an independent cross-validation analysis (see **Supplementary Note 3**).

Using the non-independent analysis method, we replicated the results of Grill-Spector and colleagues (**Fig. 1b**), showing voxel populations apparently strongly selective for animals, cars and sculptures in addition to faces. Critically, however, this result does not survive a cross-validation test, in which the selection and test data are independent. Instead, only the face voxels maintain their preference in the independent test (**Fig. 1c**). Thus, the apparent selectivity for nonface categories in the non-independent analysis must have reflected random variation, not replicable selectivity. The bias in the non-independent analysis method can be seen most clearly when this method is used to analyze a region of interest that necessarily includes only noise because it is outside the brain (**Fig. 1d**,**e**): again, this method produces apparently selective voxel populations for all categories. Because their analysis method yields similar results when analyzing noise, the findings reported by Grill-Spector and colleagues provide no evidence for strong selectivity for nonfaces (or even faces) on the fusiform gyrus.

Grill-Spector and colleagues are correct that the apparent functional homogeneity of a given cortical region based on scans performed at one imaging resolution can be trumped by higher-resolution scans that demonstrate functional heterogeneity at a finer grain. However, they have not yet demonstrated that the FFA contains any voxels strongly selective for animals, cars or sculptures.

Note: Supplementary information is available on the Nature Neuroscience website.

Chris I Baker[1], Tyler L Hutchison[2] & Nancy Kanwisher[2]

[1]Laboratory of Brain and Cognition, National Institute of Mental Health, National Institutes of Health, 10 Center Drive, Building 10, Room 4C-104, MSC1366, Bethesda, Maryland 20892, USA. [2]McGovern Institute for Brain Research and Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139, USA.
e-mail: bakerchris@mail.nih.gov

1. Grill-Spector, K., Sayres, R. & Ress, D. *Nat. Neurosci.* **9**, 1177–1185 (2006).

Measuring selectivity in fMRI data

To the editor:

Using high-resolution functional magnetic resonance imaging (fMRI), Grill-Spector and colleagues[1] recently reported that many voxels in the fusiform face area (FFA) gave a numerically higher response to nonface stimuli than to faces. Importantly, the response to nonface objects appeared to show extremely high category selectivity. As such, these findings pose a serious challenge to current models of object representation in the ventral occipitotemporal cortex[2]. The formula used by Grill-Spector to calculate selectivity, however, vastly overestimates actual voxel selectivity.

Grill-Spector calculated voxel selectivity using the following formula: Selectivity = [Preferred – Nonpreferred] / [Preferred + |Nonpreferred|]. In this formula, "Preferred indicates the amplitude of the category that yielded the maximal response and Nonpreferred indicates the average amplitude of other categories" (p. 1184), with a selectivity value of 1 indicating maximum selectivity, and 0 indicating that a voxel has no preference for any of the stimulus categories tested. This formula is similar to the standard formula used in many electrophysiological studies, with the exception that the absolute value of the average nonpreferred response was used in the denominator. Unlike spike rate data, however, negative values are commonly observed in fMRI, and this has significant consequences when used with this formula. Consider the following. Suppose that the responses in a particular voxel were 0.11 to one category and 0.10, 0.08 and –0.20 to the other categories. Using Grill-Spector's formula, this voxel's selectivity equals 1.0, even though the response difference between the preferred and nearest non-preferred category equals only ~ 3% of the voxel's total response range (0.11 to –0.20). Clearly, this voxel is not highly selective for the preferred category. In fact, using this formula, anytime

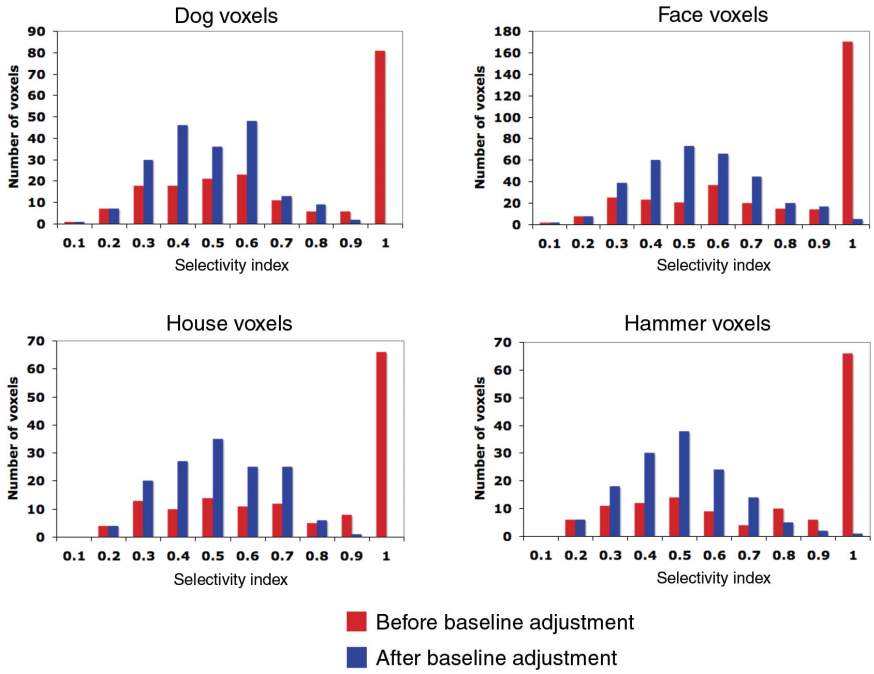

Figure 1 Histograms showing the selectivity index of FFA voxels with respect to each of four object categories, averaged across four subjects. Red bars show selectivity values before baseline adjustment. Blue bars show selectivity values for the same data after adjusting the signal baseline. (This analysis is an attempt to replicate the methods of Grill-Spector *et al.* with a corrected selectivity formula. No attempt was made to correct for the selection bias discussed by Baker *et al.* in the accompanying correspondence, which likely explains why the mean selectivity is comparable for all categories.)

© 2007 Nature Publishing Group http://www.nature.com/natureneuroscience

the response to even one category is negative, selectivity will be inflated.

To demonstrate this empirically, we scanned subjects using high-resolution fMRI (voxel size = 1.5 × 1.5 × 1.8 mm) while they viewed pictures of faces, dogs, houses and hammers. Using Grill-Spector's selectivity formula, we replicated the reported finding (red bars in **Fig. 1**). Indeed, the mean selectivity for FFA voxels in our data is nearly identical to the mean selectivity reported by Grill-Spector (0.73 versus 0.77, respectively).

To assess the selectivity distributions' sensitivities to negative values, we identified all FFA voxels in which the response to at least one category was less than 0 and added, to the response to each of the four categories, whatever value was necessary to make the smallest response across the categories equal to 0. For example, if the preferred response at a FFA voxel equaled 0.80, and the response in that same voxel to the three nonpreferred categories was 0.60, 0.60 and –1.2, we simply added 1.2 to each condition, resulting in a value of 2 for the preferred category and 1.8, 1.8 and 0 for the nonpreferred categories. Note that this procedure does not change the relationship between categories relative to the voxel's dynamic response range (0.80 to –1.2). After the baseline adjustment, we recalculated the selectivity exactly as before using Grill-Spector's selectivity formula, which is now equivalent to the standard formula. Rather than clustering near 1.0, the baseline-adjusted selectivity distributions are more Gaussian in shape, with a mean selectivity of 0.47 (blue bars in **Fig. 1** and **Supplementary Fig. 1** online).

It is possible that the FFA contains regions highly selective for nonface stimuli. This remains an open question, however, because the formula used by Grill-Spector vastly overestimates voxel selectivity. Adjusting the signal baseline to correct for negative responses reveals selectivity distributions that are very different from those they reported, and we encourage authors to consider other methods for calculating selectivity[3].

Note: Supplementary information is available on the Nature Neuroscience website.

W Kyle Simmons, Patrick S F Bellgowan & Alex Martin

Laboratory of Brain and Cognition, National Institute of Mental Health, Building 10, Room 4C-104, 10 Center Drive MSC 1366, Bethesda, Maryland 20892-1366, USA.
e-mail: alexmartin@mail.nih.gov

1. Grill-Spector, K, Sayers, R., & Ress, D. *Nat. Neurosci.* **9**, 1177–1185 (2006).
2. Haxby, J.V. *Nat. Neurosci.* **9**, 1084–1086 (2006).
3. Afraz, S.R., Kiani, R. & Esteky, H. *Nature* **442**, 692–695 (2006).

© 2007 Nature Publishing Group http://www.nature.com/natureneuroscience