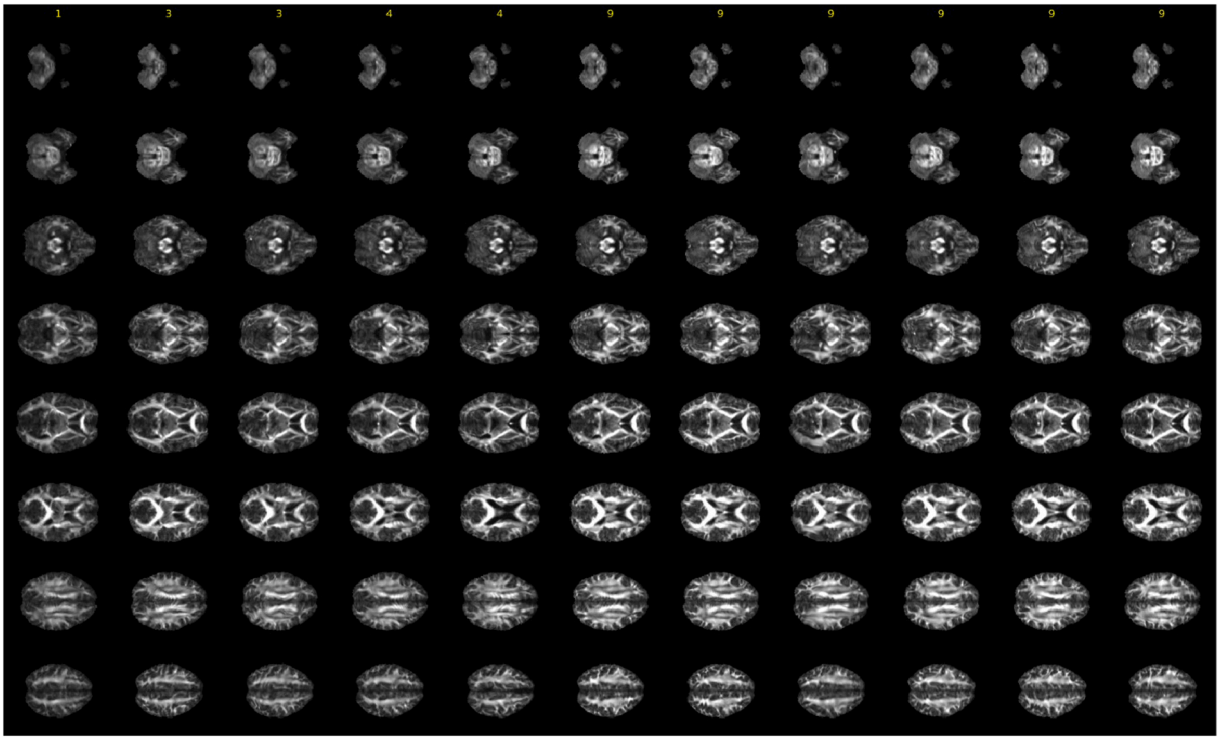


Fig. 1. Diffusion-weighted fractional anisotropy images for each of the 11 infants. Each column represents an individual infant and each row a volume. Numbers at the top denote the age of infants in months. There is good registration across the sample, illustrated by the similarity of each volume among infants.

for their chosen category (faces, places and tools), we examined functional responses from the HCP 1200 dataset. To avoid circularity between the definition of the parcels and the functional contrasts, we excluded the 210 participants used by Glasser et al. (2016) to create the parcellation, leaving N = 903. We tested if the regions displayed two kinds of selectivity. First, we tested if the regions reliably responded more to the preferred category than all of the other categories individually.

The face region responded more to faces than to places ($t(902) = 49.63$, $p < 0.0001$; 95 % CI [3.750, 4.066]) or to tools ($t(902) = 36.73$, $p < 0.0001$; 95 % CI [3.051 3.395]). The place region was more activated by places than tools ($t(902) = 53.39$, $p < 0.0001$; 95 % CI [4.146, 4.462] or faces ($t(902) = 81.90$, $p < 0.0001$; 95 % CI [8.187, 8.589316]). The "tool" region had a greater mean response to tools than faces ($t(902) = 78.83$, $p < 0.0001$; 95 % CI [6.258, 6.578]), but less to tools than to places ($t(902) = -28.76$, $p < 0.0001$; 95 % CI [− 2.255, − 1.967]), therefore exhibiting only a partially preferential response. Therefore, we were not able to identify a tool selective region.

Second, we tested if, for each functional contrast the three regions differed in their response. For the *face-average* contrast, the face region had a significantly higher mean contrast value than the place region ($t(902) = 84.41$, $p < 0001$; 95 % CI [6.239, 6.536]) or the "tool" region ($t(902) = 91.12$, $p < 0.0001$; 95 % CI [6.8897.192]). For the *place-average* contrast, the place region had greater activation than the "tool" region ($t(902) = 11.595$, $p < 0.0001$; 95 % CI [0.425, 0.599]) or the face region ($t(902) = 96.01$, $p < 0001$; 95 % CI [5.784, 6.026]). In the *tools-average* contrast, the "tool" region had greater activation than the place ($t(902) = 46.72$, $p < 0.0001$; 95 % CI [1.610, 1.752]) and the face region ($t(902) = 53.72$, $p < 0.0001$; 95 % CI [2.505, 2.695]).

In these selectively analyses we used the *category-average* contrasts. However, an alternative and equally valid contrast would have been the *category-implicit baseline*. In order to ensure the results generalize, we repeated the analysis and found similar results. The face region responded more to *faces-baseline* than *places-baseline* ($t(902) = 65.46$, $p < 0.0001$; 95 % CI [3.241, 3.441]) and again more to *faces-baseline* than *tools-baseline* ($t(902) = 41.76$, $p < 0.0001$; 95 % CI [2.196, 2.413]). Selectivity under this definition was also present for the place region, which responded more to *places-baseline* than *faces-baseline* ($t(902) = 79.77$, p-value < 0.0001; 95 % CI [5.148 5.407]) or *tools-baseline* ($t(902) = 50.10$, $p < 0.0001$; 95 % CI [2.668, 2.885]). Finally, when looking at the "tool" region, similar results are found as in the *category-average contrast*. The "tool" region displayed partial selectivity, with greater responses to the *tools-baseline* contrast than the *faces-baseline* contrast ($t(902) = 75.23$, $p < 0.0001$), but not to the *places-baseline* contrast ($t(902) = -21.35$, $p < 0.0001$; 95% CI [− 1.145, − 0.952]). Even with this new definition of selectivity, our face and place category-selective regions largely showed preference for their chosen category.

In summary, the face and the place regions defined though anatomical landmarks were found to have clear category selectivity. We acknowledge that these regions are different from the FFA and PPA as measured with functional contrasts in individuals, but the group level data from the 903 participants demonstrates clear category-level preferences, and, thus, these regions are category selective, even if their boundaries do not delimit exactly the most category selective voxels in each individual. In contrast, we were unable to identify a region that displayed a clear univariate response to tools. However, we continued to examine the connectivity of the region most selective for tools for two reasons. First, the region showing a partial preference was between the face and place selective regions (See Fig. 2A) and allowed us to characterize the medial to lateral gradient in the ventral stream, which is hypothesized to contain important dimensions of visual selectivity, including animacy and object size. Second, the region has a unique, multivariate response pattern, and this makes it a useful contrast with our univariate regions. More specifically, the region responded both to tools and places, and these categories evoke quite different patterns of activity in the place and tool regions, showing a strong stimulus by region interaction (Supplementary fig. 1). This indicates this region responds in a different way from the place region, motivating an independent analysis of its connectivity. As this region was not tool selective, we refer to it using its name from the Glasser et al. (2016)

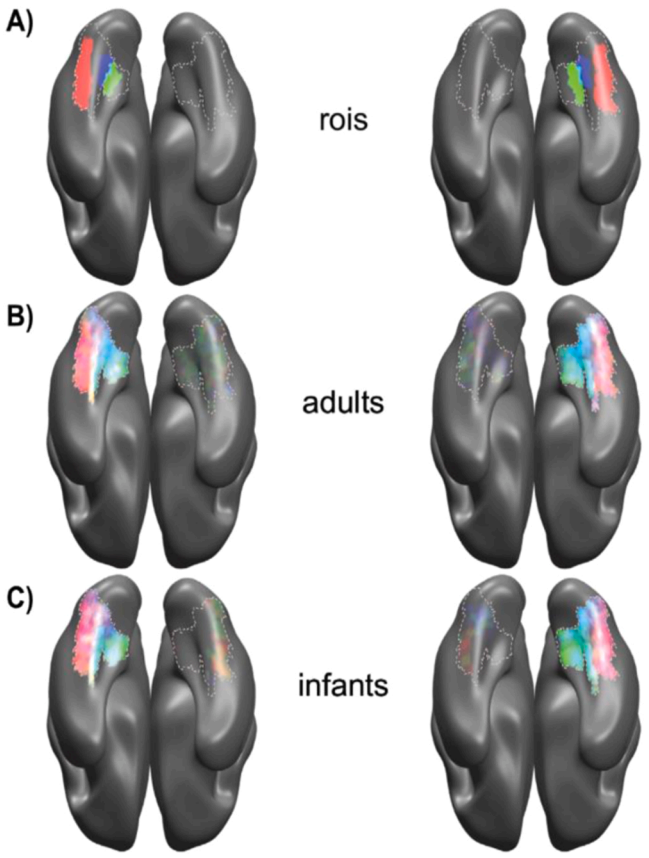


Fig. 2. A) Regions from HCP parcellation for faces, the VMV3 and for places (red, blue and green, respectively) in left and right hemispheres (left and right columns). Dotted outlines represent the VVS seed region, as defined by HCP (see methods) B) Group average overlay of voxels identified by a linear-discriminant classifier as selective for faces and places, and belonging to the VMV3 in adult participants (N = 14), based on their distinctive signature of structural connectivity with the rest of the brain. Classification was performed separately for the left and the right hemisphere, using leave-one-subject-out cross validation. The color mapping is the same as in (A). C) The distinctive signatures of structural connectivity were also present in infants (N = 11), as shown by voxels identified as category-selective by a linear-discriminant classifier trained on adult connectivity and tested in infants. Again the group average overlay is shown with the same color scheme as (A).

parcellation, ventromedial visual area 3 (VMV3), throughout the paper.

These regions were described and used as the category-selective regions in the subsequent classification analysis. The classification was performed for each voxel in the VVS, but membership in a category was defined using the contrasts from the HCP on a region basis and not voxelwise — we note there is no functional data in the adult or infant datasets described in the subsequent analyses. It should also be noted that this means that the definitions for the category selective regions were obtained from the Glasser et al. (2016) parcellation, choosing one region per category that was the most category-selective. From the current participants, only structural scans (for adult normalization) and diffusion-weighted images were used.

2.1.5. Tractography and classification

To implement probabilistic tractography, FSL's BEDPOSTX and PROBTRACKX, using 5000 streamlines per seed voxelinthe VVS, were run in each subject. The output of PROBTRACKX was then transformed to MNI space. These results were then summarized into a connectivity matrix that contained, for each voxel in the MNI ventral visual stream seed region the number of streamlines that terminated in each of the 346 target ROIs. A linear discriminant classifier was trained to distinguish face selective voxels from non-face selective voxels, based on the connectivity of each voxel. Two further classifiers were then trained to identify place and VMV3 voxels. For the adults, leave-one-subject out cross validation was then used to test whether selectivity could be predicted from connectivity for each of the three classifiers. For the infants, three classifiers trained on the full adult dataset were tested on each of the infants.

To summarize how sensitive the classifiers were in determining a voxel's category selectivity from its connectivity, we used signal detection theory, calculating values of d′ for each classifier across voxels in the seed region in each test participant. These values of d′ were then used to make comparisons between regions, infants and adults, and were correlated with age.

In a subsequent multiclass classification, a fitted discriminant analysis classifier was used to identify the fusiform complex, ventromedial visual area 2, and ventromedial visual area 3, as well as the non-category selective voxels in the HCP ventral stream visual cortex (region 4, supplementary neuroanatomical results) in adults. This was done based on voxel-wise connectivity with the 346 target regions. Leave-one-out cross validation was used to test classification accuracy, and values of d′ were calculated for each participant. For the infant version of this analysis, the fitted discriminant analysis classifier was trained on the entire adult dataset and tested on the infant data.

2.2. Results

2.2.1. HCP defined category-selective regions have distinctive signatures of connectivity in young infants

Voxels from the three regions could be robustly localized in adults based on their pattern of connectivity, with high values of d′ for the face (M = 1.77 SD = 0.25), place (M = 1.26, SD = 0.21) and VMV3 (M = 1.30, SD = 0.28) classifiers. This was confirmed using one sample t-tests (faces: $t(13) = 26.26$, $p < 0.001$ 95 % CI [1.64 1.9], places: $t(13) = 22.35$, $p < 0.001$ 95 % CI [1.15 1.37], VMV3: $t(13) = 17.17$, $p < 0.001$ 95 % CI [1.15 1.45]) (Fig. 2b). Classification performance is quantified in Fig. 3**a,** which shows values of d′ for classification of the imaging data. Classification performance was collapsed across hemispheres. Individual surface renderings that overlay voxels classified as being part of the face, place and VMV3 regions for adult participants are provided in Supplementary fig. 2.

Classifiers trained on adult data were then tested on infant data, identifying face (M = 1.74, SD = 0.24), place (M = 1.24, SD = 0.25), and VMV3 (M = 0.78, SD = 0.44) regions. The classifiers localized all three regions in infants (faces: $t(10) = 24.47$, $p < 0.001$ 95% CI [1.60 1.88], places: $t(10) = 16.54$, $p < 0.001$ 95% CI [1.10 1.39], VMV3 $t(10) = 5.95$, $p < 0.001$ 95% CI [1.04 0.52]) (Fig. 2c), using one sample t-tests. However, there was a category-by-group interaction ($F(2,46) = 6.64$, $p < 0.01$). Post-hoc tests showed this was because the face and place regions were as strongly detected in infants as they were in adults ($t(23) = 0.165$, $p > 0.05$, $t(23) = 0.257$, $p > 0.05$), but the VMV3 was detected with greater accuracy in adults than infants ($t(23) = 3.62$, $p < 0.01$) (Fig. 3b). Finally, in order to examine the developmental trajectory of the networks, infant age and classification accuracy (values of d′) were correlated– only the VMV3 network underwent significant change over the first 9 months of postnatal life (faces: $r(9) = -0.03$, CI= [− 0.64,0.82], p = 0.928; VMV3: $r(9) = 0.75$, CI= [0.34,0.97], p = 0.008; places: $r(9) = -0.01$, CI= [− 0.75, 0.52], p = 0.967) (Fig. 4).

As described in the Participants section above, this analysis contained two scans from one of the infants, separated by two months. The statistics above were therefore performed across sessions rather than strictly across participants. To ensure that our results took participant as a random variable, and therefore can be used to infer generalization to future participants, we excluded the second scan from this participant and repeated the t-tests and the correlation between infant age and classification accuracy (values of d′). Reassuringly, we found the same

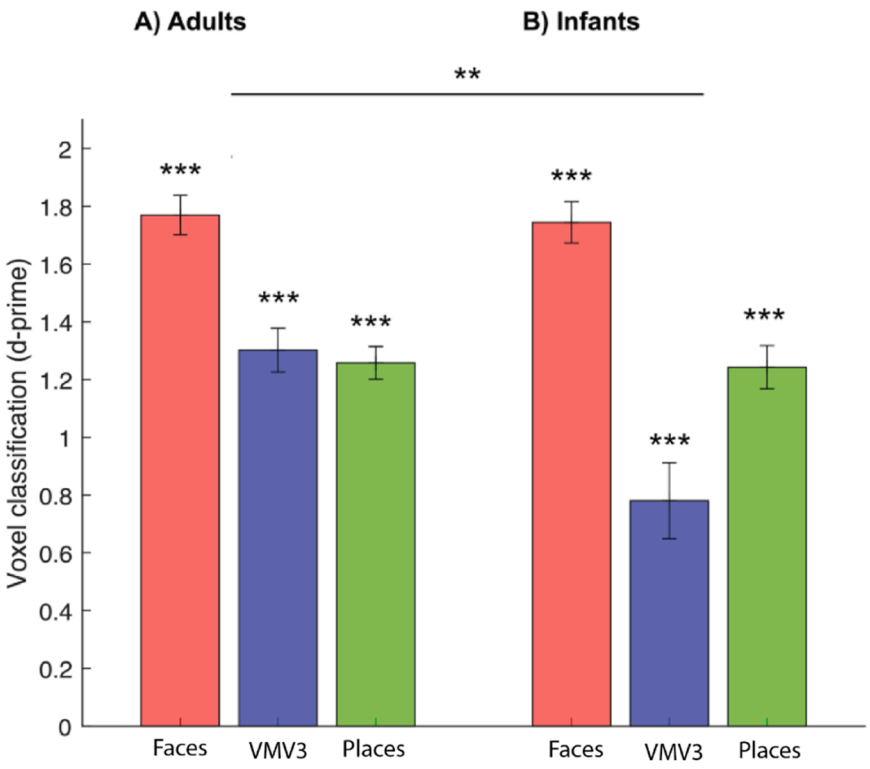


Fig. 3. a) Voxel classification performance for the adults (N = 14) quantified using d′, collapsed across hemispheres, using leave-one-out cross validation) b) Voxel classification performance for the infants (N = 11) measured using d′, and tested using one-sample t-tests. Models were trained on all adult data and tested in the infants. The three representative category-selective regions, taken from the Glasser et al. (2016) parcellation, were robustly localized in infants and adults but there was a significant difference in detection accuracy between the infant and adult VMV3 region, demonstrating the immaturity of connectivity for the VMV3 region during infancy. The mean + /- one standard error across subjects is shown. All results survive a correction for multiple comparisons across the three levels of the factor of category (face, place and VMV3).

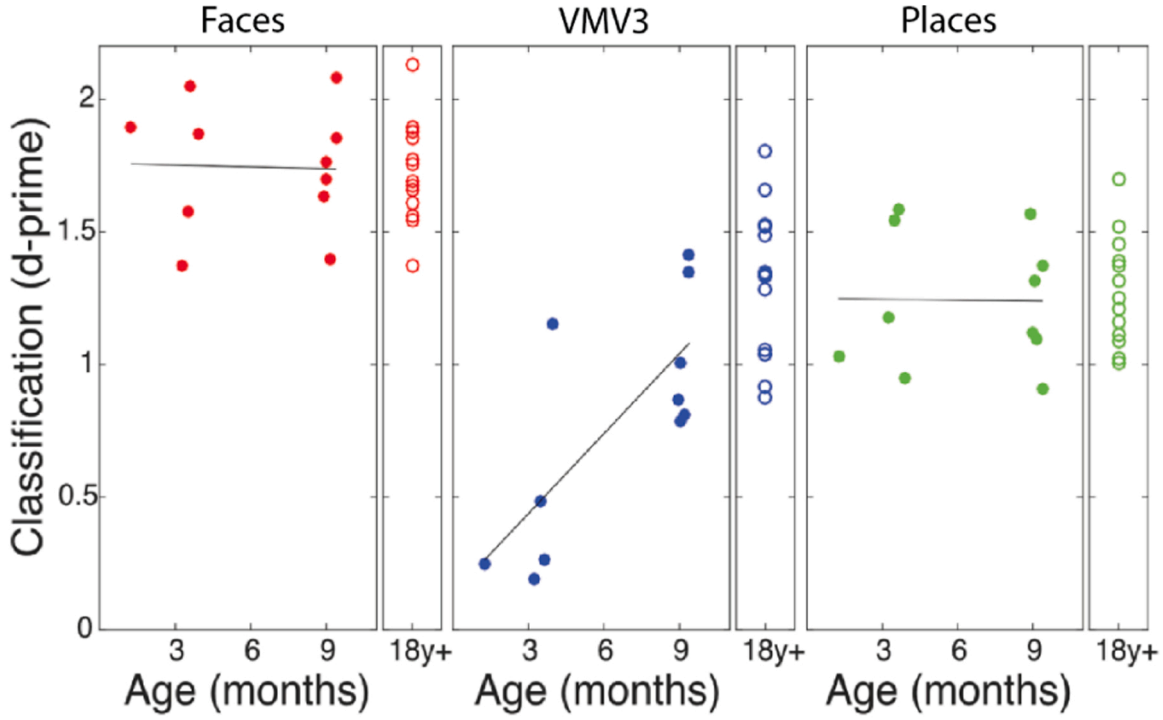


Fig. 4. The relationship between the age of participants (14 adults and 11 infants) and classification accuracy (d′) with best-fit lines. Only VMV3 classification had a significant relationship with age, demonstrating the maturation of the distinctive connectivity of the VMV3 network over the first year of postnatal life.

pattern of results. To visualize these results, we have included the same bar chart as in Fig. 3 and the same scatter plot as in Fig. 4 in Supplementary fig. 3. Again, classification accuracy was not correlated with age in the face ($r(8) = 0.11$, $p > 0.77$, 95 % CI [− 0.692,0.900] or place ($r(8) = 0.14$, $p > 0.71$, 95 % CI [− 0.748,0.610]) region. Classification in the VMV3 region continued to be correlated with age ($r(8) = 0.7$, $p < 0.019$, 95 % CI [0.113,0.978]. Additional statistics are in Supplementary fig. 3.

These results suggest the connectivity of the VMV3 region develops later than that of the face and place regions, but we also examined an alternative explanation. Could it be that tractography is more difficult in infants than adults, because of their lower signal-to-noise, and that

identification of voxels in the VMV3 region is more sensitive to this? To investigate this, first, in adults, the detection of the VMV3-selective voxels was no worse than detection of the place- selective voxels and performance was not at ceiling (Fig. 3A), suggesting that detection of the VMV3 region is not intrinsically more difficult. Second, we compared region size, which may affect performance more strongly in smaller infant brains; the place and VMV3 regions were the same size in one hemisphere and were less than 10 voxels different in the other. The hit rate for the place and VMV3 region was also not significantly different in adults ($t(13) = -1.04$, $p > 0.05$).

We next examined which target regions of connectivity most strongly influenced each of the classifiers in adults (see Supplementary fig. 4, and a rendering in Supplementary Figure 5), finding the place region had strong connectivity to regions like the hippocampus, the VMV3 to the 3rd visual area, involved in visuomotor transformations, and the face region to areas PGp and PGS, which are possibly connected via the inferior longitudinal fasciculus (Palejwala et al., 2020). Further description is provided in the SM.

The HCP VMV3 region was located in the cortex between the place and face regions. As demonstrated in the methods above, this region is not selective for tools, but is in a position in the medial to lateral ventral stream gradient that could contain small, inanimate objects. Although tool selectivity has been found before in this location (Chao et al., 1999) it is also present in other areas (Grill-Spector and Weiner, 2014) and may not be as selective as other tool regions when responding to tool vs other non-tool objects (Chen et al., 2018). From Fig. 2b, it is apparent that even in the adults, there is some blurring between category boundaries, particularly between the VMV3 and place regions. As the three classifiers were set up to each independently discriminate a single representative category selective region from all other voxels (including those that were selective for no category), these results cannot be used to quantify if pairs of categories can be distinguished from each other. To address this, we repeated classification, but with a fitted discriminant analysis classifier that allowed for multiclass classification (Guo et al., 2007). Using multiclass classification meant a single classifier aimed to predict whether a voxel was face, place, VMV3 or non-category selective. This confirmed that the three category-selective regions could be robustly discriminated from each other with even the smallest pairwise difference in d′, for the VMV3 region vs. places, reliable in adults ($t(13) = 7.05$, $p < 0.001$) and infants ($t(10) = 2.40$, $p < 0.05$). Values of d′ for the adult and infant multiclass classification analysis, for each category, can be found in Supplementary fig. 6.

3. Experiment 2

The results presented below show, using diffusion weighted imaging and tractography, that category-selective regions have distinctive patterns of connectivity with the rest of the brain in young infants, consistent with the proto-organization of category-selective regions seen by Kamps et al. (2020) with rsfMRI. When restricting our analysis to or excluding early visual regions, our results continue to extend the results of Kamps et al. (2020), and demonstrate a similar extended proto-organization emerging with multiple imaging modalities.

3.1. Methods

There are three limitations to Experiment 1. First, the sample size was small (N = 11). Second, there was only one infant who was younger than three months, by which time visual learning will have taken place. Third, the structural images from the infants in Experiment 1 were of insufficient quality to allow for normalization and the fractional anisotropy images were relied upon instead.

3.1.1. Participants

To address these limitations, we obtained diffusion weighted images from Developing Human Connectome Project (dHCP), using data from the second data release with 558 neonates, yielding 490 sessions of usable diffusion imaging data from 445 unique participants. In this experiment we analyzed the 400 infants with a single session. The distribution of ages at birth and at scan is shown in Fig. 5 and includes preterm infants.

3.1.2. Transformation of HCP ROIs to dHCP individuals

The seed and target ROIs were the same as in Experiment 1 and were transformed from the adult space to the native infant space. The dHCP release included the transformation from the individual diffusion image space to Bozek et al. (2018) 40 week template in volumetric space (https://gin.g-node.org/BioMedIA/dhcp-volumetric-atlas-groupwise), and was which applied using FSL's applywarp, nearest neighbor interpolation. ANTS, previously proven to be effective in pediatric studies (Avants et al., 2014; Jain et al., 2012; Lawson et al., 2013), was used to calculate the transformation from the skull-stripped adult T1-weighted template MNI-152 space.

In assessing registration, we found there were two problems. One was that the scalp in infant brains is thinner and closer to the cortex (see Supplementary fig. 7) and the adult cortex would register to this, and the second was the smaller cerebellum in neonates. We therefore removed the scalp using the background tissue types provided with the dHCP and removed the cerebellum with a mask using the AAL3 template (Tzourio-Mazoyer et al., 2002). The consequent ANTS registration is shown in Supplementary fig. 8. Further description can be found in the SM.

3.1.3. Connectivity signature analysis

As in Experiment 1, BEDPOSTX and FSL's PROBTRACKX were run using the above ROIs and the results were transformed back to MNI space via the 40 week template provided with the dHCP. Classification analyses could then be performed on the seed-to-target matrix in MNI space, which was common to all of the infant and adult participants. To test whether proto-organization in connectivity was present for just the visual cortex, or also other brain regions, we conducted two additional classification runs. In the first, target regions were restricted to only include those in the early visual system (networks 1,2, and 5 in the Glasser (2016) parcellation, roughly V1-V4 and MT+), and another with the inverse of the former, where target regions were all regions but excluded the early visual areas and ventral visual seed regions.

3.1.4. Gestational and scan age effects

To test whether there were effects of age-at-birth or age at scan, the

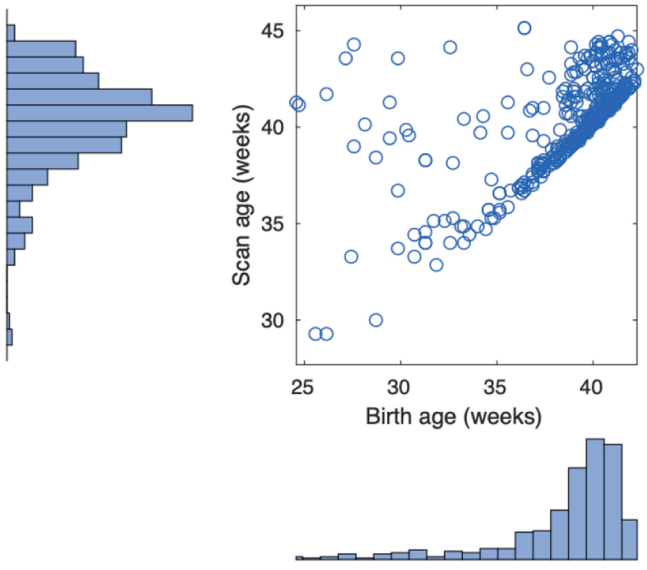


Fig. 5. The postmenstrual age at birth and at scan for the 400 infants included in Experiment 2. The modal postmenstrual age of birth is 40 weeks.

values d′ were calculated as described above for and excluding early visual regions. Then partial Spearman (rank) correlations were calculated for birth and scan age and values of d′ based on data that excluded or included targets in the early visual system. Spearman (rank) correlations were used to address the non-normality of the age measures. Bootstrapping was used to calculate confidence intervals.

3.2. Results

3.2.1. Category-selective regions have distinctive signatures of connectivity in neonates

Classifiers were able to localize all three regions in neonates, where performance was evaluated using one sample t-tests with values of d′ (Figs 6 and 7a, faces: M=1.53, SD=0.43, CI=[1.49, 1.57], $t(399) = 70.89$, $p < 10^{-6}$, Cohen's-d= 3.54, places: M= 1.07, SD= 0.43, CI= [1.03, 1.11], $p < 10^{-6}$, Cohen's-d= 2.49, $t(399) = 49.77$, VMV3: M= 0.68, SD= 0.50, CI= [0.63, 0.73], $t(399) = 27.39$, $p < 10^{-6}$, Cohen's-d= 1.37). We report the uncorrected results collapsed across hemispheres but each hemisphere was also significant. In contrast to our previous sample, faces and VMV3 had a lower classification accuracy in the neonates than the adults in both hemispheres, respectively. (faces: M_{adult}=1.80, SD_{adult}=0.45, M_{infant}=1.53, SD_{infant}=0.43, $t(412) = 2.32$ $p = 0.021$, CI(Δ)= [0.04,0.50], Cohen's-d= 0.63; places: M_{adult}= 1.80, SD_{adult}= 0.45, M_{infant}= 1.53, SD_{infant}= 0.43, $t(412) = 1.79$ $p = 0.074$, CI(Δ)= [− 0.02,0.43], Cohen's-d= 0.49; VMV3: M_{adult}= 1.27, SD_{adult}= 0.20, M_{infant}= 1.07, SD_{infant}= 0.43, $t(412) = 3.61$, $p = 0.0003$, CI(Δ)= [0.22,0.75], Cohen's-d= 0.98) (Fig. 7). This contrast with Experiment 1 could reflect less developed structural connectivity in neonates, or differences in scanners or sequences between the adult and neonate participants (Arichi et al., 2012). As in Experiment 1, we have provided a scatter plot of values of d′ and age (Supplementary Figure 9).

There is another potential confound to our analysis. In diffusion imaging there can be false positives due the streamlines terminating in neighboring regions. The analysis reported already removed the 14 ventral visual stream regions as targets, but in an additional analysis we took this a step further and removed all targets that were adjacent to any of the ventral visual stream regions. This removed areas like V1–4 and the hippocampus. After the removal of adjacent regions, all three regions could still be localized robustly in both adults and infants ($p < 0.001$). A figure depicting these results and complete statistics is available in Supplementary Figure 10. This result makes it unlikely that the localization of the regions is due to local connections and that the classifiers' performance is rather due to the connectivity of distal brain regions that are beginning to reflect adult-like category networks.

Finally, to demonstrate that the classifiers were picking up on distinct patterns of connectivity, we plotted the proportion of voxels in each ventral stream region that were classified as belonging to the face, place and VMV3 target areas. If the classifiers were accurate, for example, the FFC (our face target region) should have the highest proportion of voxels classified as belonging to the 'face' network. Reassuringly, this was the case for the three regions in adults and was broadly true in the neonates, with the exception of the VMV3, which had a lower classification accuracy than in the adults. Interestingly, qualitatively, the pattern of misclassifications also largely followed the degree of regional selectivity. Taken together, this result demonstrates that our classifiers are identifying distinct signatures of connectivity for each of our predefined regions. The plots for this analysis are found in columns two and three of Supplementary Figure 1.

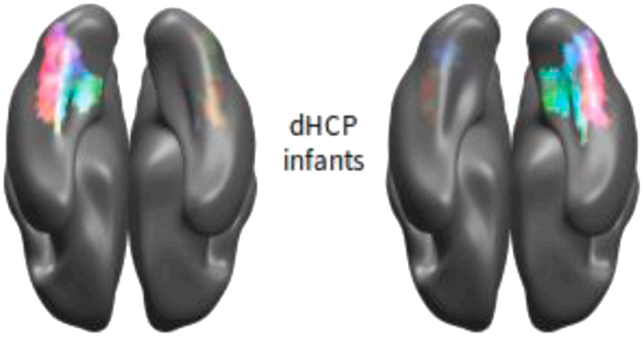


Fig. 6. Distinctive signatures of structural connectivity were also present in neonates (N = 400), as voxels could be identified as category-selective by a linear-discriminant classifier trained on adult connectivity and tested in infants. The projection is identical to Fig. 2 (with red for faces, blue for the VMV3, green for places).

3.2.2. Category selective connectivity to early visual regions vs. other brain regions

Our results are consistent with the results of Kamps et al. (2020) who found connectivity between category-selective regions and early visual regions to be established neonatally. However, Kamps et al. did not extend their analysis past the visual system. We, therefore, conducted two further analyses. These used the adult data from Experiment 1 to train the classifiers and infant data from Experiment 2 (from the dHCP) was used as the test population. In the first, we restricted the target regions to the early visual system (roughly V1-V4 and MT+) – specifically, V1-V4, Medial Superior Temporal area, Lateral Occipital area 1–3, Middle Temporal area, PH, V4t, FST, and V3CD, as defined by the Glasser parcellation (networks 1, 2, and 5 in the Glasser atlas (Glasser et al., 2016)). The three anatomically defined, category selective regions in both hemispheres could be clearly identified from their connectivity pattern to early visual regions (Fig. 7**b**) in adults (faces: M=1.70, SD=0.26, $t(13) = 24.96$, $p < 0.001$, CI= [1.56,1.85], Cohen's d= 6.67; places: M= 1.08, SD= 0.27, $t(13) = 15.08$, $p < 0.001$, CI= [0.93,1.24], Cohen's d= 4.03; VMV3: M= 1.52, SD= 0.27, $t(13) = 20.86$, $p < 0.001$, CI= [1.36, 1.67], Cohen's d= 5.58), and in infants (faces: M=1.57, SD=0.46, $t(399) = 68.07$, $p < 0.001$, CI= [1.53,1.62], Cohen's d= 3.40; places: M= 0.91, SD= 0.51, $t(399) = 35.95$, $p < 0.001$, CI= [0.86,0.96], Cohen's d= 1.80; VMV3: M= 0.82 SD = 0.64 $t(399) = 25.415582$, $p < 0.001$, CI= [0.75,0.88], Cohen's d= 1.27). These results remain significant after correcting for multiple comparisons, for 12 tests (2 ages (birth age, scan age) x 2 target subsets (exclude visual system, visual system only) x 3 categories (face, VMV3, place), and concur with Kamp's findings, extending them to tractography with diffusion weighted imaging.

In the third analysis, we selected the complement of target brain regions outside the early visual system (excluding the ventral seed regions). The three representative category selective regions in both hemispheres could again be clearly identified (Fig. 7**c**) in adults (faces: M=1.62, SD=0.22, $t(13) = 27.45$, $p < 0.001$, CI= [1.49, 1.75], Cohen's d= 7.34, places: M= 1.18, SD= 0.26, $t(13) = 17.08$, $p < 0.001$, CI= [1.03, 1.32], Cohen's d= 4.56; VMV3: M= 1.03 SD = 0.31 $t(13) = 12.42$, $p < 0.001$, CI= [0.85,1.20], Cohen's d= 3.32) and in infants (faces: M=1.54, SD=0.38, $t(399) = 79.97$, $p < 0.001$, CI= [1.50, 1.57], Cohen's d= 4.00; places: M= 1.09 SD = 0.39 $t(399) = 56.22$, $p < 0.001$, CI= [1.05,1.13], Cohen's d= 2.81; VMV3s: M= 0.67, SD= 0.55, $t(399) = 23.99$, $p < 0.001$, CI= [0.61,0.72], Cohen's d= 1.20. These results remain significant after correcting for multiple comparisons and show that anatomically defined category-selective connectivity with distal brain regions is present early in infancy and suggest it forms part of an innate proto-organization.

3.2.3. Effect of preterm birth suggests connectivity of the regions is innate

The previous result demonstrates the presence of a proto-organization containing both regions from the early visual system and long-range connectivity to the rest of the brain. However, the effect of visual experience was not measured. The varying gestational ages in the dHCP data allow us to examine whether representative category-selective networks are experience dependent or innate, developing without visual experience. If the maturation of these networks are

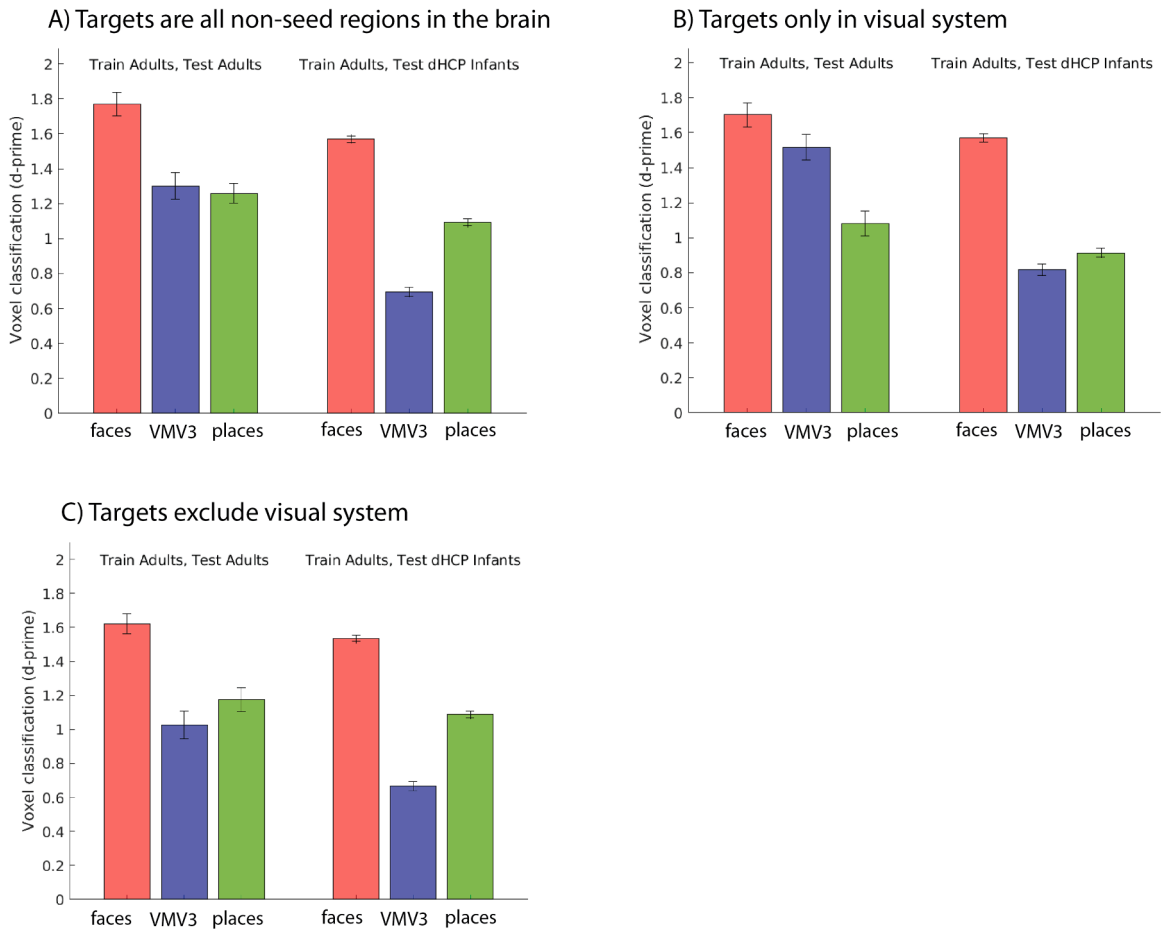


Fig. 7. Voxel classification performance quantified using d′ and tested with one sample t-tests. a) Target regions for tractography were the 346 brain regions not included in the Glasser et al. (2016) ventral visual stream. All three regions were detected robustly in both the adults and neonates. In contrast with our previous sample, both the VMV3 and face region were detected with significantly lower accuracy than adults. b) Target regions were restricted to only include those in the early visual system (networks 1,2, and 5 in the Glasser (2016) parcellation, roughly V1-V4 and MT+). All representative category-selective regions could be localized in the adults and infants, demonstrating a proto-organization evident in early life c) Target regions were the inverse of b), containing all regions but early visual areas and ventral visual seed regions. The regions were robustly localized in both the infants and the adults, demonstrating that long-range connectivity is present and part of the extended proto-organization in neonates.

experience dependent then infants who were born earlier will have greater visual experience and advanced development at the time of scan. In contrast, if the networks are innate, they will be less dependent on visual experience, and preterm birth, as it places challenges on the infant, may delay development.

The relationship between d′ and birth age, controlling for scan age, was tested using partial correlation with Spearman (rank) correlation to address the non-normality of the age measures (see distributions in Fig. 5). When considering the connections that excluded the early visual system, an uncorrected small positive effect was found for faces ($r(398) = 0.11$, $p = 0.022$, bootstrap CI= [0.02, 0.21]). An effect of similar magnitude was found for the VMV3 ($r(398) = 0.12$, $p = 0.020$, bootstrap CI= [0.02, 0.21]), but not for places ($r(398) = 0.03$, $p = 0.566$, bootstrap CI= [− 0.08, 0.14]). In contrast, when examining connections that only included the early visual regions, there was not a relationship between birth age (or gestational age) and classification accuracy for the faces ($r(398) = 0.03$, $p = 0.602$, CI= [− 0.07, 0.13]), places ($r(398) = 0.03$, $p = 0.489$, bootstrap CI= [− 0.06, 0.13]) or the VMV3 ($r(398) = 0.07$, $p = 0.156$, bootstrap CI= [− 0.02, 0.17]). None of these results survive a correction for multiple comparisons (12 tests).

When examining experience age (age at the scan - age at birth) and connectivity that excludes early visual areas, there was no relationship found for the faces ($r(398) = −0.09$, $p = 0.060$, bootstrap CI= [− 0.19, 0.00]), places ($r(398) = 0.10$, $p = 0.05$, bootstrap CI= [− 0.01, 0.21], or the VMV3 ($r(398) = −0.05$, $p = 0.317$, bootstrap CI= [− 0.15, 0.05]). In contrast, when considering only the connections to early visual regions, a relationship was found for places ($r(398) = −0.14$, $p = 0.005$, bootstrap CI= [− 0.23,− 0.05]), but not for faces ($r(398) = 0.00$, $p = 0.981$, bootstrap CI= [− 0.10, 0.10]) or the VMV3 ($r(398) = −0.08$, $p = 0.129$, CI= [− 0.17, 0.02]). The negative relationship for places could be consistent with part of the ROI being reallocated to some other category. But, none of these results survive a correction for multiple comparisons (12 tests) as so future replication would be informative.

4. Experiment 3

4.1. Method

These were the same as for Experiment 2– BEDPOSTX and FSL's PROBTRACKX were run in the infant's native space and the results were transformed back to MNI space via the 40-week template provided with the dHCP. Classification analyses could then be performed on the seed-to-target matrix in MNI space. However, a different group of participants were selected, who each had two scanning sessions, and longitudinal analysis was conducted to examine how connectivity to the representative category-selective regions changed. The longitudinal sample comprised $N = 45$ neonates from the dHCP, for which two scans were

acquired, separated by 6.8 + /- 2.4 weeks, with the postmenstrual age at the first scan of 34.5 + /- 1.7 weeks and at the second scan at around term-equivalent age 41.3 + /- 1.5 weeks. The only outcome measure of interest was the difference between the two scans in the d′ classification performance for each visual category collapsed across hemispheres. This was calculated using a paired t-test, and using a non-parametric Wilcoxon rank sum test.

4.2. Results

4.2.1. Longitudinal changes in category-selective connectivity

In the independent dataset of N = 45 neonates, using values of d′ and parametric t-tests, there was initial evidence for the strengthening face ($t(44) = 2.31$, $p < 0.026$, delta= −0.18 + /- 0.52 CI [− 0.33,− 0.02]), but not place ($t(44) = 1.65$, $p = 0.107$, delta= −0.13 + /- 0.55 CI [− 0.30,0.03]) or VMV3 connectivity ($t(44) = -0.26$, $p = 0.792$, delta= 0.03 + /- 0.69 CI [− 0.18,0.24]).

However, the data from the face network violated the normality assumption, and when reassessed with a non-parametric Wilcoxon rank sum test, there was no evidence of developmental change between the two scans (faces Wilcoxon rank sum =1921, p = 0.309). The same test was then repeated for places and the VMV3 and there was no evidence for the strengthening of connections between the scans (places Wilcoxon rank sum =1899 p = 0.232; VMV3 Wilcoxon rank sum =2030 p < 0.89). Values of d′ for both scans, separately for faces, the VMV3, and places can be found in Fig. 8.

5. Discussion

Key results of the three experiments are summarized in Table 1. Category-selective regions, defined through anatomical correspondence with adults, had distinct patterns of connectivity early in life, found to be both innate and shaped by experience. Connectivity of the region-of-interest with the strongest face selectivity in adults was the most strongly innate as it was present in neonates in both Experiment 1 and 2, and was shaped by their age at birth, suggesting it develops *in utero* prior to visual experience of faces and is slowed by preterm birth. There was no evidence that the connectivity of the face region became more distinctive with postnatal experience. Place network connectivity was also present in early life, but preliminary evidence points toward a potential role for early postnatal experience. In contrast, while the network associated with the VMV3 region was found to be disrupted by preterm birth, there was no effect of neonatal experience (Experiment 2) but preliminary evidence demonstrated a protracted development across the first 9 months of postnatal life (Experiment 1).

Category selectivity was defined at the region-of-interest level, using group-average fMRI contrasts from a large set of adults in the HCP, which differs from defining category selectivity at the voxelwise level using functional localizers in individual subjects. We did this because the infants in our study and the dHCP were scanned asleep with their eyes closed, and therefore it was not possible for them to see visual stimuli. Furthermore, given the differences between adult and infant brains, registration may not have always been precise to the voxel level. Our study cannot therefore assess the selectivity of the specific subset of most selective voxels within each individual (Osher et al., 2016). Rather, it provides the "big picture" of connectivity of the regions in the approximate anatomical location of the most selective face and place, and complementary VMV3 region in adults.

The proto-organization previously found for the ventral stream and early visual regions was extended to include long-range connections with distal brain areas making up the adult unique signature of connectivity. The neonates, who had a modal postmenstrual age of 40 weeks, had adult-like connectivity patterns for distal brain regions excluding early visual regions. This evidence, in combination with the other results above, strongly suggests that the long-range connectivity of anatomically defined category-selective regions is not developed by a prolonged statistical learning of associations, but is instead largely innate, making up part of an extended proto-organization that had previously only been seen with early visual regions. Additionally, when looking at the effect of visual experience and birth age, only long-range connectivity was disrupted by preterm birth (for faces and the VMV3), showing that within system network connectivity is less likely to be disrupted than between system network connectivity by early birth. The importance of long-range connections in semantics was recently demonstrated by multivariate decoding of white matter pathways in brain-injured patients with semantic deficits (Fang et al., 2018). It should be noted that some of these correlations are based on small sample sizes (e.g. N = 11 in 1–9-month-old infants) or did not hold up to multiple comparisons correction, and future work should seek to replicate the correlational age effects demonstrated here. These preliminary

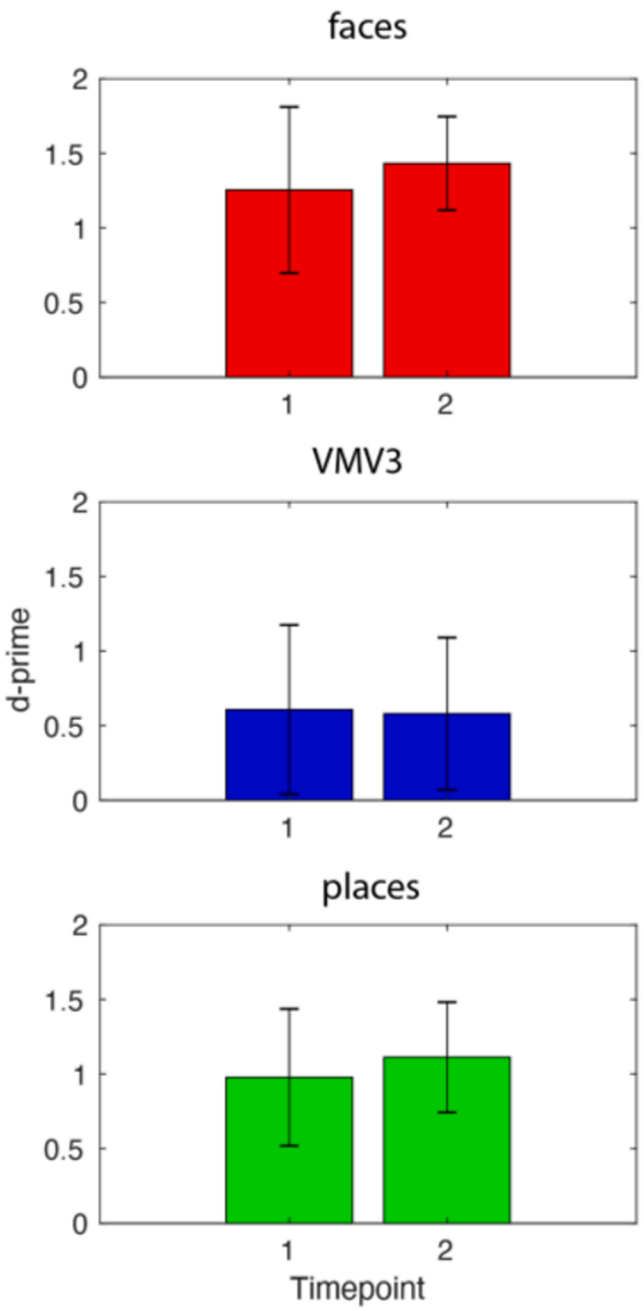


Fig. 8. Voxel classification performance quantified using d′ and tested with Wilcoxon rank sum tests. There was no evidence of developmental change for either the face (red), VMV3 (blue) or green (place) networks.

Table 1
Factors affecting category-selective connectivity.

	Exp. 1 1–9 months (N = 11, 1 scan)		Exp. 2 Neonates (N = 400, 1 scan)			Exp. 3 Neonates (N = 45, 2 scans)
	Presence	Effect of age at scan	Presence	Effect of preterm birth	Effect of experience age	Longitudinal effect of age
Faces	*** (***)	(-)	*** (***)	(*)	(-)	(-)
Places	*** (***)	(-)	*** (***)	(-)	(**)	(-)
VMV3	*** (***)	* (**)	*** (***)	(*)	(-)	(-)

Note: This table illustrates the findings from Experiment 1–3. Uncorrected significance levels are represented by * 's, and are italicized within brackets. * = < 0.05, ** = <0.01, and *** = < 0.001. In Experiment 1, Bonferroni corrections were performed for three tests (face, place and the VMV3). In Experiment 2, corrections were performed for 12 tests (2 (birth age, scan age) x 2 (exclude visual system, visual system only) x 3 (face, VMV3, place)). Corrected statistics levels are detonated by the * in a larger font without brackets, and follow the same convention as above, where * = < 0.05, ** = <0.01, and *** = < 0.001 at corrected levels. Presence of category-selective connectivity was found in all Experiments and in all regions and remained significant after multiple comparison correction. However, the VMV3 experienced postnatal change, while both face and VMV3 regions were found to be affected by preterm birth. The place network was found to be affected by postnatal experience. In Experiment 2, the relationships between age and connectivity did not survive a correction for multiple comparisons.

results extend those from fMRI (Deen et al., 2017) and functional connectivity (Kamps et al., 2020) in infants and structural connectivity in adults (Saygin and Kanwisher, 2014) to reveal a proto-organization of category-selective regions in the ventral visual stream. The early maturity and specificity of the structural networks is further reflected in the connections identified by the classifiers. For example, many of the strongest connections for the place region are part of the place processing network—the hippocampus, entorhinal cortex and parahippocampal regions (see SM) (Epstein, 2008) However, although we find consistent evidence of network maturity, there is substantial research that microstructure associated with these networks continues to mature (Gomez et al., 2017).

Although the results suggest that the face specific network matures early, even *in utero*, a broadly mature structure does not imply that visual experience is not necessary to acquire a functional face representation. Work in monkeys and with individuals who had cataracts as children demonstrate the need for experience to fully develop face processing capabilities, and infants under a month spend an estimated quarter of their waking hours with faces taking up the majority of their visual field (Arcaro et al., 2017; Jayaraman et al., 2015; Le Grand et al., 2001). Reid et al. (2016) showed that even fetuses will preferentially orient to face-like patterns in the third trimester, demonstrating a system that is already tuned to face-like input, ready to identify opportunities to refine instead of fully acquire face representations in the ventral visual stream. However, these results have generated controversy (Scheel et al., 2018), and some have argued that a longer period of experience would be needed to induce the face preferences Reid et al. tried to demonstrate. Place connectivity was affected by age at scan (postnatal age, Experiment 2) but did not show maturation in Experiment 1. This suggests that the broad structure that underlies the place network might show early, substantial refinement in postnatal life that does not continue across the first year. Young infants get substantial experience viewing scenes (and due to their long supine hours, perhaps particularly ceilings). Scene representations in the ventral visual stream are biased towards the periphery, and the retinal temporal hemifield, representing the periphery, develops before the nasal hemifield (representing more foveal representations) (Grill-Spector and Weiner, 2014; Lewis and Maurer, 1992) The place-selective network was found to contain structures like the entorhinal cortex, which is involved both in navigation and memory. It is likely that the complexity of creating multiple objects in space to form a coherent representation substantial enough to navigate requires a period of visual experience, and that early experience in the neonatal period might be substantial enough to form these connections.

The network associated with VMV3 was found to change with experience and was disrupted by preterm birth. Possibly reflective of this late, experience driven maturation, this region did not show perfect selectivity in the adult HCP participants, and its multivariate nature may mean that maturation continues across the first 9 months of postnatal life. However, we would caution overinterpreting the results based on the limited sample size in Experiment 1.

As the VMV3 region exhibited the strongest tool-average contrast (See Fig. 1, first column in blue) and its location in the ventral stream, we consider its connectivity in terms of the development of tool and object selectivity and representation. Young infants are perhaps less likely to see tools than faces and places. Infants younger than three months have primarily frontal views of faces, while older infants at 8–10 months have a mixture of body parts from nearby caregivers, some of which are acting on objects (Bambach et al., 2018; Fausey et al., 2016). These infants see a variety of objects at mealtimes and it has been suggested that the presence of these objects in the environment allow infants of the same age to look towards objects when they are named (Yu et al., 2019). These infants are able to sit up, crawl, and increasingly manipulate objects, although their tool-use behavior is not as sophisticated as that of toddlers.

An alternative interpretation of our results is that we have demonstrated a broader gradient that follows the lateral-medial distinction in the temporal cortex and may relate to the gradients proposed for object size, animacy or retinotopic organization. While we did not model the entire ventral visual stream in a multivariate way, our three regions represent a medial-to-lateral gradient, and we selected the VMV3 region partially on the basis that it was between the face and place region. Our results may reflect an animacy distinction, as would any identification of a face, place, medial tool or object region, and the same is true for retinotopy and object size, as these gradients are superimposed on top of category selectivity. Category-selective regions are part of the broad gradients that define the ventral visual stream.

Additionally, Glasser et al. (2016) acknowledge that the regions in their parcellation are not identical to those in previous work, citing both methodological and sample size differences. For example, previous research has found the face region to be more patchy, but Glasser's contrasts identified a strip-like region. This may be a result of sample size, the functional contrasts in the HCP, or myelin and other structural data used to define the regions, but it does not mean that these regions are overall less category-selective.

Using structural data and linear discriminant classifiers to identify connections that were most predictive of whether a voxel was part of a category-selective region was done for the first time in infants. While the connectivity patterns were selected by the classifiers, this method is fitting for answering hypothesis driven questions. The regions used were determined by independent, apriori data and hypothesis, and others have already shown these individual patterns of connectivity, generated from parcellations, can predict face, place, and object selectivity in the areas of cortex studied here (Osher et al., 2016). Therefore, the results

shown speak to the development of this aspect of category-selectivity, demonstrating the maturity of long range networks across the first year of postnatal life.

A limitation is the lack of functional localizers and diffusion data in the same participants, which would allow us to more accurately identify individual patterns of selectivity, and, in infants, whether they have functional patterns that are different from adults. Completing the analysis with functional contrasts would answer the more exploratory question to which degree to which voxels in the infant ventral stream are category-selective, which would then need to be compared to adults. It may be challenging to determine the exact correspondence between the groups, as it's not clear if there is a one-to-one mapping of functional responses across development see (Cusack et al., 2018) for discussion.

The identification of individual patterns of selectivity is still an important question. Future research may be able to more closely identify the trajectory of network maturation of category-selective regions with a comprehensive longitudinal study. One advantage to using the regions from the HCP is their generalizability between adult participants, something that would be challenging to do in a local sample. The HCP region definitions are also based on multiple types of data (structural, functional and diffusion data), which would also be challenging to acquire locally in large numbers. Nevertheless, these efforts would be worthwhile and could help answer outstanding questions.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We thank the families and the radiographers, both in London, ON and those who participated in the Developing Human Connectome Project (ERC Grant Agreement no. 319456) for their contribution. We are grateful for both our Canadian (NSERC/CIHR CHRP(201110CPG)) and European (ERC Advanced Grant ERC-2017-ADG, FOUNDCOG, 787981) funding sources, as well as the grants that supported Dr. Cabral's Postdoctoral Training (5T15 LM007059–35 Pittsburgh Biomedical Informatics Training Program; T32MH018951 Child & Adolescent Mental Health Research Pittsburgh). We would like to acknowledge the many stimulating conversations we have had with colleagues, including our reviewers, surrounding this work. Science is at its best when there is intellectual debate and we can equally contribute to progress, and their feedback has moved this effort forward.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.dcn.2022.101179.

References

Almeida, J., Fintzi, A.R., Mahon, B.Z., 2013. Tool manipulation knowledge is retrieved by way of the ventral visual object processing pathway. Cortex 49, 2334–2344.
Arcaro, M.J., Livingstone, M.S., 2017. A hierarchical, retinotopic proto-organization of the primate visual system at birth. eLife 6.
Arcaro, M.J., Livingstone, M.S., 2021. On the relationship between maps and domains in inferotemporal cortex. Nat. Rev. Neurosci. 22, 573–583.
Arcaro, M.J., Schade, P.F., Vincent, J.L., Ponce, C.R., Livingstone, M.S., 2017. Seeing faces is necessary for face-domain formation. Nat. Neurosci. 20, 1404–1412.
Arichi, T., et al., 2012. Development of BOLD signal hemodynamic responses in the human brain. Neuroimage 63, 663–673.
Avants, B.B., et al., 2014. The Insight ToolKit image registration framework. Front. Neuroinform. 8, 44.
Bambach, S., Crandall, D.J., Smith, L.B., Yu, C., 2018. Toddler-inspired visual object learning. In: Proceedings of the 32nd Conference on Neural Information Processing Systems.
Belin, P., Zatorre, R.J., Lafaille, P., Ahad, P., Pike, B., 2000. Voice-selective areas in human auditory cortex. Nature 403, 309–312.
Binder, J.R., Desai, R.H., Graves, W.W., Conant, L.L., 2009. Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. Cereb. Cortex 19, 2767–2796.
Bozek, J., et al., 2018. Construction of a neonatal cortical surface atlas using multimodal surface matching in the developing human connectome project. Neuroimage 179, 11–29.
Chao, L.L., Haxby, J.V., Martin, A., 1999. Attribute-based neural substrates in temporal cortex for perceiving and knowing about objects. Nat. Neurosci. 2, 913–919.
Chen, Q., Garcea, F.E., Jacobs, R.A., Mahon, B.Z., 2018. Abstract representations of object-directed action in the left inferior parietal lobule. Cereb. Cortex 28, 2162–2174.
Coccia, M., Bartolini, M., Luzzi, S., Provinciali, L., Ralph, M.A.L., 2004. Semantic memory is an amodal, dynamic system: evidence from the interaction of naming and object use in semantic dementia. Cogn. Neuropsychol. 21, 513–527.
Cohen, L., et al., 2002. Language-specific tuning of visual cortex? Functional properties of the visual word form area. Brain 125, 1054–1069.
Cusack, R., et al., 2014. Automatic analysis (aa): efficient neuroimaging workflows and parallel processing using Matlab and XML. Front. Neuroinform. 8, 90.
Cusack, R., McCuaig, O., Linke, A.C., 2018. Methodological challenges in the comparison of infant fMRI across age groups. Dev. Cogn. Neurosci. 33, 194–205.
Deen, B., et al., 2017. Organization of high-level visual cortex in human infants. Nat. Commun. 8, 13995.
Dehaene, S., Cohen, L., 2011. The unique role of the visual word form area in reading. Trends Cogn. Sci. 15, 254–262.
Epstein, R., Kanwisher, N., 1998. A cortical representation of the local visual environment. Nature 392, 598–601.
Epstein, R.A., 2008. Parahippocampal and retrosplenial contributions to human spatial navigation. Trends Cogn. Sci. 12, 388–396.
Fang, Y., et al., 2018. Semantic representation in the white matter pathway. PLOS Biol. 16, e2003993.
Fausey, C.M., Jayaraman, S., Smith, L.B., 2016. From faces to hands: changing visual input in the first two years. Cognition 152, 101–107.
Glasser, M.F., et al., 2016. A multi-modal parcellation of human cerebral cortex. Nature 536, 171–178.
Gomez, J., et al., 2017. Microstructural proliferation in human cortex is coupled with the development of face processing. Science 355, 68–71.
Gomez, J., Barnett, M., Grill-Spector, K., 2019. Extensive childhood experience with Pokémon suggests eccentricity drives organization of visual cortex. Nat. Hum. Behav. https://doi.org/10.1038/s41562-019-0592-8.
Goodale, M.A., Milner, A.D., 1992. Separate visual pathways for perception and action. Trends Neurosci. 15, 20–25.
Grill-Spector, K., Weiner, K.S., 2014. The functional architecture of the ventral temporal cortex and its role in categorization. Nat. Rev. Neurosci. 15, 536–548.
Guo, Y., Hastie, T., Tibshirani, R., 2007. Regularized linear discriminant analysis and its application in microarrays. Biostatistics 8, 86–100.
Huth, A.G., de Heer, W.A., Griffiths, T.L., Theunissen, F.E., Gallant, J.L., 2016. Natural speech reveals the semantic maps that tile human cerebral cortex. Nature 532, 453–458.
Jain, V., et al., 2012. Longitudinal reproducibility and accuracy of pseudo-continuous arterial spin-labeled perfusion MR imaging in typically developing children. Radiology 263, 527–536.
Jayaraman, S., Fausey, C.M., Smith, L.B., 2015. The faces in infant-perspective scenes change over the first year of life. PLOS One 10, e0123780.
Kamps, F.S., Hendrix, C.L., Brennan, P.A., Dilks, D.D., 2020. Connectivity at the origins of domain specificity in the cortical face and place networks. Proc. Natl. Acad. Sci. USA 117, 6163–6169.
Kanwisher, N., Yovel, G., 2006. The fusiform face area: a cortical region specialized for the perception of faces. Philos. Trans. R. Soc. Lond. B Biol. Sci. 361, 2109–2128.
Kanwisher, N., Woods, R.P., Iacoboni, M., Mazziotta, J.C., 1997. A locus in human extrastriate cortex for visual shape analysis. J. Cogn. Neurosci. 9, 133–142.
Lawson, G.M., Duda, J.T., Avants, B.B., Wu, J., Farah, M.J., 2013. Associations between children's socioeconomic status and prefrontal cortical thickness. Dev. Sci. 16, 641–652.
Le Grand, R., Mondloch, C.J., Maurer, D., Brent, H.P., 2001. Neuroperception. Early visual experience and face processing. Nature 410, 890.
Lerma-Usabiaga, G., Carreiras, M., Paz-Alonso, P.M., 2018. Converging evidence for functional and structural segregation within the left ventral occipitotemporal cortex in reading. Proc. Natl. Acad. Sci. USA 115, E9981–E9990.
Lewis, T.L., Maurer, D., 1992. The development of the temporal and nasal visual fields during infancy. Vis. Res. 32, 903–911.
Li, J., Osher, D.E., Hansen, H.A., Saygin, Z.M., 2020. Innate connectivity patterns drive the development of the visual word form area. Sci. Rep. 10, 18039.
Mahon, B.Z., Caramazza, A., 2011. What drives the organization of object knowledge in the brain? Trends Cogn. Sci. 15, 97–103.
Mahon, B.Z., Anzellotti, S., Schwarzbach, J., Zampini, M., Caramazza, A., 2009. Category-specific organization in the human brain does not require visual experience. Neuron 63, 397–405.
Malach, R., et al., 1995. Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. Proc. Natl. Acad. Sci. USA 92, 8135–8139.
Martin, A., 2016. GRAPES-Grounding representations in action, perception, and emotion systems: How object properties and categories are represented in the human brain. Psychon. Bull. Rev. 23, 979–990.
McKone, E., Crookes, K., Jeffery, L., Dilks, D.D., 2012. A critical review of the development of face recognition: experience is less important than previously believed. Cogn. Neuropsychol. 29, 174–212.

Meyer, K., et al., 2010. Predicting visual stimuli on the basis of activity in auditory cortices. Nat. Neurosci. 13, 667–668.
Osher, D.E., et al., 2016. Structural connectivity fingerprints predict cortical selectivity for multiple visual categories across cortex. Cereb. Cortex 26, 1668–1683.
Palejwala, A.H., et al., 2020. Anatomy and white matter connections of the fusiform gyrus. Sci. Rep. 10, 13489.
Patterson, K., Nestor, P.J., Rogers, T.T., 2007. Where do you know what you know? The representation of semantic knowledge in the human brain. Nat. Rev. Neurosci. 8, 976–987.
Peelen, M.V., Downing, P.E., 2017. Category selectivity in human visual cortex: beyond visual object recognition. Neuropsychologia 105, 177–183.
Reid, V.M., et al., 2016. The human fetus preferentially engages with face-like visual stimuli. Curr. Biol. 27, 1825–1828 e3.
Saygin, Z., Kanwisher, N., 2014. Structural and functional connectivity fingerprints for face, body, scene, and object perception. J. Vis. 14 (603–603).
Scheel, AM, Ritchie, SJ, Brown, NJL, Jacques, SL, 2018. Methodological problems in a study of fetal visual perception. Curr Biol. 28 (10), R594–R596. https://doi.org/10.1016/j.cub.2018.03.047. PMID: 29787718.
Smith, S.M., et al., 2004. Advances in functional and structural MR image analysis and implementation as FSL. Neuroimage 23 (Suppl 1), S208–S219.
Smith, S.M., et al., 2006. Tract-based spatial statistics: voxelwise analysis of multi-subject diffusion data. Neuroimage 31, 1487–1505.
Tzourio-Mazoyer, N., et al., 2002. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. Neuroimage 15, 273–289.
Yu, C., Suanda, S.H., Smith, L.B., 2019. Infant sustained attention but not joint attention to objects at 9 months predicts vocabulary at 12 and 15 months. Dev. Sci. 22, e12735.