Brief Communications

Nature versus Nurture in Ventral Visual Cortex: A Functional Magnetic Resonance Imaging Study of Twins

Thad A. Polk,[1] **Joonkoo Park,**[1] **Mason R. Smith,**[1] **and Denise C. Park**[2]

[1]Department of Psychology, University of Michigan, Ann Arbor, Michigan 48109, and [2]The Beckman Institute, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801

Using functional magnetic resonance imaging, we estimated neural activity in twins to study genetic influences on the cortical response to categories of visual stimuli (faces, places, and pseudowords) that are known to elicit distinct patterns of activity in ventral visual cortex. The neural activity patterns in monozygotic twins were significantly more similar than in dizygotic twins for the face and place stimuli, but there was no effect of zygosity for pseudowords (or chairs, a control category). These results demonstrate that genetics play a significant role in determining the cortical response to faces and places, but play a significantly smaller role (if any) in the response to orthographic stimuli.

Introduction

It is now well established that different categories of visual stimulus elicit distinct patterns of neural activity in ventral visual cortex. Faces activate parts of the midfusiform gyrus, often bilaterally (Kanwisher et al., 1997), whereas buildings, outdoor scenes, and other places in the environment activate parts of the parahippocampal gyrus bilaterally (Aguirre et al., 1998; Epstein and Kanwisher, 1998). Words, pronounceable pseudowords, and letters activate parts of the left fusiform gyrus and collateral sulcus (Puce et al., 1996; Cohen et al., 2000; Polk and Farah, 2002; Polk et al., 2002; Baker et al., 2007). Although there is debate about whether the neural representation of these visual categories is localized to relatively small cortical regions or is distributed across a much larger area (Haxby et al., 2001), these stimulus categories definitely elicit distinctive and reproducible patterns of neural activity in the ventral visual cortex.

In this study, we assessed whether genetics play a significant role in determining the neural activity patterns associated with these visual categories. Is the cortical response to faces, places, and orthographic stimuli at least partly innate? In the domain of face recognition, the evidence is mixed. On the one hand, brain damage immediately after birth can produce selective and lasting deficits in face recognition (Farah et al., 2000), suggesting that some of the neural substrates of face recognition can be established before any postnatal experience. However, Le Grand et al. (2003) found that early deprivation of visual input to the right, but not left, hemisphere undermines the development of normal face processing expertise. Their interpretation was that the neural

circuitry underlying face processing is not prespecified, but requires early visual experience to develop. Consistent with this view, when car and bird experts view cars and birds, they exhibit significant activity in the same ventral area that is activated by faces (Gauthier et al., 2000), and substantial training with novel stimuli can also lead to activation of this area (Gauthier et al., 1999). These latter findings are also consistent with the hypothesis that this activity is not entirely prespecified by genetics, but is significantly influenced by experience.

The influence of genetic similarity on the cortical response to places and orthographic stimuli is also unknown. We know of no existing studies that shed light on this question for place recognition. Visual word recognition is an acquired skill that does not develop without substantial training. It is a recent development on an evolutionary scale, and it is not shared with any other species. It therefore seems unlikely that the cortical response to visually presented orthographic stimuli would be genetically determined. Reading may make unique demands on more general perceptual processes, however (e.g., recognizing specific combinations of shapes quickly and in parallel) (Farah and Wallace, 1991), and if the cortical response to reading words is a reflection of those general processes then genetics could certainly play a significant role in determining that response. Consistent with this view, recent evidence suggests that genetics do play a role in at least some cases of developmental dyslexia (Fisher and DeFries, 2002).

Twin studies provide a direct approach to assessing the role of genetics in neural organization. Previous studies have demonstrated that some aspects of structural brain organization are more similar in monozygotic (MZ) twins than in dizygotic (DZ) twins (Thompson et al., 2001). Our focus is on functional brain organization. If the patterns of neural activity associated with a given visual category are significantly more similar in MZ twins than in DZ twins, then that is strong evidence that genetics played a significant role in determining that activity pattern. Conversely,

if the patterns in MZ twins are no more similar than in DZ twins, it provides evidence that genetics did not play a significant role.

Materials and Methods

Participants. We tested 13 pairs of right-handed MZ twins (nine female pairs, four male pairs, ages 18–29 with a mean age of 21.3) and 11 pairs of DZ twins (seven female pairs, four male pairs, ages 18–23, mean age 19.9). Zygosity was determined by comparing seven to eight highly variable DNA markers (D5S818, D13S317, D7S820, D16S539, vWA, TH01, TPOX, CSF1PO) from the buccal cells of twins collected by swabbing the cheek of each participant. DNA was amplified using the polymerase chain reaction technique. Twins in whom all the markers matched were classified as monozygotic and twins in whom some markers mismatched were classified as dizygotic.

Experimental design. During functional MRI scanning, participants performed a one-back matching task (press a button with the right, middle finger if the current stimulus matches the last stimulus, press a button with the right, index finger if they do not match) on gray-scale pictures from five categories: faces, places (pictures of houses), pseudowords (pronounceable nonword letter strings), chairs, and phase-scrambled control images (see Fig. 1). Pseudowords, instead of words, were used to ensure that there were no effects of word semantics on activation patterns (and pseudowords and words are known to elicit similar activation patterns in ventral visual cortex) (Polk and Farah, 2002). In keeping with many previous studies, houses were used to activate the so-called parahippocampal place area (Epstein and Kanwisher, 1998). The control images were versions of the experimental images in which the phase information was scrambled, but spatial frequency was preserved (i.e., the power spectra were identical). There were three runs with 15 20 s blocks per run (three blocks of each of the five categories of stimuli in pseudorandom order) with a 20 s rest period preceding each run. Each block consisted of 10 items from the same category presented for 1500 ms each, followed by a 500 ms intertrial interval.

fMRI data acquisition. Structural images were T1-weighted images collected in axial slices parallel to the anterior commissure–posterior commissure line with a resolution of 0.9375 × 0.9375 × 5.0 mm. Neural activity was estimated based on the blood oxygen level-dependent (BOLD) signal measured by a GE 3T scanner using a spiral acquisition sequence (2000 ms repetition time, 30 5 mm thick axial slices with an in-plane resolution of 3.75 × 3.75 mm, 24 cm field of view, 30 ms echo time, 90 degree flip angle).

Preprocessing. The functional images (T2*-weighted echo planar images) for each participant underwent reconstruction, slice timing correction, and realignment as part of preprocessing. These images were then spatially smoothed with a Gaussian kernel of 8 × 8 × 8 mm and normalized into a standard MNI (the Montreal Neurological Institute) space and resampled into a resolution of 3 × 3 × 3 mm per voxel. The anatomical images (T1-weighted images) for each participant were also normalized and resampled into a resolution of 3 × 3 × 3 mm. Normalized images were resampled for computational efficiency for the calculation of effective sample size (see below), and it was verified that the similarity results were the same whether images were resampled or not.

ROI. An anatomical mask was first constructed that included the parahippocampal, fusiform, inferior temporal, and inferior occipital gyri bilaterally using the PickAtlas software toolbox (Maldjian et al., 2003). Then, the functional data from the first run (Run1) were analyzed on a voxel-by-voxel basis using a general linear model corrected for temporal autocorrelation with regressors corresponding to each of the four experimental conditions (using the VoxBo software package, http://www.voxbo.org). A region of interest was then defined as the union of voxels within this mask that were significantly active for any of the four major contrasts (faces, places, pseudowords, or chairs vs phase-scrambled control) in a random-effects group analysis that included all participants ($p < 10^{-6}$, uncorrected). The region of interest consisted of 1237 voxels.

Subregions of the region of interest (ROI) were also separately defined to test whether anatomical similarity was uniform across different functional subregions of ventral visual cortex. Voxels within the anatomical ROI that were significantly active for each of the four stimulus categories versus control in the group analysis ($p < 10^{-6}$, uncorrected) were de-

fined as the four subregions. For example, the face subregion was defined as the voxels significantly active for the contrast of faces versus phase-scrambled control, and it included 414 voxels. Likewise, the house subregion included 1139 voxels, the pseudoword subregion included 67 voxels, and the chair subregion included 764 voxels.

Functional maps of activation. The functional data from the second and third run (Run2 and Run3) were analyzed on a voxel-by-voxel basis using a general linear model corrected for temporal autocorrelation with regressors corresponding to each of the four experimental conditions. Then, functional statistical parametric maps of t-values and β-values for each of the four major contrasts were computed. Thus, the data used to generate the functional maps was independent of the data used to generate the ROI. An initial examination of the functional maps of all participants indicated that one participant from the MZ twin group did not show any activation within the anatomical mask for the face and house contrasts and another participant from the same group did not show any activation for the house contrast, even when a liberal activation threshold ($p < 0.01$, uncorrected) was adopted. These participants and their siblings were therefore excluded from further analysis.

Similarity analysis. The similarity of each twin pair for all four contrasts was assessed by computing a Pearson correlation coefficient directly on the t values (and also, β weights) across the entire region of interest (see Fig. 1). To conduct inferential tests on the functional similarities of MZ versus DZ twins, the correlation coefficient r for each pair was transformed into a z-score using Fisher's r-to-Z transformation. In addition, because the maps are spatially autocorrelated, the effective sample size is smaller than the total number of voxels in the region of interest. We therefore computed an effective sample size for a given pair of maps using the technique described by Clifford et al. (1989). A Z-statistic corresponding to the contrast between z-scores for MZ and DZ twins was then computed as follows (Cohen and Cohen, 2003):

$$ Z = \frac{\sum \lambda_i z_i}{\sqrt{\sum \lambda_i^2 \dfrac{1}{\hat{M}_i - 3}}} , $$

where λ_i is the contrast weight (e.g., 1 for MZ twins, -1 for DZ twins), z_i is the z-score corresponding to the similarity of the functional maps from twin pair i, and $\hat{M}$ is the effective sample size for twin pair i. Statistical significance for the difference between MZ and DZ pairs' similarity measures was calculated based on a one-tailed test of the resulting Z-statistic unless otherwise noted.

The similarity of the two T1-weighted anatomical maps from each twin pair was assessed using the same method as the similarity analysis of the functional maps. A Pearson correlation coefficient between the two anatomical maps was computed to estimate the similarity, and the inferential test was based on the Z-statistics. A second analysis of anatomical similarity was conducted using T2*-weighted echo-planar scout images, rather than T1 images, in case the T2*-weighted images captured other aspects of anatomical structure not captured by the T1-weighted images. The similarity of the two T2*-weighted anatomical maps from each twin pair was assessed using the same method described previously.

Within-subject analysis. Within-subject correlations were computed to estimate the reliability of the experiment and to estimate the potential maximum correlation values within twin pairs. For each participant, a functional activation map for each contrast was computed from the second run (Run2) and another functional map for each contrast was computed from the third run (Run3). A correlation coefficient was then computed as described previously.

Unrelated pairs. Each participant was randomly paired with another unrelated participant to create a hypothetical group of twenty-two unrelated pairs. The functional and anatomical similarities of these unrelated pairs were computed using the same method used to compute the similarities of the related pairs.

Results

The patterns of neural activity for four stimulus categories (faces, houses, pseudowords, and chairs) within a functionally defined
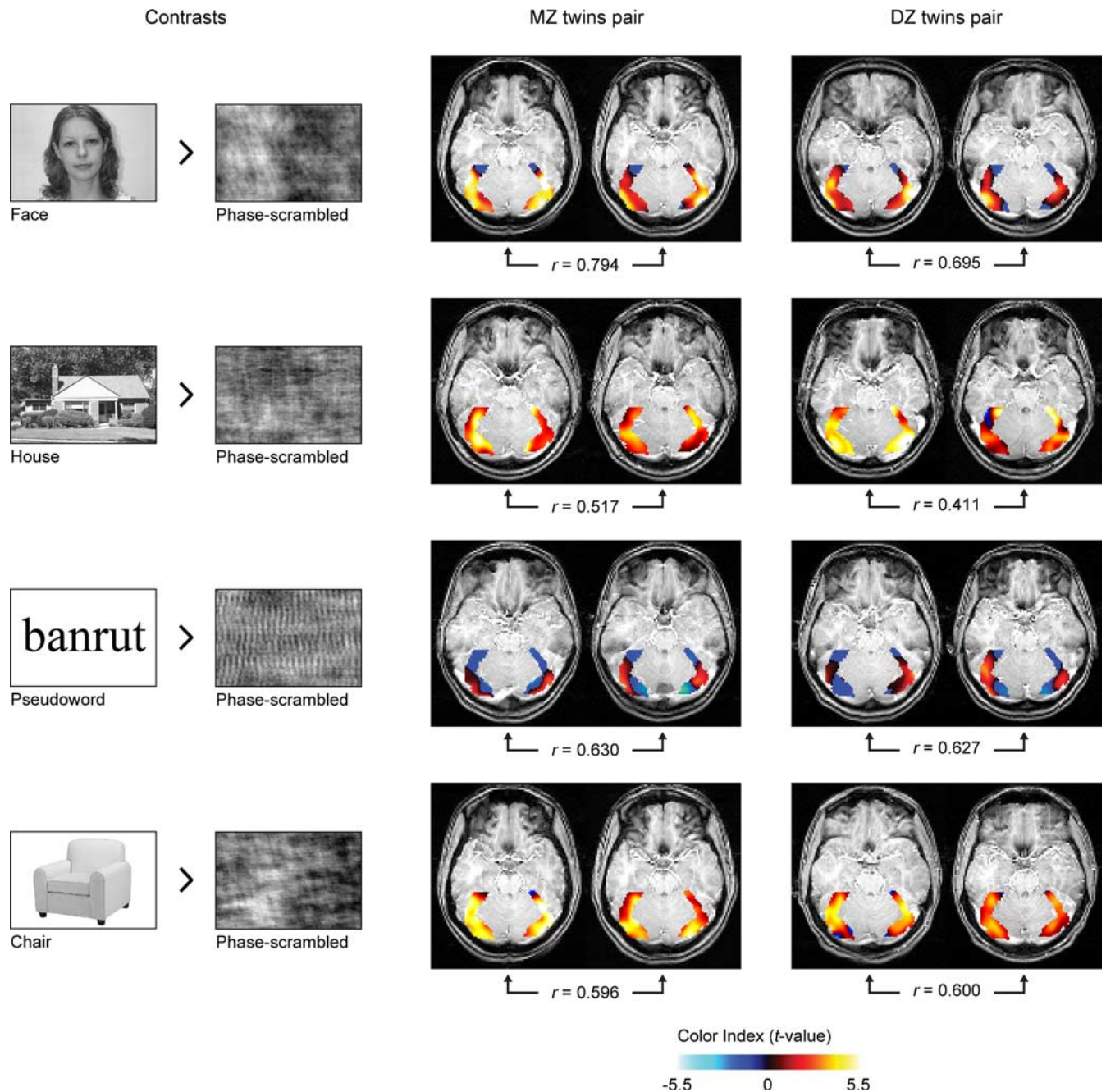
Figure 1. Patterns of estimated neural activity when viewing the four stimulus categories (axial slice, z = −18, MNI coordinate). Functional activation maps were computed for the four contrasts of interest (faces, houses, pseudowords, and chairs relative to the phase-scrambled control condition), and the similarity measures (r) between these functional maps were computed for each twin pair. Twins pairs that best represented the mean similarity measure (Fig. 2) of the group were selected for illustration.

region in ventral visual cortex were estimated for each participant, and the similarity of the two patterns from each twin pair for all four contrasts was assessed by the correlation coefficient (Fig. 1).

The neural activation patterns in response to both faces and houses were significantly more similar in MZ twins than in DZ twins (faces, $Z = 2.508, p = 0.006$; houses, $Z = 2.741, p = 0.003$), but there was no significant difference between the twin groups for pseudowords or chairs (pseudowords, $Z = 0.125, p = 0.450$; chairs, $Z = 0.677, p = 0.249$) (Fig. 2). In other words, zygosity had a significant effect on the pattern of neural activity in response to faces and houses (places), but not pseudowords or chairs. Furthermore, the effect of zygosity was significantly larger

in the face and house conditions than in the pseudoword and chair conditions ($Z = 2.194; p = 0.014$). This interaction between zygosity and stimulus category implies that the results cannot be attributed to any general differences between MZ and DZ twins (e.g., differences in overall anatomical similarity) because the effects of zygosity were selective to a subset of stimulus categories. And, as might be expected, MZ twins were indeed more anatomically similar than were DZ twins ($Z = 3.064; p = 0.001$). The effect of zygosity in the face and house, but not pseudoword or chair, conditions was replicated when β weights, rather than t-values, were used to assess the similarity of twin pairs (faces, $Z = 3.441, p < 0.001$; houses, $Z = 3.357, p < 0.001$; pseudowords, $Z = 0.511, p = 0.305$; chairs, $Z = 0.991, p = 0.161$; interaction
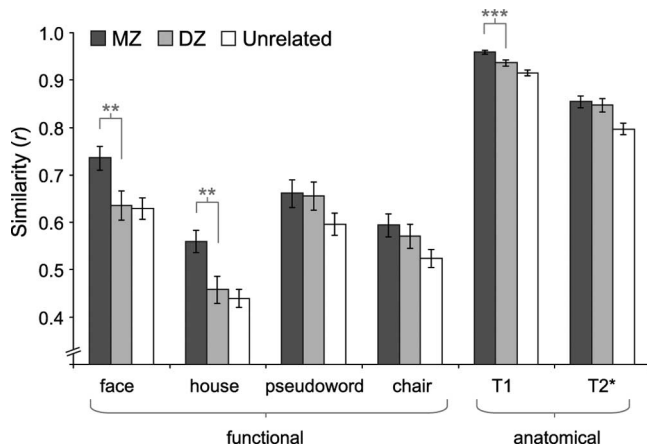
Figure 2. Mean functional similarity between MZ, DZ, and unrelated pairs for each of the four conditions (**$p < 0.01$; ***$p < 0.001$ between MZ and DZ pairs) and mean anatomical similarity. Error bars denote SEM.

Figure 3. *A*, Mean anatomical similarity (*r*) measured by T1-weighted images in functional subregions of the anatomical mask. *B*, Mean anatomical similarity (*r*) measured by T2*-weighted images in functional subregions of the anatomical mask (*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$ between MZ and DZ pairs).

between zygosity and stimulus category, $Z = 2.729$, $p = 0.003$). The results were also replicated when an anatomical, rather than functional, ROI was used (faces, $Z = 3.889$, $p < 0.001$; houses, $Z = 2.465$, $p = 0.007$; pseudowords, $Z = 1.560$, $p = 0.059$; chairs, $Z = 1.294$, $p = 0.098$; interaction between zygosity and stimulus category, $Z = 1.766$, $p = 0.039$). Furthermore, we also found a significant effect of zygosity for the direct face vs house contrast ($Z = 1.648$, $p < 0.05$ using *t*-value maps; $Z = 2.230$, $p = 0.013$ using β maps).

We also computed the same similarity measures within subjects and between unrelated pairs of subjects to estimate reasonable upper and lower bounds on the measure. The mean ($\pm$ SE) within-subject correlations for the face, house, pseudoword, and chair conditions were 0.797 (± 0.010), 0.748 (± 0.008), 0.849 (± 0.008), and 0.799 (± 0.007), respectively, which were all significantly higher than the similarity between MZ twins (faces, $Z = 2.397$, $p = 0.008$; houses, $Z = 8.626$, $p < 0.001$; pseudowords, $Z = 7.851$, $p < 0.001$; chairs, $Z = 9.710$, $p < 0.001$). The similarity of activation maps in unrelated pairs was also significantly smaller than in related pairs (MZ and DZ twins combined) for all four stimulus categories (faces, $Z = 2.001$, $p = 0.023$; houses, $Z = 2.670$, $p = 0.004$; pseudowords, $Z = 1.983$, $p = 0.024$; chairs, $Z = 2.277$, $p = 0.011$). Although the similarity of activation maps in DZ twins was higher than in unrelated pairs for all four categories, none of these differences were statistically significant (faces, $Z = 0.175$, $p = 0.431$; houses, $Z = 0.532$, $p = 0.297$; pseudowords, $Z = 1.526$, $p = 0.064$; chairs, $Z = 1.458$, $p = 0.072$), perhaps because of the relatively small sample size.

One potential concern is that zygosity could have different effects on anatomical similarity in different subregions of ventral visual cortex. For example, suppose MZ twins are more anatomically similar than DZ twins in areas involved in processing faces and houses, but not in areas involved in processing pseudowords and chairs. In that case, the observed interaction between zygosity and stimulus category on functional similarity could simply be the result of a corresponding interaction on anatomical similarity. To address this concern, we computed anatomical similarity in regions significantly activated by each of the four stimulus categories. As can be seen in Figure 3*A*, there was a significant effect of zygosity on anatomical similarity measured in T1-weighted images in all four subregions (faces, $Z = 3.459$, $p < 0.001$; houses, $Z = 2.859$ $p = 0.002$; pseudowords, $Z = 1.761$, $p = 0.039$; chairs, $Z = 2.950$, $p = 0.002$), and there was no significant
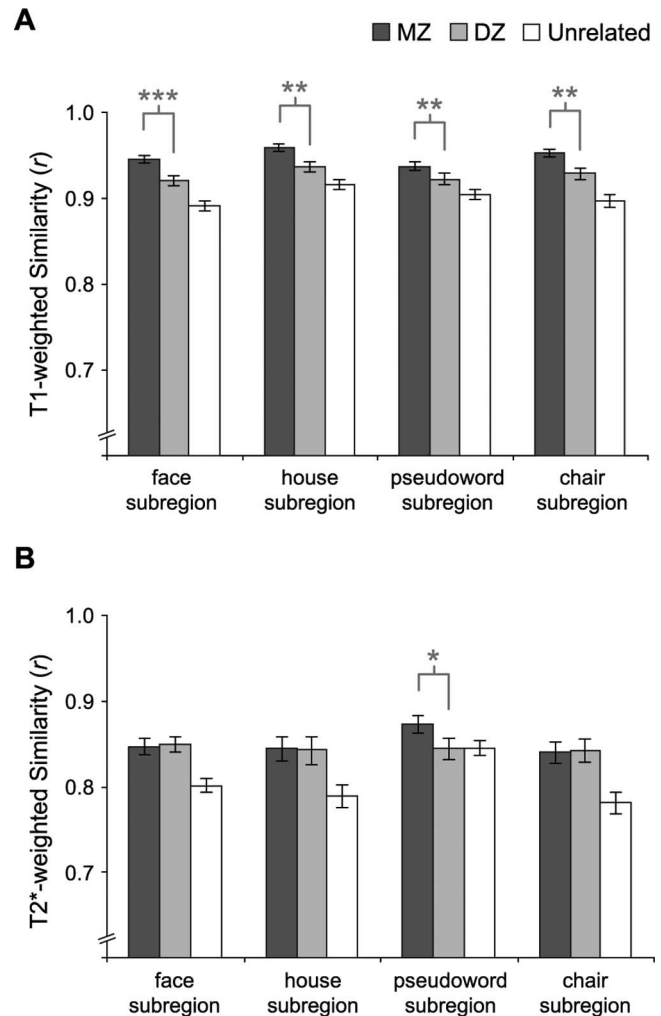
interaction between zygosity and stimulus category (the effect of zygosity in the face and house conditions vs in the pseudoword and chair conditions, $Z = 0.677$, $p = 0.249$). Anatomical similarity measured by T2*-weighted images (Fig. 3*B*) showed that the only significant effect of zygosity was in the pseudoword subregion ($Z = 1.788$; $p = 0.037$), but of course pseudowords were one of the conditions in which zygosity had no effect on functional similarity.

Discussion
We found that ventral visual activity patterns associated with face and place processing were significantly more similar in MZ twins than in DZ twins. In contrast, there was no significant difference in the similarity of activation patterns associated with pseudoword and chair processing. Furthermore, there was a significant interaction between zygosity and stimulus category: zygosity had a significantly larger effect on similarity for the face and place conditions than for the pseudoword and chair conditions. These results were consistent whether the region was defined functionally or anatomically and whether activation was measured using response magnitude (β values) or statistical reliability (*t* values). They also cannot be attributed to differences in

anatomical similarity because the effect of zygosity was specific to a subset of stimulus categories. Moreover, the difference between MZ and DZ twins in anatomical similarity was consistent across different functional subregions of ventral visual cortex (areas most activated by each stimulus category), providing additional evidence that the observed effects of zygosity on functional similarity cannot be attributed to differences in anatomical similarity.

The results of this study demonstrate that genetics play a significant role in determining the cortical response to faces and places. Of course, these findings do not imply that experience plays no role in determining the observed activity. To take just one example, genes that affect social behavior could potentially lead some people to look at faces and places more than other people, and the resulting difference in experience could lead to changes in the neural circuitry (we thank one of the anonymous reviewers for this example). The results simply demonstrate that genetics do play a crucial role. The results also show that genetics play a significantly smaller role in determining the cortical response to visually presented orthographic stimuli. Overall, the findings are consistent with the view that the cortical substrates of face recognition and place recognition are partially innately specified, but that the cortical response to orthographic stimuli is more dependent on experience. Face and place recognition are older than reading on an evolutionary scale, they are shared with other species, and they provide a clearer adaptive advantage. It is therefore plausible that evolution would shape the cortical response to faces and places, but not orthographic stimuli.

References

Aguirre GK, Zarahn E, D'Esposito M (1998) An area within human ventral cortex sensitive to "building" stimuli: evidence and implications. Neuron 21:373–383.

Baker CI, Liu J, Wald LL, Kwong KK, Benner T, Kanwisher N (2007) Visual word processing and experiential origins of functional selectivity in human extrastriate cortex. Proc Natl Acad Sci USA 104:9087–9092.

Clifford P, Richardson S, Hemon D (1989) Assessing the significance of the correlation between two spatial processes. Biometrics 45:123–134.

Cohen J, Cohen J (2003) Applied multiple regression/correlation analysis for the behavioral sciences. Mahwah, NJ: L. Erlbaum.

Cohen L, Dehaene S, Naccache L, Lehericy S, Dehaene-Lambertz G, Henaff MA, Michel F (2000) The visual word form area: spatial and temporal characterization of an initial stage of reading in normal subjects and posterior split-brain patients. Brain 123:291–307.

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

Farah MJ, Wallace MA (1991) Pure alexia as a visual impairment: a reconsideration. Cognitive Neuropsychology 8:313–334.

Farah MJ, Rabinowitz C, Quinn GE, Liu GT (2000) Early commitment of neural substrates for face recognition. Cogn Neuropsychol 17:117–123.

Fisher SE, DeFries JC (2002) Developmental dyslexia: genetic dissection of a complex cognitive trait. Nat Rev Neurosci 3:767–780.

Gauthier I, Tarr MJ, Anderson AW, Skudlarski P, Gore JC (1999) Activation of the middle fusiform "face area" increases with expertise in recognizing novel objects. Nat Neurosci 2:568–573.

Gauthier I, Skudlarski P, Gore JC, Anderson AW (2000) Expertise for cars and birds recruits brain areas involved in face recognition. Nat Neurosci 3:191–197.

Haxby JV, Gobbini MI, Furey ML, Ishai A, Schouten JL, Pietrini P (2001) Distributed and overlapping representations of faces and objects in ventral temporal cortex. Science 293:2425–2430.

Kanwisher N, McDermott J, Chun MM (1997) The fusiform face area: a module in human extrastriate cortex specialized for face perception. J Neurosci 17:4302–4311.

Le Grand R, Mondloch CJ, Maurer D, Brent HP (2003) Expert face processing requires visual input to the right hemisphere during infancy. Nat Neurosci 6:1108–1112.

Maldjian JA, Laurienti PJ, Kraft RA, Burdette JH (2003) An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. Neuroimage 19:1233–1239.

Polk TA, Farah MJ (2002) Functional MRI evidence for an abstract, not perceptual, word-form area. J Exp Psychol Gen 131:65–72.

Polk TA, Stallcup M, Aguirre GK, Alsop DC, D'Esposito M, Detre JA, Farah MJ (2002) Neural specialization for letter recognition. J Cogn Neurosci 14:145–159.

Puce A, Allison T, Asgari M, Gore JC, McCarthy G (1996) Differential sensitivity of human visual cortex to faces, letterstrings, and textures: a functional magnetic resonance imaging study. J Neurosci 16:5205–5215.

Thompson PM, Cannon TD, Narr KL, van Erp T, Poutanen VP, Huttunen M, Lonnqvist J, Standertskjold-Nordenstam CG, Kaprio J, Khaledy M, Dail R, Zoumalan CI, Toga AW (2001) Genetic influences on brain structure. Nat Neurosci 4:1253–1258.