Category-Selective Visual Regions Have Distinctive Signatures of Connectivity in Early Infancy

Laura Cabral* [1], Leire Zubiaurre [2], Conor Wild [1] Annika Linke [3] Rhodri Cusack [4]

[1] Brain and Mind Institute, Western Interdisciplinary Research Building, Western University

London, Ontario, Canada N6A 3K7

[2] Department of Methods and Experimental Psychology, Faculty of Psychology and Education,

University of Deusto, Bilbao, Spain, 48007.

[3] Brain Development Imaging Laboratories, San Diego State University, San Diego,

California, USA, 92120.

[4] Trinity College Institute of Neuroscience, Trinity College Dublin, College Green

Dublin 2, Ireland.

Corresponding Author:
Laura Cabral
Brain and Mind Institute
Western Interdisciplinary Research Building
Western University
London, Ontario, Canada
N6A 3K7
lcabral@uwo.ca

Running head: Maturity of connectivity in category-selective visual regions

Subject areas: Neuroscience, Developmental Biology

Maturity of connectivity in category-selective visual regions

Abstract

By four months, infants can form categories of similar looking objects, but it is unclear when they begin to make the rich cross-modal, motoric and affective associations that are characteristic of adult visual categories. These associations are thought to be encoded by long-range brain connectivity and are reflected in the distinctive signature of connectivity of each category-selective region in the ventral visual stream. Category-selective ventral visual regions are already functioning in young infants, but their long-range connectivity has not been investigated. Therefore we used MRI diffusion tractography to characterize the connectivity of face, place and tool regions in 1-9 month infants. Using a linear discriminant classifier, we found that the face and place regions had adult-like connectivity throughout infancy, but the tool-network underwent significant maturation until 9 months. This suggests that young infants might already be forming rich associations, but that different categories are maturing with different developmental trajectories.

Key words: infant, vision, neuroimaging, diffusion MRI, tractography, machine learning

Introduction

Forming categories is a core part of human development as it allows us to improve our inferences about the environment and make better predictions. At 3-4 months old, infants are able to form visual, perceptual categories (e.g. cats vs. dogs) (Eimas & Quinn, 1994; Quinn, Eimas, & Rosenkrantz, 1993). If they are presented with a series of cats, they will preferentially orient towards a subsequently presented dog rather than another cat, which shows that they can form a stable representation of the distribution of perceptual features of the cat stimuli (Sloutsky & Fisher, 2004; Wilcox, 1999) and detect deviations from the category. The visual processing that leads to these perceptual categories is thought to be located in the ventral occipital and temporal cortex. In adults, distinct regions are selective for perceptual categories such as faces, places or objects (Epstein, 2008; Epstein & Kanwisher, 1998; Kanwisher & Yovel, 2006; Kanwisher, Woods, Iacoboni, & Mazziotta, 1997; Malach et al., 1995). Functional neuroimaging has found that at 4-6 months infants already have regions that are selective for perceptual categories in the ventral temporal cortex as well (Deen et al., 2017).

In adults, the pattern of connectivity to other brain areas is distinct for each category-selective region in the ventral temporal cortex (Osher et al., 2016; Saygin & Kanwisher, 2014). These long-range connections are thought to encode the cross-modal, motoric and affective associations characteristic of rich semantic categories (Binder, Desai, Graves, & Conant, 2009; Huth, de Heer, Griffiths, Theunissen, & Gallant, 2016; Patterson, Nestor, & Rogers, 2007). As concrete examples, seeing a silent video of a dog barking evokes the representation of its sound in auditory cortex (Meyer, Kaplan, Essex, Webber, Damasio & Damasio, 2010), and for tools and objects, category representations in the ventral stream are integrated with action representations in the dorsal stream (Almeida, Fintzi, & Mahon, 2013; Coccia, Bartolini, Luzzi, Provinciali, & Ralph, 2004; Goodale & 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).

It has not been established when infants begin go beyond visual, perceptual categories to make rich associations. It is difficult to measure, as conceptual understanding is usually probed verbally, but infants in the first year have very limited language. Could they already, however, be forming associations through a process of unsupervised statistical learning? In this study we use neuroimaging to investigate this question, by examining the maturity of structural connectivity of

category-selective regions in infants. To measure connectivity, we used diffusion-weighted imaging and tractography. We extracted the characteristic signatures of connectivity of three category-selective regions in adults using a machine learning approach, and then tested for generalisation to infants.

Results

In order to measure structural connectivity, diffusion-weighted images were obtained from 14 adults and 11 infants. Probabilistic tractography was performed using seed and target regions taken from the parcellation by the Human Connectome Project (HCP) (Glasser et al., 2016). Each voxel in the ventral visual stream, as defined by the HCP (see Methods), was used as a seed, while the brain areas outside the ventral stream acted as targets for tractography. To define category-selective regions in the ventral stream, contrast maps from the HCP fMRI localizers were used to determine the regions within the parcellation that were most selective for face, places and tools. These regions were the fusiform complex, the ventromedial visual area 2, and ventromedial visual area 3, respectively. Regions of interest are displayed in **Figure 1a.**
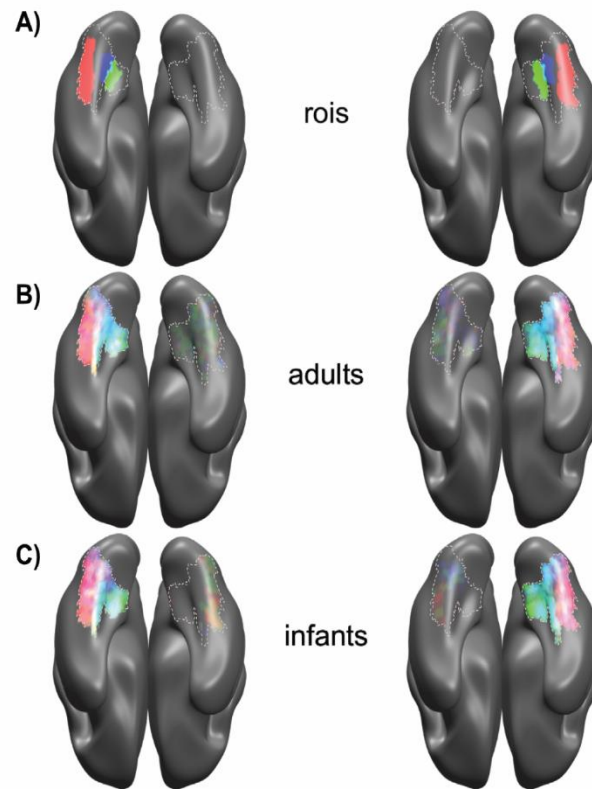
Figure 1. This figure represents the regions used for classification and group overlays of the voxels selected by the classifiers as part of the face, place, and tool regions of interest based on their structural connectivity with other brain regions. A) Regions from the HCP parcellation that were most selective for faces, tools and places (red, blue and green, respectively) in the left and right hemispheres (left and right columns). Dotted outlines represent the ventral stream seed region, as defined by the HCP (see methods) B) Voxels identified by a linear-discriminant classifier as selective for faces, places, and tools 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. Group average overlay maps are shown with the same color mapping as (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.

The connectivity pattern for the category-selective regions as then probed using three linear-discriminant classifiers, one for each visual category. Using leave-one-subject-out cross-validation, a classifier was trained to differentiate voxels from the category selective regions from the other voxels in the ventral stream, based on their structural connectivity with the rest of the brain. The classifier's performance was then tested on the left out subject. Using signal detection theory, d-primes were calculated for each participant, to determine how sensitive the classifiers were in locating voxels in the face, place and tool regions. All three regions could be robustly localized in adults (t(13)=26.26, p<0.001, t(13)=22.35, p<0.001, t(13)=17.17, p<0.001) (**Figure 1b**). Classification performance is quantified in **Figure 2a,** which shows the d-primes for classification of the imaging data.

In order to characterize the connectivity of the infant ventral stream, probabilistic tractography was also performed on the infant diffusion data, using the same seed and target regions as in adults (see Methods for details of two-stage adult-to-infant normalization procedure). To determine whether the category-selective regions that were present in adults were present in infants, linear discriminant classifiers were trained on the entire adult dataset in the manner described above. These classifiers were then tested on the infant data. The classifiers localized all three regions in the infants (t(10)=24.47, p<0.001**,** t(10)=16.54, p<0.001, t(10)=5.95, p<0.001) **(Figure 1c)**. 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, $N.S.$, t(23)=0.257, $N.S.$), but the tool region was detected with greater accuracy in adults than in infants $(t$(23)=3.62, p<0.01**) (Figure 2b)**. Finally, in order to examine the developmental trajectory of the networks, infant age and classification accuracy (d-primes) were correlated-- only the tool network underwent significant change over the first 9 months of postnatal life (faces: r(9)=-0.03, $N.S.$; tools: r(9)=0.75, p<0.01; places: r(9)=-0.01, $N.S.$) **(Figure 3)**.
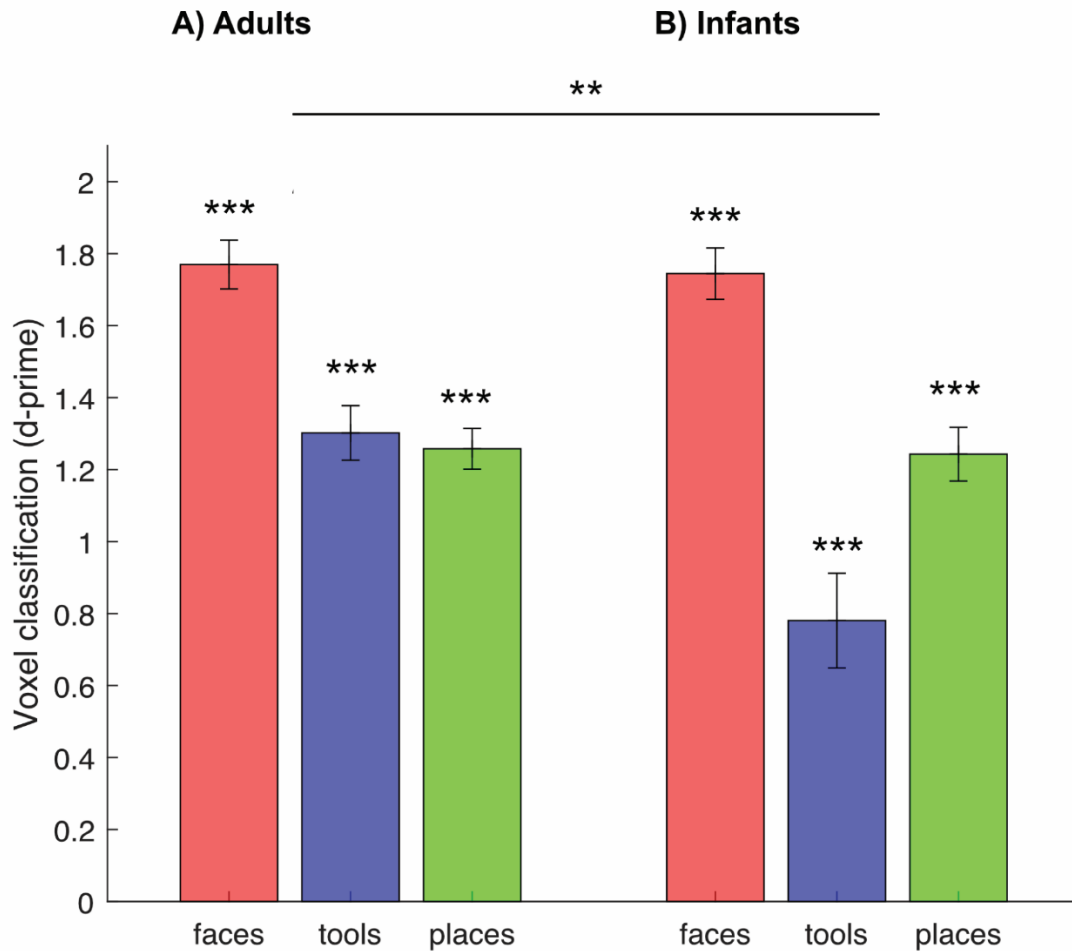
Figure 2. a) Voxel classification performance for the adults (N=14) quantified using d-prime, collapsed across hemispheres. b) Voxel classification performance for the infants (N=11) again measured using d-prime. All regions were robustly localized in infants and adults but there was a significant difference in detection accuracy between the infant and adult tool region, demonstrating the immaturity of connectivity for the tool region during infancy. The mean +/- one standard error across subjects is shown.
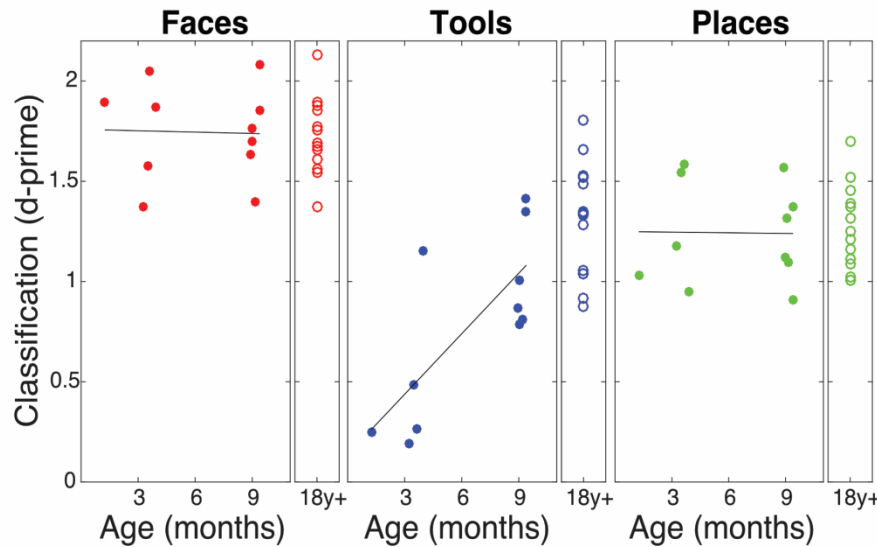
Figure 3. The relationship between the age of participants (14 adults and 11 infants) and classification accuracy (d-prime) with best-fit lines. Only tool classification had a significant relationship with age, demonstrating the maturation of the distinctive connectivity of the tool network over the first year of postnatal life

These results suggest the connectivity of the tool 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 or smaller brains, and that identification of voxels in the tool region is more sensitive to this? Two analyses were conducted to investigate this. First, in adults, the detection of the tool-selective voxels was no worse than detection of the place- selective voxels and performance was not at ceiling (Fig 2A), suggesting that detection of the tool region is not intrinsically more difficult. Second, we compared region size, which may affect performance more strongly in smaller infant brains; the place and tool 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 tool region was also not significantly different in adults ($t(13)$=-1.04, *N.S*).

We next examined which target regions of connectivity most strongly influenced each of the classifiers (see Supplementary Table 1). For the place area, this was connectivity to a network strongly associated with navigation, including the hippocampus, parahippocampal areas, and the entorhinal cortex (R. A. Epstein, 2008; Hafting, Fyhn, Molden, Moser, & Moser, 2005). The face region's strongest-weighted structural connections were to area PH, which is strongly

deactivated in the HCP functional face contrasts (Glasser et al., 2017). The tool region showed strong weighting of connections to visual regions, and accordingly, tools often have distinctive basic perceptual features (Tyler & Moss, 2001). Following this tools showed strong connections to the 4th visual area, which is associated with color processing. Tools also showed strong connections to the third visual area, which is connected to the posterior parietal regions that are associated with visuomotor transformations. Finally, the tool region's strong connectivity with the posterior orbitofrontal complex (OFC) may be driven by top-down tool classification (Bar et al., 2006).

The HCP tool region was located in the cortex between the place and face regions. Although tool selectivity has been found before in this location (Chao, Haxby, & Martin, 1999) it is also present in other areas (Grill-Spector & Weiner, 2014). From Figure 1b, it is apparent that even in the adults, there is some blurring between category boundaries, particularly between the tool and place regions. As the three classifiers were set up to each independently discriminate a single 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 the classification, but with a fitted discriminant analysis classifier that allowed for multiclass classification (Guo, Hastie, & Tibshirani, 2007). Using multiclass classification meant a single classifier aimed to predict whether a voxel was face, place, tool 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-prime, for tools vs. places, reliable in adults ($t(13)$=7.05, $p<0.001$) and infants ($t(10)$=2.40, $p<0.05$).

Discussion

All three category-selective regions could be robustly localized in adults. These results extend the fMRI results found by Deen et al. (2017) and the structural connectivity patterns found in adults by Saygin and Kanwisher (2014), demonstrating adult ventral stream category-selective regions can be delimited based on their distinctive signature of structural connectivity with the rest of the brain. Additionally, the broader networks associated with the category-selective regions were found to be completely or partially mature in infants, with the tool network experiencing a longer maturational time course, extended over the first 9 months of postnatal life. The early maturity of the distinctive networks associated with category-selective

regions suggests that infants might be going beyond perceptual categories before they are able to express themselves verbally. The associations of the three perceptual categories are 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 S1) (Epstein, 2008). However, we do acknowledge that mature connectivity profiles do not necessitate that infants have fully formed semantic, multimodal or conceptual representations. In fact, it's likely that they do not, and that networks continue to mature well into the synaptic pruning that occurs during adolescence.

It is likely that experience is playing a substantial role in the transition from perceptually based categories to conceptual maturation. For example, by one month, infants have had considerable experience with faces, spending an estimated quarter of their waking hours with faces taking up the majority of their visual field (Jayaraman, Fausey, & Smith, 2015). In contrast with this experience-driven maturation hypothesis, other researchers have proposed an innate face processing module in the brain. The most recent support for innate face processing comes from (Reid et al., 2017) who demonstrated that fetuses will preferentially orient to face-like patterns in the third trimester. However, these results have generated controversy (Scheel, Ritchie, Brown, & Jacques, n.d.) and does not align with a study in infant monkeys, which found that experience with faces is necessary to develop a typical adult like face processing system (Arcaro, Schade, Vincent, Ponce, & Livingstone, 2017).

Infants also have substantial experience viewing scenes (and due to their long supine hours, perhaps particularly ceilings). The 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 & Weiner, 2014; Lewis & Maurer, 1992). This combination of experience and mature brain circuitry might drive the mature place network connectivity, and therefore possible conceptual maturation, seen in the study.

In contrast, young infants are perhaps less likely to see tools than faces and places. Tools also are a less homogenous category, which will make category-level recognition more difficult, especially in comparison with faces, which have high similarity between exemplars. Compared with categories that are passively perceived, tool use also requires integration between sensory and motor representations, which might require more extensive experience with the environment.

By 9 months, infants are able to differentiate between textures (Bushnell & Boudreau, 1993), and hold spoons correctly during self-feeding (Bushnell & Boudreau, 1993). Using head mounted eye trackers, researchers have determined that once infants learn to reach, they'll often hold an object quite close to their face, which has been shown to be an ideal training stimulus for neural networks to recognize objects (Bambach, Crandall, Smith, & Yu, 2018). In line with this evidence and the prolonged maturation of the tool network observed here, researchers have proposed a perception-action theory, where interactions between perception and motor experience gradually accrued over development explain the maturation of tool use (Kahrs & Lockman, 2014; Lockman, 2000). This theory is supported by behavioral evidence that experience with tools drives tool use behavior (Barrett, Davis, & Needham, 2007).

The new methods developed in this work were used for the first time in infant neuroimaging. These methods have the potential as a diagnostic method for infants at high risk of developing impairments due to neurological abnormalities. In adults and children, connectivity has been shown to be predictive of many brain disorders, including mental health disorders, and can be more predictive than the standard measures doctors will use to prescribe treatment (Fox, 2018; Whitfield-Gabrieli et al., 2016). Early diagnosis could facilitate early intervention, when the brain is most plastic, which may have promising potential for infants, their families and society as a whole.

A limitation of our results is the lack of functional and diffusion data in the same participants. Future research may be able to more closely identify the network maturation of category-selective regions with a comprehensive longitudinal study, where functional localizers are acquired in both awake infant and adult participants. However, one advantage to using the regions derived from the HCP is their generalizability across a large group of 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 in infants and adults. Nevertheless, these efforts would be worthwhile and could help answer many outstanding questions.

Methods

Data Acquisition

For both the adult and infant participants, high-quality diffusion-weighted MRI data were acquired using a 3T Siemens Prisma Magnetron Scanner at the Centre 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 sec). Using monopolar diffusion encoding gradients creates larger eddy currents, which distort the magnetic field and cause image distortion. Using the solution developed by the HCP, two scans were acquired with opposite phase-encoding polarities (left-to-right and one right-to-left). Combining these images during analysis with FSL's TOPUP calculates the susceptibility distortion, while EDDY corrects for eddy current-induced distortions and participant movement (28).

During the scan, younger infants were swaddled and wrapped in a Medvac pillow bag to help them remain still (29). Infants older than 6 months were not swaddled. All infants wore Mini Muffs adhesive sound protection (Natus, 7dB attenuation) and ear defenders (29 dB attenuation). Infants were scanned during natural sleep. Adult participants wore standard ear plugs and ear defenders and were requested to be as still as possible.

Participants

Diffusion-weighted MRI acquisitions were available from 11 sleeping infants as part of a larger infant imaging project with 51 participants. Infants were recruited either through public advertising or through clinical collaborators at the neonatal intensive care unit in London, Ontario. Diffusion MRI was acquired in 14 infants but three were subsequently excluded because of apparent brain injury. This left six healthy controls and four low-risk infants born preterm. One infant was scanned twice, but as the scans were two months apart, they were treated as separate participants in the analysis, making for a total of 11 infant datasets. Clinical information for the premature infants was obtained from medical records and a radiologist reviewed each scan for suspected brain injury. Infants were between 1 and 9 months old (corrected-age for infants born preterm, M=6.4 months, SD=3.2 months).

Diffusion-weighted MRI was also acquired from 16 adults at Western University. Participants were between 18 and 40 years old (M=22.75, SD=4.89). Author LC participated in

the study and her data is included in the analysis. One participant was excluded because of an incidental finding, while another was excluded due to technical difficulties.

Approval for the study was provided by the Western University's Health Sciences Research Ethics Board. All parents provided informed consent before infants were scanned. All adult participants also provided informed consent.

Preprocessing

The data was analysed with a pipeline built from the automatic analysis (*aa*) software(30), FSL(31), and custom Matlab (R2016a). *aa* divides the description of the analysis into a user script that describes what data should be analysed, the study specific settings, and a task-list. This user script then calls then aa engine, which runs the processing pipeline, ensuring that only the stages not already completed are run, and that when possible modules are executed in parallel. The task list describes which processing modules should be used to analyse the data. For this analysis, the modules identified the DICOM files and organized them based on header information (*aamod_autoidentifyseries_timtrio, aamod_get_dicom_diffusion*) and converted them to NIFTI format (*aamod_convert_diffusion_phaseencode_direction*). Then, *aamod_diffusion_extractnodif* identified the 10 volumes where b=0 in the diffusion data. The following six stages called components of the FSL diffusion processing pipeline. To combine the negative and positive phase encoding diffusion data into a single image and reduce distortion, *aamod_diffusion_topup* (TOPUP). *aamod_bet_diffusion* then removed non-brain tissue in the b=0 image (BET). In order to correct for any residual distortions due to eddy currents or head motion, *aamod_diffusion_eddy* (EDDY). *aamod_diffusion _dtifit* was then run to model diffusion tensors at the voxel level (DTIFIT).

In order to obtain mappings between individual brains to standard (MNI) space for the infants and adults, the normalization procedure for FSL's tract-based spatial-statistics (TBSS)(32) was run. This normalizes the fractional anisotropy (FA) image to a mean FA tract skeleton, using non-linear registration. Normalizing the FA image resulted in a good registration for both the infant and adult data. Conventional normalization using a structural (T1 or T2) image to a template was not possible for the infants, as a number of the structural images were of poor quality due to participant motion. To ensure that normalizing with FA was not introducing an artifact into our results, additional analyses (not shown) in the adults where good structural

images were also available, confirmed that very similar results were obtained if normalization was performed using the structural rather than diffusion images.

Human Connectome Project

In order to identify seed and target regions for the diffusion analysis, the parcellation from the HCP was used. The HCP parcellation segments the brain into 180 distinct regions in each hemisphere, based on structural, functional and diffusion data. To identify the seed regions, the HCP definition of the ventral stream visual cortex (region 4, supplementary neuroanatomical results) was used to identify the 14 regions that make up the ventral visual stream (10). The individual voxels that were part of the 14 regions in the ventral visual stream were used as seeds and were excluded from tractography targets. The other 346 regions from the parcellation served as the target regions in the analysis. These seed and target regions were projected from the cortical surface into volumetric MNI space. The normalization parameters from TBSS were then used to project these regions from MNI space into each subject's individual diffusion data space for tractography.

To select the regions in the ventral visual stream which were most responsive to faces, places, and tools, the functional MRI localizers from the HCP project were used. The category-average contrasts were used to select regions. These regions were the fusiform complex, the ventromedial visual area 2, and ventromedial visual area 3, respectively. Regions can be seen in Figure 1A. These regions were used as the category-selective regions in the subsequent classification analysis.

Tractography and Classification

Using the data from aa's *aamod_diffusion_bedpostx* module (BEDPOSTX), probabilistic tractography was performed in the infants and adults using FSL's PROBTRACKX, using 5,000 streamlines per seed voxel in the individual subject's ventral visual stream. 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.

Three linear discriminant classifiers were then trained to identify the fusiform complex, the ventromedial visual area 2, and ventromedial visual area 3 in adults, based on connectivity with the 346 target regions. For the adults, leave-one-subject out cross validation was used to test whether selectivity could be predicted from connectivity. For the infants, three classifiers trained

in a similar way on the full adult dataset were then tested on the infant diffusion data. D-primes were calculated to evaluate the accuracy of the classifiers in both infants and adults. To test for a relationship between age and connectivity in the infants, Pearson correlations were calculated between age and d-prime scores for each category.

For the multiclass classification, a fitted discriminant analysis classifier (www.mathworks.com/help/stats/fitcdiscr.html) was used to identify the fusiform complex, the 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 d-primes 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. D-primes were used to calculate classification accuracy.

Statistics

For the single class analysis, the following statistics were conducted. Six one-sample, two-tailed t-tests were conducted (one for each category, in both the infants (N=11) and adults (N=14)) to test if the d-primes were reliably different than zero. A repeated-measures ANOVA examined main effects of and interactions between age group (infants (N=11) vs adults (N=14)), hemisphere, and category. Three two-tailed, repeated-measures t-tests were conducted to compare adult d-primes to infant d-primes; one test was conducted per category. Three Pearson's correlation coefficients were calculated to test if there was a relationship between age and d-prime in the infants (N=11); one was conducted per category. One paired-samples, two-tailed t-test was conducted to compare the hit rate between the place and tool region in the adults (N=14). For the multi-class analysis, tool d-primes were compared to place d-primes in a paired-samples, two-tailed t-test. This was done for both the infants (N=11) and adults (N=14).

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; a Journal Devoted to the Study of the Nervous System and Behavior*, *49*(9), 2334–2344.

Arcaro, M. J., Schade, P. F., Vincent, J. L., Ponce, C. R., & Livingstone, M. S. (2017). Seeing faces is necessary for face-domain formation. *Nature Neuroscience*, *20*(10), 1404–1412.

Bambach, S., Crandall, D. J., Smith, L. B., & Yu, C. (2018). Toddler-Inspired Visual Object Learning. *32nd Conference on Neural Information Processing Systems*. Retrieved from http://vision.soic.indiana.edu/papers/diversity2018nips.pdf

Bar, M., Kassam, K. S., Ghuman, A. S., Boshyan, J., Schmid, A. M., Schmidt, A. M., … Halgren, E. (2006). Top-down facilitation of visual recognition. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(2), 449–454.

Barrett, T. M., Davis, E. F., & Needham, A. (2007). Learning about tools in infancy. *Developmental Psychology*, *43*(2), 352–368.

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. *Cerebral Cortex* , *19*(12), 2767–2796.

Bushnell, E. W., & Boudreau, J. P. (1993). Motor development and the mind: the potential role of motor abilities as a determinant of aspects of perceptual development. *Child Development*, *64*(4), 1005–1021.

Chao, L. L., Haxby, J. V., & Martin, A. (1999). Attribute-based neural substrates in temporal cortex for perceiving and knowing about objects. *Nature Neuroscience*, *2*(10), 913–919.

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. *Cognitive Neuropsychology*, *21*(5), 513–527.

Deen, B., Richardson, H., Dilks, D. D., Takahashi, A., Keil, B., Wald, L. L., … Saxe, R. (2017). Organization of high-level visual cortex in human infants. *Nature Communications*, *8*, 13995.

Eimas, P. D., & Quinn, P. C. (1994). Studies on the formation of perceptually based basic-level categories in young infants. *Child Development*, *65*(3), 903–917.

Epstein, R. A. (2008). Parahippocampal and retrosplenial contributions to human spatial

navigation. *Trends in Cognitive Sciences*, *12*(10), 388–396.

Epstein, R., & Kanwisher, N. (1998). A cortical representation of the local visual environment. *Nature*, *392*(6676), 598–601.

Fang, Y., Wang, X., Zhong, S., Song, L., Han, Z., Gong, G., & Bi, Y. (2018). Semantic representation in the white matter pathway. *PLoS Biology*, *16*(4), e2003993.

Fox, M. D. (2018). Mapping Symptoms to Brain Networks with the Human Connectome. *The New England Journal of Medicine*, *379*(23), 2237–2245.

Glasser, M. F., Coalson, T. S., Robinson, E. C., Hacker, C. D., Harwell, J., Yacoub, E., … Van Essen, D. C. (2016). A multi-modal parcellation of human cerebral cortex. *Nature*, *536*(7615), 171–178.

Glasser, M. F., Coalson, T. S., Robinson, E. C., Hacker, C. D., Harwell, J., Yacoub, E., … Van Essen, D. C. (2017). Neuroanatomical Supplementary Results For A Multi-modal Parcellation of Human Cerebral Cortex. *Nature*, *536*, 1–97.

Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in Neurosciences*, *15*(1), 20–25.

Grill-Spector, K., & Weiner, K. S. (2014). The functional architecture of the ventral temporal cortex and its role in categorization. *Nature Reviews. Neuroscience*, *15*(8), 536–548.

Guo, Y., Hastie, T., & Tibshirani, R. (2007). Regularized linear discriminant analysis and its application in microarrays. *Biostatistics* , *8*(1), 86–100.

Hafting, T., Fyhn, M., Molden, S., Moser, M.-B., & Moser, E. I. (2005). Microstructure of a spatial map in the entorhinal cortex. *Nature*, *436*(7052), 801–806.

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*(7600), 453–458.

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*(5), e0123780.

Kahrs, B. A., & Lockman, J. J. (2014). Building Tool Use From Object Manipulation: A Perception-Action Perspective. *Ecological Psychology: A Publication of the International Society for Ecological Psychology*, *26*(1-2), 88–97.

Kanwisher, N., Woods, R. P., Iacoboni, M., & Mazziotta, J. C. (1997). A locus in human extrastriate cortex for visual shape analysis. *Journal of Cognitive Neuroscience*, *9*(1), 133–

142.

Kanwisher, N., & Yovel, G. (2006). The fusiform face area: a cortical region specialized for the perception of faces. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *361*(1476), 2109–2128.

Lewis, T. L., & Maurer, D. (1992). The development of the temporal and nasal visual fields during infancy. *Vision Research*, *32*(5), 903–911.

Lockman, J. J. (2000). A Perception–Action Perspective on Tool Use Development. *Child Development*, *71*(1), 137–144.

Malach, R., Reppas, J. B., Benson, R. R., Kwong, K. K., Jiang, H., Kennedy, W. A., … Tootell, R. B. (1995). Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *92*(18), 8135–8139.

Meyer, K., Kaplan, J., Essex, R., Webber, C., Damasio, H., & Damasio, A. (2010). Predicting visual stimuli on the basis of activity in auditory cortices. *Nature Neuroscience, 13,* 667-668.

Osher, D. E., Saxe, R. R., Koldewyn, K., Gabrieli, J. D. E., Kanwisher, N., & Saygin, Z. M. (2016). Structural Connectivity Fingerprints Predict Cortical Selectivity for Multiple Visual Categories across Cortex. *Cerebral Cortex* , *26*(4), 1668–1683.

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. *Nature Reviews. Neuroscience*, *8*(12), 976–987.

Quinn, P. C., Eimas, P. D., & Rosenkrantz, S. L. (1993). Evidence for representations of perceptually similar natural categories by 3-month-old and 4-month-old infants. *Perception*, *22*(4), 463–475.

Reid, V. M., Dunn, K., Young, R. J., Amu, J., Donovan, T., & Reissland, N. (2017). The Human Fetus Preferentially Engages with Face-like Visual Stimuli. *Current Biology: CB*, *27*(12), 1825–1828.e3.

Saygin, Z., & Kanwisher, N. (2014). Structural and functional connectivity fingerprints for face, body, scene, and object perception. *Journal of Vision*, *14*(10), 603–603.

Scheel, A. M., Ritchie, S. J., Brown, N. J. L., & Jacques, S. L. (n.d.). Through a womb, darkly: Methodological problems in a recent study of fetal visual perception.

Sloutsky, V. M., & Fisher, A. V. (2004). Induction and categorization in young children: a similarity-based model. *Journal of Experimental Psychology. General*, *133*(2), 166–188.

Tyler, L. K., & Moss, H. E. (2001). Towards a distributed account of conceptual knowledge. *Trends in Cognitive Sciences*, *5*(6), 244–252.

Whitfield-Gabrieli, S., Ghosh, S. S., Nieto-Castanon, A., Saygin, Z., Doehrmann, O., Chai, X. J., … Gabrieli, J. D. E. (2016). Brain connectomics predict response to treatment in social anxiety disorder. *Molecular Psychiatry*, *21*(5), 680–685.

Wilcox, T. (1999). Object individuation: infants' use of shape, size, pattern, and color. *Cognition*, *72*(2), 125–166.