# Awake fMRI reveals a specialized region in dog temporal cortex for face processing

Daniel Dilks, Peter Cook, Samuel Weiller, Helen Berns, Mark Spivak, Gregory Berns

Recent behavioral evidence suggests that dogs, like humans and monkeys, are capable of visual face recognition. But do dogs also exhibit specialized cortical face regions similar to humans and monkeys? Using functional magnetic resonance imaging (fMRI) in six dogs trained to remain motionless during scanning without restraint or sedation, we found a region in the canine temporal lobe that responded significantly more to movies of human faces than to movies of everyday objects. Next, using a new stimulus set to investigate face selectivity in this predefined candidate dog face area, we found that this region responded similarly to images of human faces and dog faces, yet significantly more to both human and dog faces than to images of objects. Such face selectivity was not found in dog primary visual cortex. Taken together, these findings: 1) provide the first evidence for a face-selective region in the temporal cortex of dogs, which cannot be explained by simple low-level visual feature extraction; 2) reveal that neural machinery dedicated to face processing is not unique to primates; and 3) may help explain dogs' exquisite sensitivity to human social cues.

# Awake fMRI reveals a specialized region in dog temporal cortex for face processing

Daniel D. Dilks[1], Peter Cook[1], Samuel K. Weiller[1], Helen P. Berns[1], Mark Spivak[2], Gregory S. Berns[1*]

**Affiliations:**

[1]Department of Psychology, Emory University, Atlanta, GA 30322.

[2]Comprehensive Pet Therapy, Atlanta, GA 30328.

**\*Correspondence to:**

Gregory S. Berns
Emory University
36 Eagle Row
Atlanta, GA 30322
gberns@emory.edu

**Awake fMRI reveals a specialized region in dog temporal cortex for face processing**

Daniel D. Dilks, Peter Cook, Samuel K. Weiller, Helen P. Berns, Mark Spivak, Gregory S. Berns

## Abstract

Recent behavioral evidence suggests that dogs, like humans and monkeys, are capable of visual face recognition. But do dogs also exhibit specialized cortical face regions similar to humans and monkeys? Using functional magnetic resonance imaging (fMRI) in six dogs trained to remain motionless during scanning without restraint or sedation, we found a region in the canine temporal lobe that responded significantly more to movies of human faces than to movies of everyday objects. Next, using a new stimulus set to investigate face selectivity in this predefined candidate dog face area, we found that this region responded similarly to images of human faces and dog faces, yet significantly more to both human and dog faces than to images of objects. Such face selectivity was not found in dog primary visual cortex. Taken together, these findings: 1) provide the first evidence for a face-selective region in the temporal cortex of dogs, which cannot be explained by simple low-level visual feature extraction; 2) reveal that neural machinery dedicated to face processing is not unique to primates; and 3) may help explain dogs' exquisite sensitivity to human social cues.

## Introduction

For social animals, faces are immensely important stimuli, carrying a wealth of information, such as identity, sex, age, emotions, and communicative intentions of other individuals (Bruce and Young, 1998; Tate et al., 2006; Leopold and Rhodes, 2010). Given the importance of face recognition for social animals, it is perhaps not surprising that humans and monkeys have dedicated neural machinery for processing visual face information discrete from the neural machinery responsible for processing nonface visual information, such as for scenes, bodies, and objects (Gross et al., 1972; Desimone et al., 1984; Perrett et al., 1988; Tsao et al., 2008; Kanwisher and Dilks, 2013). But what about other social animals, especially non-primates, like dogs? Dogs are a special case because they are both highly social with each other and have an additional evolutionary history with humans through domestication. As such, dogs may have evolved mechanisms especially tuned to social cues and therefore may have specialized neural machinery for face processing (Hare and Tomasello, 2005; Kaminski et al., 2012; Miklosi and Topal, 2013).

Behavioral evidence suggests that dogs may indeed process facial information (Racca et al., 2010; Somppi et al., 2014), but the neural mechanisms underlying the dogs' behavior could be very

2

different than humans or monkeys. For example, face recognition in dogs might rely on purely associative mechanisms, associating a face with a meaningful outcome (e.g., food). If so, then one would not expect face-specific processing in visual cortical areas, but rather activation in reward areas. Alternatively, dogs may have evolved specialized neural machinery for face recognition, and thus one would expect face-selective regions in visual cortex.

To test these competing hypotheses about face-specific processing, using fMRI, we scanned six awake, unrestrained dogs (Berns et al., 2012). To obtain high-quality fMRI data, each dog i) completed 2-4 months of behavioral training to teach them to hold still during scanning, and ii) had a custom-made chinrest to help minimize head movement. During scanning, dogs were presented with movie clips of human faces, objects, scenes, and scrambled objects (dynamic stimuli) and static images of human faces, dog faces, objects, scenes, and scrambled faces (static stimuli) on a projection screen placed in the rear of the magnet (Fig. 1 and Movie S1).

## Materials and Methods

### Participants

Participants were dogs (n=8; 5 neutered males, 3 spayed females) from the Atlanta community. All were pets and/or released service dogs whose owners volunteer their time for fMRI training and experiments. For participation in previous experiments (Berns et al., 2012; Berns et al., 2013; Cook et al., 2014), dogs took part in a training program using behavior shaping, desensitization, habituation, and behavior chaining to prepare them to be comfortable with the physical confines of the MRI bore and the loud noise produced by scanning. Accordingly, all dogs had demonstrated an ability to remain still during training and scanning for periods of 60 seconds or greater.

This study was performed in strict accordance with the recommendations in the Guide for the Care and Use of Laboratory Animals of the National Institutes of Health. The study was approved by the Emory University IACUC (Protocol #DAR-2001274-120814BA), and all dogs' owners gave written consent for participation in the study.

### Training

All dogs had previously undergone training which involved the presentation of images on a computer screen (Cook et al., 2014). Thus, prior to participation in the current experiment, the dogs were accustomed to viewing images on a screen in the MRI. Prior to actual scanning, all dogs underwent practice sessions with a complete run through of all stimuli (described below), which were presented in

3

a mock scanner on a computer screen. Dogs were approved for scanning by demonstrating that they could remain motionless for the duration of at least two, 20s-blocks of black and white images of human faces, dog faces, objects, scenes, and scrambled faces, with no actual human in view. Between image blocks, dogs were either praised or rewarded with a food treat for holding still.

**Experimental Design**

In the current experiment, a blocked fMRI design was used in which the dogs viewed either movie clips (dynamic stimuli) or static images (static stimuli). Each dog completed 3 dynamic stimuli runs and 3-4 static stimuli runs, depending on their tolerance of the MRI. In the dynamic runs, dogs were presented with 3-s color movie clips of human faces, objects (toys), scenes, and scrambled objects (Fig. 1 and Supplemental movie S1). The scrambled object movies were constructed by dividing each object movie clip into a 15 by 15 box grid and spatially rearranging the location of each of the resulting movie frames. There were 7 movie clips for each category. Each run contained two sets of four consecutive stimulus blocks in palindromic order (e.g., faces, objects, scenes, scrambled objects, scrambled objects, scenes, objects, faces), to make two blocks per stimulus category per run. In the static runs, dogs were presented with black and white images of human faces, dog faces, objects, scenes, and scrambled faces (Pitcher et al., 2011). The scrambled face images were constructed using the steerable pyramid method (Simoncelli and Freeman, 1995). Each image was presented for 600 ms followed by a 400 ms black screen interstimulus interval. There were 20 images for each category. Each run contained two sets of five consecutive stimulus blocks in palindromic order (e.g., human faces, dog faces, objects, scenes, scrambled faces, scrambled faces, scenes, objects, dog faces, human faces), to make two blocks per stimulus category per run.

**MRI Scanning**

All scanning was conducted with a Siemens 3 T Trio whole-body scanner. Head movement was minimized by a custom chinrest for each dog that allowed the dog to achieve consistent stationing in the MRI coil (a standard neck coil) (Berns et al., 2013). All participants wore ear plugs during scanning. Each scan session began with a 3s, single image localizer in the sagittal plane (SPGR sequence, slice thickness=4mm, TR=9.2 ms, TE=4.15ms, flip angle=40degrees, 256x256 matrix, FOV=220mm). A T2-weighted structural image was previously acquired during one of our earlier experiments using a turbo spin-echo sequence (25-30 2 mm slices, TR=3940 ms, TE=8.9ms, flip angle=131 degrees, 26 echo trains, 128x128 matrix, FOV=192mm), which lasted ~30 s.

Functional scans used a single-shot echo-planar imaging (EPI) sequence to acquire volumes of 24 sequential 3 mm slices with a 10% gap (TE=28 ms, TR=1400 ms, flip angle=70 degrees, 64x64 matrix, 3 mm in-plane voxel size, FOV=192mm). Slices were oriented dorsally to the dog's brain (coronal to the magnet, as, in the sphinx position, the dogs' heads were positioned 90 degrees from the usual human orientation) (Fig. 1) with the phase-encoding direction right-to-left. Sequential slices were used to minimize between-plane offsets from participant movement, and the 10% slice gap minimized the crosstalk that can occur with sequential scan sequences.

Stimuli were presented using Python 2.7.9 and the Expyriment library. Each stimulus block was manually triggered by an observer at the rear of the magnet. This ensured that the dog was properly stationed at the beginning of each block. Importantly, no actual human was in view during any of the stimulus presentation blocks. The center of each stimulus was presented binocularly, and at eye level in front of the dog, such that each stimulus fell in the center of the visual field when the dog was looking forward. Each dog completed 3 dynamic stimuli runs and 3-4 static stimuli runs, depending on their tolerance of the MRI.

**Functional Data Preprocessing and Analysis**

Preprocessing was conducted using AFNI (NIH) and its associated functions, and steps were identical to those described previously (Berns et al., 2012; Berns et al., 2013). In brief, 2-pass, 6-parameter affine motion correction was used with a hand-selected reference volume for each dog. Because dogs moved between blocks (and when rewarded), aggressive censoring was carried out, relying on a combination of outlier voxels in terms of signal intensity and estimated motion. Censored files were inspected visually to be certain that bad volumes (e.g., when the dog's head was out of the scanner) were not included. The majority of censored volumes followed the consumption of food. If less than 33% of the volumes were retained, we excluded that subject (Berns et al., 2013). This resulted in the exclusion of two dogs. For the remaining six dogs, 47-76% of volumes were retained.

EPI images were then smoothed and normalized to %-signal change. Smoothing was applied using 3dmerge, with a 6 mm kernel at Full-Width Half-Maximum (FWHM). The resulting images were then input into the General Linear Model. For each subject, a General Linear Model was estimated for each voxel using 3dDeconvolve. For the dynamic stimuli runs, the task-related regressors in this model were: (1) human faces, (2) objects, (3) scenes, and (4) scrambled objects. For the static stimuli runs, the task-related regressors in this model were: (1) human faces, (2) dog faces, (3) objects, (4) scenes, and (5) scrambled faces. Because our previous work measuring the hemodynamic response function (hrf) in

5

dogs on this task revealed a peak response at 4-6 s after signal onset (Berns et al., 2012), the six task regressors were modeled as variable duration events and were convolved with a single gamma function approximating the hrf. Motion regressors generated by the motion correction were also included in the model to further control for motion effects. A constant and linear drift term was included for each run.

To identify a candidate dog face area (DFA), we used the dynamic stimuli runs (contrast: faces > objects). Next, we identified the area of peak activation in the temporal lobe ventral and caudal to the splenium (Datta et al., 2012). We then placed a spherical ROI of 5 mm radius around the peak voxel in this region for subsequent testing of the static stimuli. For V1, we identified the area of peak activation in the dorsal occipital region when comparing the mean signal of all categories in the dynamic condition to baseline. The mean signal was then extracted from these ROIs for each dog in each of the 5 conditions of static stimuli. We used a mixed-effect model (SPSS 21, IBM) to determine whether there was a significant effect of stimulus category in both the putative DFA and V1 (maximum-likelihood estimation, diagonal covariance structure for repeated effects).

**Results**

The dynamic stimuli runs served as a functional localizer for a candidate dog face area (DFA) in inferior temporal cortex (Pitcher et al., 2011). The DFA was defined in each dog as a contiguous cluster of voxels peaking in the temporal cortex, ventral and caudal to the splenium that responded more to movies of human faces than to movies of objects. The DFA was detected in the right hemisphere in all six dogs, but differed slightly in the medial-lateral direction: Four dogs exhibited a DFA more medially, while the other two dogs exhibited a DFA more laterally. Next, a region-of-interest (ROI) of 5 mm radius was centered over the peak voxel within the predetermined DFA for each dog (Fig. 2), and the activity of this ROI was compared across the stimulus categories in the static stimuli runs. Crucially, data from the static stimuli runs served as the test data and were independent from the data used to define the DFA. First, we found no significant difference in activation between human and dog faces in the DFA [$t(5)$ = 1.21, $p = 0.28$), and thus collapsed across these two categories. Second, to investigate face selectivity, we then compared the response to images of faces to images of objects, scenes, and scrambled faces in DFA, and found a significant category effect [$F(3,9.4) = 10.01$, $p = 0.003$], with a significantly greater response to images of faces compared to objects ($p_{unc} < 0.001$; $p_{Bonf} < 0.001$) and scenes ($p_{unc} = 0.022$; $p_{Bonf} = 0.067$), but not scrambled faces ($p_{unc} = 0.26$; $p_{Bonf} = 0.78$) (Fig. 3A). These findings not only reveal within-subject replicability across paradigms, but also the face selectivity of the DFA.

Given the similarity in low-level features between faces and scrambled faces, it is not surprising

6

that the DFA might respond to scrambled faces, albeit less reliably than to the images of faces themselves. But might the face selectivity in putative DFA be explained entirely by retinotopic information simply inherited from early visual cortex? To address this possibility, we defined the primary visual cortex (V1) (contrast: average of all stimulus categories versus baseline) using the dynamic stimuli runs for each dog. For all subjects, we found a region dorsally in the caudal portion of the marginal and endomarginal gyri, consistent with the known location of dog V1 (Beitz and Fletcher, 1993). Next, an ROI of 5 mm radius was centered over the peak voxel within the predetermined V1 for each dog, and the activity of this ROI was compared across the stimulus categories in the static stimuli runs. The response profile of this V1 ROI was compared to the response profile of the DFA. A 2 (ROI: DFA, V1) x 4 (Condition: faces, objects, scenes, scrambled) mixed-effect model revealed a significant interaction [$F(3,20.2) = 4.17$, $p = 0.02$), indicating that the response profile of the DFA was not like that of V1, and thus not strictly a result of low-level feature extraction (Fig. 3B).

**Discussion**

Taken together, the above results provide the first evidence for a region in temporal cortex of dogs involved in the processing of faces. Indeed, while there is ample behavioral evidence that dogs respond to faces, our results demonstrate an evolutionary continuity in the neural substrates of a key aspect of social behavior: a face-selective region in dog cortex located in an area similar to that of humans and monkeys. The commonality of location is consistent with the commonality and importance of face processing in social species and is found in visual cortex, suggesting that dogs' ability to process faces is not simply the result of learned associations. Our finding that dogs, like humans and monkeys, exhibit specialized cortical face regions is also consistent with two other studies demonstrating that neural machinery dedicated to face processing may not be unique to primates, having been observed in sheep (Kendrick and Baldwin, 1987) and crows (Marzluff et al., 2012).

In addition to behavioral evidence suggesting dedicated neural machinery for face processing in dogs, one previous study suggested a neural signature for such processing (Tornqvist et al., 2013). Using visual event-related potentials (ERPs), this study reported differences in two ERP components between the responses to human and dog faces. This study gave the first hint of a neural substrate for face processing but also raised several questions, namely the degree to which visual recognition of any object (given only face stimuli were tested), or low-level feature extraction (given the low-level visual differences between the images of human and dog faces) might explain the results. Moreover, the limited spatial resolution of EEG precluded the precise localization of putative face-selective machinery,

7

which is relatively small and restricted to specific regions of occipital and temporal cortex in primates. Our fMRI results build on these ERP findings and offer strong evidence for a face-selective region in dog temporal cortex, responding significantly more to images of faces than to images of other stimulus categories (i.e., objects, scenes, and scrambled faces). Furthermore, the face selectivity of the DFA was not found in dog primary visual cortex, ruling out simple low-level feature extraction as explanations for the face-selective response in DFA.

The principal limitation of our study stems from the small effect size. The average differential BOLD response was well less than 1%, which is consistent with human fMRI studies. Comparable animal fMRI studies, however, overcome the signal limitation by immobilizing the subject and scanning for much longer periods of time to decrease the effects of noise, but this approach often uses a small number of subjects – typically two. In contrast, our approach is to use awake, trained dogs who cooperatively enter the scanner and hold still for periods up to several minutes without restraint. And while the dogs do extraordinarily well, the data quality cannot approach that obtained from a sedated, immobilized monkey. The trade-off is noisier data. We compensate by using more subjects than a typical monkey study, here reporting the data from six dogs. Although we have studied larger cohorts of dogs in previous studies, watching images on a flat screen is not a natural behavior for dogs, and only a subset of the MRI-trained dogs would do so, even after months of training. Even so, the data we report here show a high degree of within-subject replicability, with some inter-subject variation in the location of DFA, some of which may be due to noise and some due to the existence of multiple face-sensitive patches.

In summary, the existence of a face-selective region in temporal dog cortex opens up a whole range of new questions to be answered about their social intelligence: What is the relative impact of early socialization on dog versus human face processing? Do face regions in dogs process emotional content? Given canids' reliance on body posture for communication, are there corresponding body-selective regions? We do not know whether face-selective cortex in dogs is a result of the domestication process and dogs' resultant reliance on humans, or whether such face regions predate domestication and exist widely in other social carnivores. But the relatively small size of the dog brain, and the dedication of face processing to specific regions, highlights the importance of face processing to this species, and may explain dogs' exemplary skill at interspecies communication.

8

**References**

Beitz AJ, Fletcher TF (1993) The brain. In: Miller's Anatomy of the Dog, Third Edition (Evans HE, ed). Philadelphia: W.B. Saunders Company.

Berns GS, Brooks AM, Spivak M (2012) Functional MRI in awake unrestrained dogs. PLoS ONE 7:e38027.

Berns GS, Brooks A, Spivak M (2013) Replicability and heterogeneity of awake unrestrained canine fMRI responses. PLoS ONE 8:e81698.

Bruce V, Young A (1998) In the Eye of the Beholder: The Science of Face Perception. New York: Oxford University Press.

Cook PF, Spivak M, Berns GS (2014) One pair of hands is not like another: caudate BOLD response in dogs depends on signal source and canine temperament. PeerJ 2:e596.

Datta R, Lee J, Duda J, Avants BB, Vite CH, Tseng B, Gee JC, Aguirre GD, Aguirre GK (2012) A digital atlas of the dog brain. PLoS ONE 7:e52140.

Desimone R, Albright TD, Gross CG, Bruce C (1984) Stimulus-selective properties of inferior temporal neurons in the macaque. J Neurosci 4:2051-2062.

Gross CG, Rocha-Miranda CE, Bender DB (1972) Visual properties of neurons in inferotemporal cortex of the macaque. J Neurophys 35:96-111.

Hare B, Tomasello M (2005) Human-like social skills in dogs? Trends Cogn Sci 9:439-444.

Kaminski J, Schulz L, Tomasello M (2012) How dogs know when communication is intended for them. Developmental Science 15:222-232.

Kanwisher N, Dilks DD (2013) The functional organization of the ventral visual pathway in humans. In: The New Visual Neurosciences (Chalupa LM, Werner JS, eds). Cambride, MA: MIT Press.

Kendrick KM, Baldwin BA (1987) Cells in temporal cortex of conscious sheep can respond preferentially to the sight of faces. Science 236:448-450.

Leopold DA, Rhodes G (2010) A comparative view of face perception. Journal of Comparative Psychology 124:233-251.

Marzluff JM, Miyaoka R, Minoshima S, Cross DJ (2012) Brain imaging reveals neuronal circuitry underlying the crow's perception of human faces. Proc Natl Acad Sci U S A 109:15912-15917.

Miklosi A, Topal J (2013) What does it take to become 'best friends'? Evolutionary changes in canine social competence. Trends Cogn Sci 17:287-294.

Perrett DI, Mistlin AJ, Chitty AJ, Smith PAJ, Potter DD, Broennimann R, Harries M (1988) Specialized face processing and hemispheric asymmetry in man and monkey: Evidence from single unit and reaction time studies. Behav Brain Res 29:245-258.

Pitcher D, Dilks DD, Saxe RR, Triantafyllou C, Kanwisher N (2011) Differential selectivity for dynamic versus static information in face-selective cortical regions. NeuroImage 56:2356-2363.

Racca A, Amadei E, Ligout S, Guo K, Meints K, Mills D (2010) Discrimination of human and dog faces and inversion responses in domestic dogs (Canis familiaris). Animal Cognition 13:525-533.

Simoncelli EP, Freeman WT (1995) The steerable pyramid: a flexible architecture for multi-scale derivative computation. In: International Conference on Image Processing, pp 444-447: IEEE Comput. Soc. Press.

Somppi S, Tornqvist H, Hanninen L, Krause CM, Vainio O (2014) How dogs scan familiar and inverted faces: an eye movement study. Animal Cognition 17:793-803.

Tate AJ, Fischer H, Leigh AE, Kendrick KM (2006) Behavioural and neurophysiological evidence for face identity and face emotion processing in animals. Phil Trans R Soc (Lond) B 361:2155-2172.

Tornqvist H, Kujala MV, Somppi S, Hanninen L, Pastell M, Krause CM, Kurjala J, Vainio O (2013) Visual event-related potentials of dogs: a non-invasive electroencephalography study. Animal Cognition 16:973-982.

Tsao DY, Moeller S, Freiwald WA (2008) Comparing face patch systems in macaques and humans. Proc
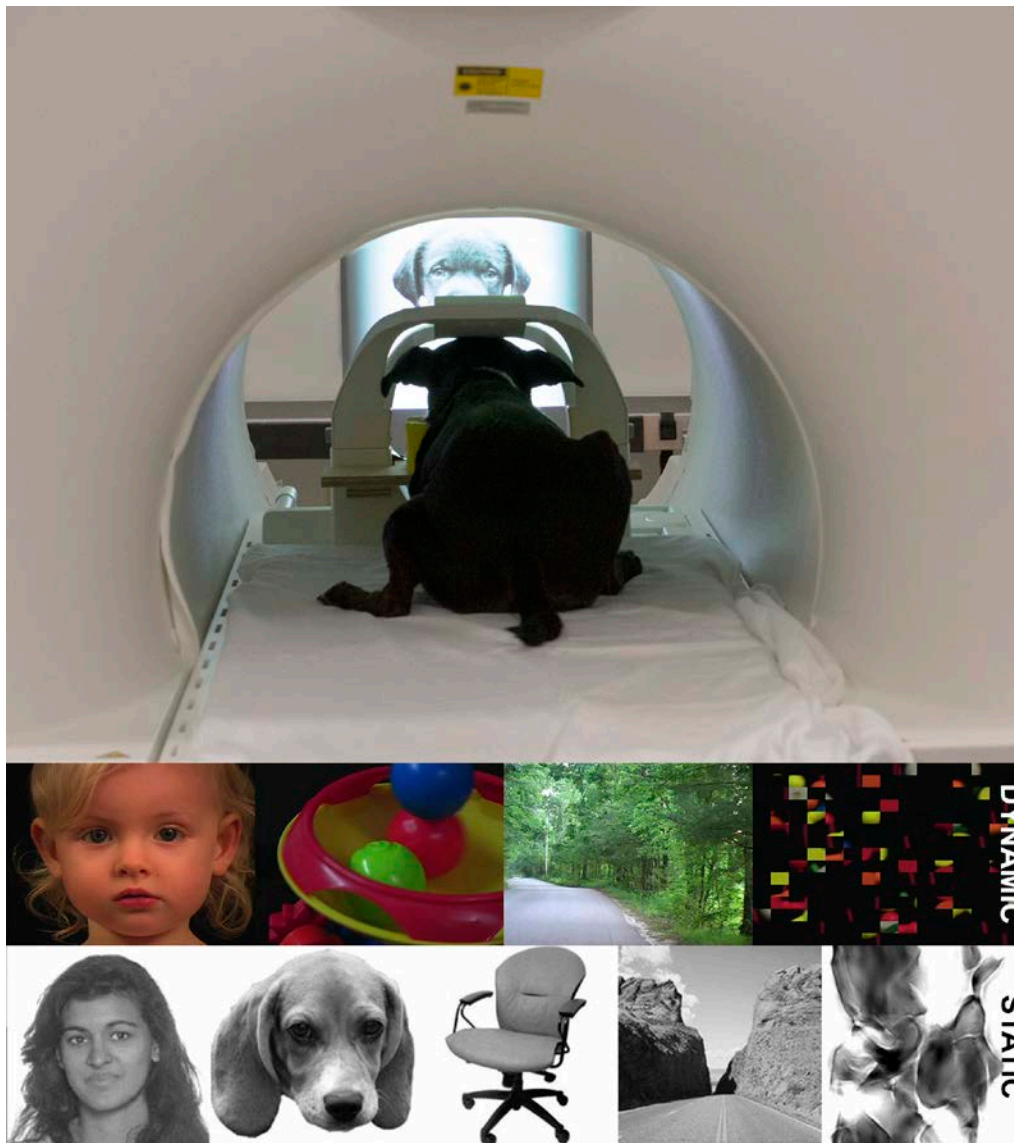    Natl Acad Sci U S A 105:19514-19519.

**Fig. 1. Experimental setup in MRI.** Dogs were trained to station within an individually customized chin rest placed inside a stock human neck coil. The upper surface coil was located just superior to the dog's head. Images were rear projected onto a translucent screen placed at the end of the magnet bore. In the dynamic stimuli runs, color movies clips (3-s each) were shown in 21 s blocks of human faces, objects (toys), scenes, and scrambled objects. In the static stimuli runs, black and white images (600 ms on, 400 ms off) were shown in 20 s blocks of human faces, dog faces, everyday objects, scenes, and scrambled faces. The dynamic stimuli runs were used to localize a candidate face region in the temporal cortex of dogs, and then the static stimuli runs were used to independently test the face selectivity of this region.
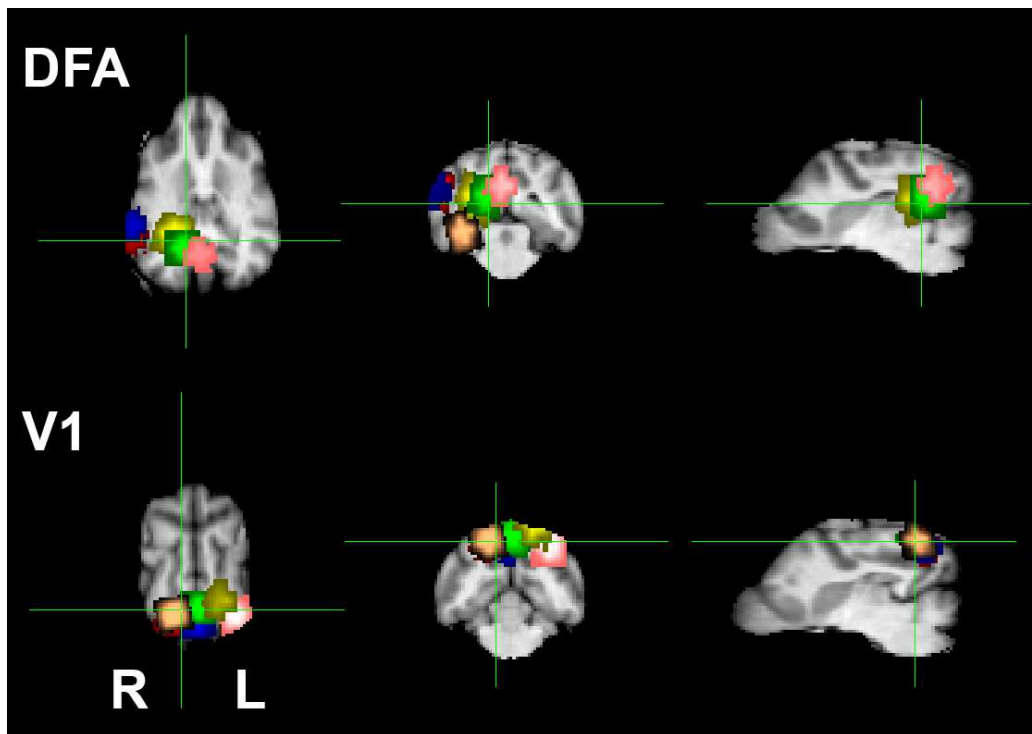
11

**Fig. 2. ROI locations for the dog face area (DFA) and primary visual cortex (V1).** The DFA (*top*) was identified by the contrast of faces versus objects during the dynamic stimuli runs. Each color represents the ROI of one dog. For visualization and comparison of location, the ROIs have been spatially normalized and overlaid on a high resolution dog brain atlas (Datta et al., 2012). The location of the DFA was localized to the medial bank of the ventrocaudal temporal lobe in 4 of the 6 dogs, with the other 2 localized more laterally. V1 was identified by the average of all dynamic run conditions (face, objects, scenes, scrambled) relative to baseline. In each dog, a dorsal area of activation in the caudal portions of the marginal/endomarginal gyri was identified and corresponded to the known location of primary visual cortex.
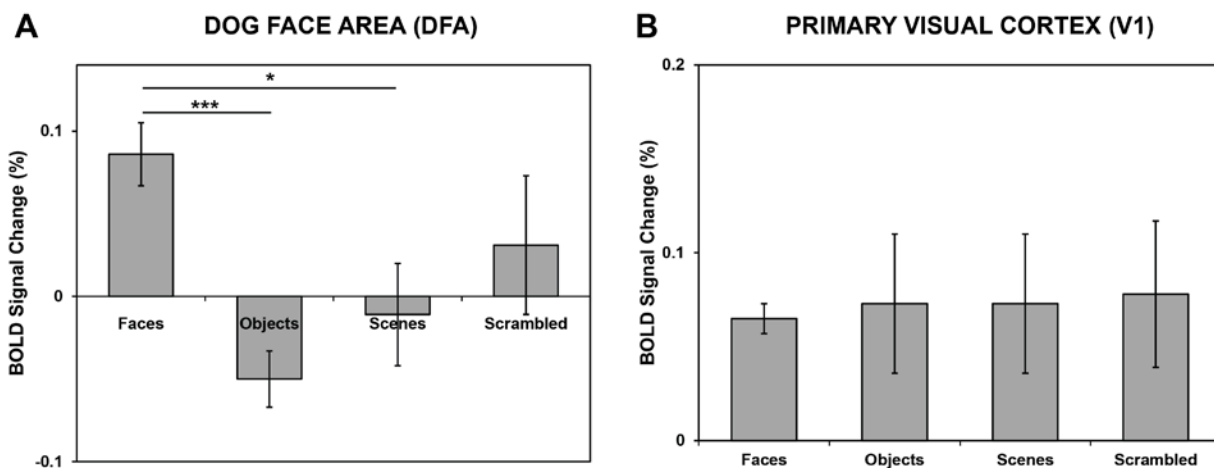
**Fig. 3. Average percent signal change in DFA and V1**. Error bars indicate the standard error of the mean (n = 6 subjects). (A) We found a significant category effect [$F_{(3,9.4)}$ = 10.01, p = 0.003], with a significantly greater response to images of faces compared to objects (***p < 0.001) and scenes (*p = 0.022). (B) V1 had a similar level of response to all stimulus categories [$F_{(3,12.0)}$ = 0.059, p = 0.98], and crucially was significantly different from DFA [$F_{(3,20.2)}$ = 4.17, p = 0.02].