

Temporal stability of visually selective responses in intracranial field potentials recorded from human occipital and temporal lobes

Arjun K. Bansal, Jedediah M. Singer, William S. Anderson, Alexandra Golby, Joseph R. Madsen and Gabriel Kreiman

J Neurophysiol 108:3073-3086, 2012. First published 5 September 2012;
doi: 10.1152/jn.00458.2012

You might find this additional info useful...

This article cites 48 articles, 19 of which you can access for free at:
<http://jn.physiology.org/content/108/11/3073.full#ref-list-1>

Updated information and services including high resolution figures, can be found at:
<http://jn.physiology.org/content/108/11/3073.full>

Additional material and information about *Journal of Neurophysiology* can be found at:
<http://www.the-aps.org/publications/jn>

This information is current as of December 13, 2012.

Temporal stability of visually selective responses in intracranial field potentials recorded from human occipital and temporal lobes

Arjun K. Bansal,¹ Jedediah M. Singer,¹ William S. Anderson,² Alexandra Golby,³ Joseph R. Madsen,⁴ and Gabriel Kreiman^{1,5,6}

¹Department of Ophthalmology, Boston Children's Hospital, Harvard Medical School, Boston, Massachusetts; ²Department of Neurosurgery, Johns Hopkins School of Medicine, Baltimore, Maryland; ³Department of Neurosurgery, Brigham and Women's Hospital, Boston, Massachusetts; ⁴Department of Neurosurgery, Boston Children's Hospital, Harvard Medical School, Boston, Massachusetts; ⁵Center for Brain Science, Harvard University, Cambridge, Massachusetts; and ⁶Swartz Center for Theoretical Neuroscience, Harvard University, Cambridge, Massachusetts

Submitted 31 May 2012; accepted in final form 1 September 2012

Bansal AK, Singer JM, Anderson WS, Golby A, Madsen JR, Kreiman G. Temporal stability of visually selective responses in intracranial field potentials recorded from human occipital and temporal lobes. *J Neurophysiol* 108: 3073–3086, 2012. First published September 5, 2012; doi:10.1152/jn.00458.2012.—The cerebral cortex needs to maintain information for long time periods while at the same time being capable of learning and adapting to changes. The degree of stability of physiological signals in the human brain in response to external stimuli over temporal scales spanning hours to days remains unclear. Here, we quantitatively assessed the stability across sessions of visually selective intracranial field potentials (IFPs) elicited by brief flashes of visual stimuli presented to 27 subjects. The interval between sessions ranged from hours to multiple days. We considered electrodes that showed robust visual selectivity to different shapes; these electrodes were typically located in the inferior occipital gyrus, the inferior temporal cortex, and the fusiform gyrus. We found that IFP responses showed a strong degree of stability across sessions. This stability was evident in averaged responses as well as single-trial decoding analyses, at the image exemplar level as well as at the category level, across different parts of visual cortex, and for three different visual recognition tasks. These results establish a quantitative evaluation of the degree of stationarity of visually selective IFP responses within and across sessions and provide a baseline for studies of cortical plasticity and for the development of brain-machine interfaces.

prosthetic applications; temporal stability; visual cortex; visual object recognition; field potentials

OUR BRAINS PROVIDE A SUBSTRATE for learning and adaptation while at the same time maintaining a stable representation of information over long time periods. The neurophysiological signals underlying this representation can be studied at multiple spatial scales and multiple temporal scales (from milliseconds to years). Partly due to technical limitations, our understanding of neurophysiological signals and changes over long periods is rather limited. Most single-neuron and local field potential studies have focused on the “instantaneous” neuronal preferences examined over periods of seconds to minutes to hours.

Several investigators have begun to use chronic multi-electrode arrays to record action potentials over periods spanning days or, in some cases, even months (e.g., Nicolelis et al. 2003; Porada et al. 2000). These chronic recordings give rise to the

possibility of evaluating the extent to which neuronal preferences change over long periods. Some studies recording either action potentials or field potentials in the monkey's motor and visual system have described stable neuronal signals over the course of several weeks (Bondar et al. 2009; Chao et al. 2010; Nicolelis et al. 2003; Suner et al. 2005; Tolias et al. 2007). In the visual system, single units from monkey primary visual cortex followed over several days maintained their orientation tuning preferences (Tolias et al. 2007). A recent study at the other end of the visual cortical hierarchy showed that neurons in macaque areas TE and superior temporal sulcus (STS) maintain their visual selectivity over at least a couple of weeks (Bondar et al. 2009). These results suggest that visual functions such as object recognition could rely on stable neuronal representations despite inherent neuronal variability and plasticity in the neocortical circuitry.

Little is known about the temporal stability of neurophysiological preferences in response to visual stimuli along the human ventral visual stream at a spatial scale of ensembles of neurons. Here, we set out to examine the feature and tuning stability of visually selective intracranial field potential (IFP) signals recorded from the human neocortex in subjects implanted with electrodes for clinical reasons (Engel et al. 2005; Kreiman 2007). We studied IFP responses to brief flashes of visual stimuli while subjects performed three object discrimination tasks (Agam et al. 2010; Liu et al. 2009) in multiple recording sessions spanning 2 h to over 200 h. Despite minor variations across sessions, we found a remarkable degree of stability in visual IFP responses over multiple days. The quantification presented here provides a rigorous comparison for future studies of plasticity and also for the training and development of decoding algorithms used in brain-machine interfaces.

METHODS

Subjects

Subjects were 27 patients (13 females, 23 right handed, ages 8.5–46.5 yr) with pharmacologically intractable epilepsy. The subjects were admitted to Boston Children's Hospital or Brigham and Women's Hospital to localize their seizure foci for potential surgical resection. All the experiments described in this report were approved by the respective Institution Review Boards at each hospital and were carried out with each subject's informed consent. Data from 15 of

Address for reprint requests and other correspondence: G. Kreiman, 1 Blackfan Cr., Karp 11, Boston, MA 02115 (e-mail: gabriel.kreiman@childrens.harvard.edu).

these patients have been reported previously in other studies (Agam et al. 2010; Liu et al. 2009). For this study, we only analyzed data from subjects who performed the same task in at least two sessions separated by at least 1 h.

IFP Recordings

The recording procedures were as described previously (Agam et al. 2010; Liu et al. 2009). Briefly, subjects were implanted with intracranial electrodes (Ad-Tech, Racine, WI; 2.3-mm diameter, 1-cm separation between electrode centers, impedance $\ll 1$ k Ω) to localize the seizure foci. Electrode contacts were arranged in grids or strips containing between 4 and 64 contacts. The total number of recording sites per subject ranged from 34 to 126 (80 ± 23 , mean \pm SD). The signal from each recording site was amplified ($\times 2,500$) and sampled at 256, 500, or 2,000 Hz (XLTEK, Oakville, ON, Canada; Bio-Logic, Knoxville, TN). A notch filter was applied at 60 Hz, along with a bandpass filter from 0.1 to 100 Hz. Throughout the text, we refer to the recorded signal as IFP. The time interval between two consecutive sessions was between 1.03 and 202.25 h (20.5 ± 29.8 h, mean \pm SD). The time interval between the first and last recorded session was between 2.58 and 247.87 h (41.1 ± 47.6 h, mean \pm SD). Electrode localization was performed by aligning each subject's MRI into Talairach space, coregistering computed tomography (CT) and magnetic resonance (MR) images, and assigning each electrode to 1 of 74 different regions with FreeSurfer software (Dale et al. 1999) based on the 2009 atlas (Destrieux et al. 2010). In 1 of 27 subjects, the coregistration was not performed due to incomplete data. By coregistering the preoperative structural MR with the CT images showing the electrodes, we find out each electrode's locations. We report anatomic regions for all electrode locations in Table 4. As reported previously (e.g., Liu et al. 2009), only a subset of all the electrode locations yielded visually selective responses as defined below (numbers presented below).

Stimulus Presentation and Tasks

We considered three different tasks (Table 1). In all cases, an image was presented on a laptop computer and subjects were asked to perform a recognition task dependent on the image. The stimulus presentation times, sizes and behavioral responses varied with the task and are described below and in Table 1. In all cases, images were contrast normalized and the presentation order was pseudorandomized. Here, we only consider those cases where we repeated the same task in at least two recording sessions separated by at least 1 h. Separate recording sessions included the same stimuli and tasks (the presentation order was different across sessions). We recorded IFP responses from all implanted electrodes during the tasks.

Task 1: single object presentation. The data from this task were described in Liu et al. (2009). Nine of the 11 subjects in the original study were considered in this study (in the other subjects, there was only 1 session). Subjects were presented with 25 objects (5 categories \times 5 exemplars); each object was presented for 200 ms, and there was a

600-ms blank in between images. Subjects performed a one-back matching task indicating whether or not an object was repeated. The mean behavioral performance was $86 \pm 7\%$.

Task 2: object pairs. The data from this task were described in Agam et al. (2010). Six of the nine subjects in the original study were considered in this study (in the other subjects, there was only 1 session). Subjects were presented with 25 objects (5 categories \times 5 exemplars). Here, we analyzed only those trials that included either a single object or two objects from the same category. In exemplar-level analyses (see Figs. 2 and 3), we only analyzed single-object presentation trials. Subjects performed a target identification task indicating whether the image contained an object belonging to a prespecified target category (target category changed in blocks of 50 trials). The mean behavioral performance was $92 \pm 12\%$.

Task 3: asynchronous object presentation. Twelve subjects were presented with five exemplars. Subjects were asked to indicate with a button press whether a test image matched whole images or image halves presented in the middle of flickering noise. Each trial started with a 500-ms fixation period. After fixation, phase-scrambled flickering noise was presented at 10–30 Hz (depending on the subject). After 500 ms, a whole image or an image half was presented in place of one noise image (33–100 ms), followed by more flickering noise. In those trials where an image half was presented, the complementary image half was presented after another 33–300 ms. Five hundred milliseconds after the onset of the final piece of image content, the flickering noise disappeared and the test image appeared. Subjects reported whether or not the test image matched the previously shown image content. The test image remained visible until the subject responded, at which point a high-contrast mask was flashed, followed by the fixation point for the next trial. Here, we considered only the IFP responses to the test images. Because there were only five exemplars, the analyses shown in Figs. 2 and 3 do not include data from this task. Performance was $70 \pm 9\%$ ($75 \pm 14\%$ for the easiest condition with whole images presented at 0-ms asynchrony).

Data Analyses

Responsiveness. We first evaluated whether each electrode showed a significant deviation in IFP elicited by the visual stimulus onset irrespective of the stimulus identity. For each electrode and in each recording session, we compared the IFP response amplitude [$\max(\text{IFP}) - \min(\text{IFP})$] in the baseline period (defined from 100 ms before stimulus onset until 50 ms after stimulus onset) against the IFP response amplitude [$\max(\text{IFP}) - \min(\text{IFP})$] from 100 to 250 ms after stimulus onset. To avoid potential artifacts due to movement or interictal discharges, trials in which the visual response amplitude was more than four standard deviations greater than the mean response amplitude across all trials were discarded as potential outliers. On average, 2.2 ± 2.5 trials ($0.44 \pm 0.41\%$) were discarded on the basis of this criterion. An electrode was considered visually "responsive" if the stimulus and baseline amplitudes were significantly different (2-sided *t*-test, $P < 0.01$) in at least one session. No information about the identity of the visual stimulus was used to define responsiveness.

Table 1. Summary of experimental parameters for each of the three tasks

	Task 1	Task 2	Task 3
Task name	Single objects	Object pairs	Asynchronous objects
No. of subjects	9	6	12
No. of stimuli	5 categories \times 5 exemplars	5 categories \times 5 exemplars	5 exemplars
Stimulus presentation time, ms	200	100	$1,024.6 \pm 256$
Stimulus size, deg visual angle	$\sim 1.5\text{--}6$	~ 3.4	$\sim 3.8\text{--}10.8$
Blank time, ms	600	1,000	500 (flickering Fourier noise)
Subject response	One-back match	Target identification	Image match

Data for single objects were described in Liu et al. (2009), and data for object pairs were described in Agam et al. (2010). The presentation time is variable (mean \pm SD) in task 3 because the stimulus was presented until the subject responded.

Selectivity. An electrode was considered to be visually “selective” if it satisfied all the following conditions in at least one session: 1) the mean IFP response amplitude to the preferred category between 50 and 300 ms after stimulus onset was $>50 \mu\text{V}$; 2) the across-category IFP amplitude response variance was significantly greater than the within-category IFP amplitude response variance (1-way ANOVA, $P < 0.01$, $df = 4$, $N = 473 \pm 317$ trials per electrode per session); and 3) decoding performance was significant ($P < 0.01$; see below for details on decoding algorithm). These analyses were independent of the ones described above to determine whether an electrode was visually responsive or not.

Single-trial decoding analyses. For each trial, the IFP waveform between 50 and 300 ms after image presentation was extracted. Principal components analysis (PCA) was performed to reduce the dimensionality of the waveforms across all trials. Briefly, eigenvector decomposition was performed for the waveforms extracted from all trials. The waveforms were then projected onto the top n eigenvectors, where n was the number of eigenvectors that explained 95% of the variance in the data. On average, $n = 8.6 \pm 3.3$ eigenvectors were used. We used a linear discriminant analysis (LDA) classifier to determine the image content on the basis of physiological responses on single trials. We followed a cross-validation procedure whereby the classifier was trained on 70% of the data, and the decoding performance results shown throughout were evaluated using the test data in the remaining 30% of the data. Depending on the question, different sessions and data sets were used for training and testing (see Figs. 6 and 7).

Decoding performance (defined as the fraction of correctly classified trials) was separately analyzed for each channel and category. In *task 3*, instead of category level decoding, we examined exemplar-level decoding (since 5 objects, instead of 5 objects per category, were used). In the analysis shown in Fig. 6, *A–G*, we used a binary classifier where chance performance was 0.5. For each binary classifier, we randomly subsampled the data to use an equal number of trials from the category to be decoded (50%) and the four other categories (50%). To ensure that potential changes in decoding performance across sessions were not merely due to different numbers of training trials across sessions, we randomly subsampled the data to use the same numbers of trials across sessions. To compute the mean decoding performance, 50 cross-validation iterations were run for each classifier. In each iteration, a subsample of trials was selected to evaluate decoding performance as described above. Data were randomly assigned to the training set or the test set. There was no overlap between the training data and the test data to avoid overfitting.

To determine whether the decoding performance values were significantly different from chance levels, we performed the same procedure described above after randomly shuffling the labels in the training data. We computed the distribution of decoding performance values after shuffling using 1,000 iterations. If the empirical probability of observing a mean decoding performance greater than the actual mean decoding performance in those 1,000 iterations was <0.01 , then the electrode was considered to have significant decoding performance for that category. The mean and standard deviation of the decoding performance values with the shuffled labels were 0.500 and 0.023, respectively.

In addition to considering the principal components of the waveforms, we also evaluated decoding performance in single trials using other features of the data, including different frequency bands. A three-pole Butterworth filter was used for each of the following frequency bands: <4 , 4–8, 8–12, 12–35, 35–50, and 70–100 Hz. The feature plots in Fig. 4 show the average log power (power evaluated as the squared amplitude of the corresponding filtered signal) in the 50- to 300-ms window for the preferred category (as determined using PCA–LDA-based decoding). For Fig. 4, *G–I*, the average log power in the corresponding frequency band during a baseline period from 100 to 0 ms before stimulus onset for each trial was subtracted. For the decoding procedure using each frequency band, we used 5 features

comprising the average log power in 5 nonoverlapping bins of 50 ms each from 50 to 300 ms after image onset.

The best decoding performance values were typically obtained using PCA on the IFP waveform (see Fig. 6*A*). This is not surprising given that the identity of the electrodes that passed the selectivity criteria depended on PCA-based decoding. Nevertheless, for this study the key analysis is not the absolute performance of the classifier, but rather the degree of stationarity in decoding performance across sessions once a criterion for selectivity is defined.

In Fig. 6*H*, we also considered a multiclass classifier where the task was to determine in each trial which of the five different object categories was presented (Hung et al. 2005, Liu et al. 2009). For multiclass classifiers, we randomly sampled from all five categories (20% each); hence, chance decoding performance is 0.20. We used the same number of trials across sessions to ensure that potential changes in decoding performance across sessions were not merely due to different numbers of training trials. To compute the mean decoding performance, 50 cross-validation iterations were run for each classifier. In each iteration, we performed cross-validation by selecting a subsample of trials to evaluate decoding performance as described above. Data were randomly assigned to the training set or the test set. There was no overlap between the training data and the test data to avoid overfitting. The chance level performance was determined with 1,000 iterations of the above procedure but with shuffled labels (0.20 ± 0.014).

Stationarity. An electrode was defined as “stationary” if it satisfied the criteria for selectivity in two or more sessions. It was defined as “nonstationary” if it satisfied the criteria for selectivity in only one session. If an electrode showed a selective response to more than one category according to the above criteria, then the category that yielded the maximum decoding performance in the first selective session was considered for analyzing stationarity. Therefore, in all the analyses, we are comparing the responses to the same exemplars or categories across sessions, and we do not consider an electrode to be stationary if it responded only to one exemplar/category in one session and only to a different exemplar/category in another session.

We also analyzed stationarity at the exemplar level. This analysis was run on *task 1* and *task 2* data (see Figs. 2 and 3) where we used a degree of overlap measure to quantify stationarity. “Overlap” was defined as the number of exemplars in the top 5 (of 25) IFP responses in each session that were common across 2 consecutive sessions: $\text{Overlap}_{1,2} = \# \{ {}_s S_1 \cap {}_s S_2 \} / 5$, where ${}_s S_1$ and ${}_s S_2$ indicate the set of top 5 exemplars in recording *session 1* and recording *session 2*, respectively. $\text{Overlap}_{1,2}$ takes a value of 1 if the top 5 exemplars are the same in both sessions and a value of 0 if the top exemplars in *session 1* are completely different from those in *session 2*. The expected degree of overlap by chance is 0.2.

RESULTS

We recorded IFP responses from 2,195 electrodes recorded from 27 subjects in 3 different visual recognition tasks. In all three tasks, a contrast normalized grayscale visual stimulus was briefly presented on the screen and subjects were asked to make a behavioral response depending on the identity of the stimulus. The stimuli, presentation times, visual angles, and behavioral responses varied across the different tasks and are described in METHODS and summarized in Table 1. Here, we analyzed recordings where the same task was performed in two or more recording sessions (3.0 ± 1.1 sessions, mean \pm SD; mode 2, range 2–6 sessions). The interval between sessions ranged from 1.0 to 202.3 h (20.5 ± 29.8 h, mean \pm SD; median 14.4 h). There were $1,650 \pm 847$ trials per electrode summing across sessions. The stimulus order was pseudorandomized and was therefore different in each recording session.

Table 2. Number of total electrodes, visually selective electrodes, and stationary and nonstationary electrodes

	Task 1	Task 2	Task 3	Total
Total no. of electrodes	760	448	987	2,195
Total no. of selective electrodes	33	18	40	91
No. of stationary electrodes	30	15	36	81
No. of nonstationary electrodes	3	3	4	10

An electrode was considered visually selective if a 1-way ANOVA test across categories yielded $P < 0.01$, decoding performance using a statistical classifier yielded $P < 0.01$, and the response amplitude was $>50 \mu\text{V}$. Further details about the criteria used to place an electrode in each category are described in METHODS. In Figs. 4–6, but not in this table, when there were more than 2 sessions for an electrode that was selective in at least 1 session, each pair of consecutive sessions contributed 1 point.

To evaluate the degree of temporal stability in the stimulus-evoked responses, we considered those electrodes that showed strong visually selective responses, that is, those electrodes that discriminated among the different images. Other electrodes showed either no response or stimulus-evoked responses that were not selective or only weakly selective (Agam et al. 2010; Liu et al. 2009). The majority (1,877 of 2,195) of the electrodes were not visually responsive (see METHODS for definition of responsiveness). The temporal stability for these electrodes is hard to define in the context of visual responses (e.g., most electrodes were located in nonvisual areas and yielded consistent flat responses across sessions). Considering the electrodes that showed a visually evoked response ($n = 318$; see METHODS), there was a strong correlation in the IFP waveforms across sessions: 0.62 ± 0.02 (mean \pm SE). The most stringent and interesting test of stability is based on evaluating whether selective responses are maintained over time. The degree of stability in the responses of electrodes showing marginal selectivity can be strongly influenced by signal-to-noise ratios and is more difficult to interpret. Therefore, we used strict criteria for visual selectivity (defined in METHODS) and focused on 91 electrodes that could discriminate among the different stimuli. Eighty-six of these 91 electrodes (94.5%) also satisfied the criteria for selectivity when a multiclass classifier instead of a binary classifier was used for decoding (see METHODS). Importantly, selectivity was analyzed separately in each session, and therefore the degree of stability or lack of stability was not a criterion used to select the electrodes for analyses. The use of different criteria to determine visual selectivity yielded qualitatively similar results. The number of electrodes showing visual selectivity for each task is shown in Table 2. The most

common locations for the electrodes showing visual selectivity were inferior occipital gyrus and sulcus, fusiform gyrus, inferior temporal gyrus, occipital pole, parahippocampal gyrus, and middle occipital gyrus (Table 3).

Examples of three visually selective responses are shown in Fig. 1 (A and B, task 1; C, task 2). In Fig. 1A, we illustrate the selective responses of an electrode located in the left temporal pole. The two subplots show the responses recorded in two sessions that were 21 h apart. To analyze the visual responses, we considered the interval between 50 and 300 ms after stimulus onset (Hung et al. 2005; Liu et al. 2009). The overall response properties were consistent across the two sessions. The preferred category was the same in the two sessions (indicated by the cyan color) and the average amplitude to the preferred category was $408 \pm 109 \mu\text{V}$ (mean \pm SD; $n = 48$) in the first session and $391 \pm 87 \mu\text{V}$ ($n = 78$) in the second session. The times of the response peaks in the 50- to 300-ms interval for the preferred category were 244 ± 56 and 226 ± 65 ms in sessions 1 and 2, respectively. The times of the response minima in the same interval for the preferred category were 166 ± 84 and 143 ± 75 ms in sessions 1 and 2, respectively. The responses were not identical: for example, the red curve shows a distinct trough in the second session, particularly after 300 ms. Figure 1B shows an example from the same task (task 1) for an electrode in a different location (right fusiform gyrus) and with different stimulus preferences. An example from task 2 is shown in Fig. 1C. The insets in Fig. 1 show responses from adjacent or nearby electrodes (within 1 or 2 cm). Although the biophysical origin of the IFP signals is not clearly understood, the distinct IFPs from those nearby electrodes suggest that the responses are relatively local and do not reflect changes in large volumes of cortex (see also Vidal et al. 2012). Figure 1 shows responses averaged over all the stimuli belonging to a given category. In Fig. 2, we show the responses from the same electrodes in Fig. 1, A and B, to each of the 25 different exemplar objects in each category. Although in some cases there were clear differences across sessions (see e.g., responses to the 1st exemplar in the 4th row in Fig. 2A), overall, the responses during the two sessions were quite consistent (see e.g., responses to all exemplars in the 3rd row in Fig. 2B). These examples suggest a strong overall degree of stability in the visually selective responses.

To begin to quantify the degree of stability in the responses across sessions, we asked whether there was any overlap between the top five exemplar images that elicited the strongest responses

Table 3. Electrode locations

Locations With ≥ 4 Selective Electrodes	Total No. of Electrodes	Total No. of Selective Electrodes	No. of Stationary Electrodes	No. of Nonstationary Electrodes
Inferior occipital gyrus and sulcus	26	18	17	1
Fusiform gyrus	64	17	17	0
Inferior temporal gyrus	194	14	12	2
Occipital pole	27	9	9	0
Parahippocampal gyrus	39	8	7	1
Middle occipital gyrus	40	6	6	0
Total	390	72	68	4
Other locations (< 4 selective electrodes)	1,805	19	13	6

Data are only reported for those locations with ≥ 4 visually selective electrodes to avoid reporting a percentage for locations with a very small number of electrodes. The locations reported encompass 79% (72/91) of the visually selective electrodes and 84% (68/81) of the stationary electrodes. Table 4 reports the locations of all recorded electrodes (see METHODS).

Table 4. Location of all electrodes analyzed in this study conveying the range of sampled locations

Area	<i>n</i>	Area	<i>n</i>
Middle temporal gyrus	230	Transverse frontopolar gyri and sulci	10
Inferior temporal gyrus	194	Superior temporal sulcus	9
Lateral aspect of superior temporal gyrus	169	Parieto-occipital sulcus	8
Temporal pole	128	Posterior-dorsal part of the cingulate gyrus	6
Supramarginal gyrus	97	Posterior-ventral part of the cingulate gyrus	6
Middle frontal gyrus	96	Temporal plane of the superior temporal gyrus	6
Orbital gyri	81	Central sulcus	6
Precentral gyrus	77	Anterior part of the cingulate gyrus and sulcus	5
Angular gyrus	77	Planum polare of the superior temporal gyrus	5
Superior frontal gyrus	70	Posterior transverse collateral sulcus	4
Subcentral gyrus (central operculum) and sulci	69	Middle-anterior part of the cingulate gyrus and sulcus	4
Fusiform gyrus	64	Calcarine sulcus	4
Triangular part of the inferior frontal gyrus	58	Inferior part of the precentral sulcus	4
Lingual gyrus	57	Inferior temporal sulcus	4
Opercular part of the inferior frontal gyrus	56	Paracentral lobule and sulcus	3
Postcentral gyrus	55	Short insular gyri	3
Middle occipital gyrus	40	Middle frontal sulcus	2
Parahippocampal gyrus	39	Superior frontal sulcus	2
Occipital pole	27	Sulcus intermedius primus (of Jensen)	2
Inferior occipital gyrus and sulcus	26	Lateral orbital sulcus	2
Superior parietal gyrus	24	Subparietal sulcus	2
Precuneus	23	Middle-posterior part of the cingulate gyrus and sulcus	1
Orbital part of the inferior frontal gyrus	17	Anterior transverse temporal gyrus (of Heschl)	1
Orbital sulci	17	Middle occipital sulcus and lunatus sulcus	1
Fronto-marginal gyrus (of Wernicke) and sulcus	16	Anterior occipital sulcus	1
Cuneus	13	Lateral occipito-temporal sulcus	1
Inferior frontal sulcus	13	Medial occipito-temporal sulcus	1
Gyrus rectus	12	Postcentral sulcus	1
Anterior transverse collateral sulcus	12	Remaining	223
Superior occipital gyrus	11	Total	2,195

Further details about each of these locations and the corresponding anatomical landmarks can be found in Destrieux et al. (2010). Average Talairach coordinates for the electrodes in *task 1* were reported in Liu et al. (2009).

(METHODS). For this purpose, we considered the amplitude of the IFP signals between 50 and 300 ms. For the example electrode shown in Figs. 1A and 2A, four of the top five exemplars were the same between the two sessions (overlap = 0.8; Fig. 3A), whereas for the example electrode in Figs. 1B and 2B, all five top exemplars were the same across the two sessions (overlap = 1.0; Fig. 3B). The distribution of the degree of overlap across sessions for all the visually selective electrodes is shown in Fig. 3, C (*task 1*) and D (*task 2*). In *task 3*, there were only five exemplars, and therefore we did not perform this analysis. These distributions were significantly different from the distribution expected by chance (chi-squared goodness of fit test: $P < 10^{-6}$). The mean degree of overlap was 0.49 ± 0.29 (*task 1*) and 0.43 ± 0.23 (*task 2*). Selecting the top 3 exemplars or the top 10 exemplars also yielded a degree of overlap that was significantly above chance levels. To put these values in perspective, we considered the degree of short-term stability due to fluctuations within a session. Trial-to-trial variability within a session constrains the amount of stability that can be expected over long time scales. We focused on *task 1* (there were not enough trials for this analysis in *task 2*) and computed the degree of short-term stability by comparing the top five exemplars from the first half of each session with those in the second half. The mean degree of overlap was 0.49 ± 0.26 . These results suggest that the stability in exemplar preferences across sessions was as strong as the stability within a session.

To further quantify the stability of the IFP response waveforms across sessions, we considered 9 different features extracted from the 50- to 300-ms interval. These features were

the IFP amplitude [$\max(\text{IFP}) - \min(\text{IFP})$], time of maximum IFP, time of minimum IFP, and the log power (see METHODS) in 6 different frequency bands: <4, 4–8, 8–12, 12–35, 35–50, and 70–100 Hz. In Fig. 4, we compare these features between consecutive sessions. We evaluated the degree of stability for each feature by computing the coefficient of determination (R^2) across consecutive sessions. Most of these features yielded strong R^2 values ranging from 0.69 to 0.93, except for the 35- to 50-Hz frequency band ($R^2 = 0.44$). The slopes ranged from 0.79 to 0.93 (except 0.53 for the 35- to 50-Hz frequency band). A slope of 1 would indicate identical features across sessions. We compared these values with the within-session short-term stability to estimate the upper limit in the R^2 and slope values imposed by trial-to-trial variability. When comparing the feature values from the first half of each session with those in the second half (data not plotted in Fig. 4), the R^2 values ranged from 0.65 to 0.96, except for the 35- to 50-Hz frequency band ($R^2 = 0.41$). The corresponding slopes ranged from 0.8 to 0.99, except for the 35- to 50-Hz frequency band (slope = 0.7). The degree of stability as evaluated by the R^2 values and the slopes for each of these features were comparable for the three different tasks (compare the 3 different colors in Fig. 4).

We analyzed the responses as a function of the time elapsed between recording sessions. There was no clear increase in the fraction of nonstationary recordings over time (Fig. 5A). In fact, Fig. 5A reveals a small decrease in the fraction of nonstationary session pairs; we assume that this decrease reflects random variation given the small number of nonstationary electrodes ($n = 10$; Table 2). We also analyzed whether

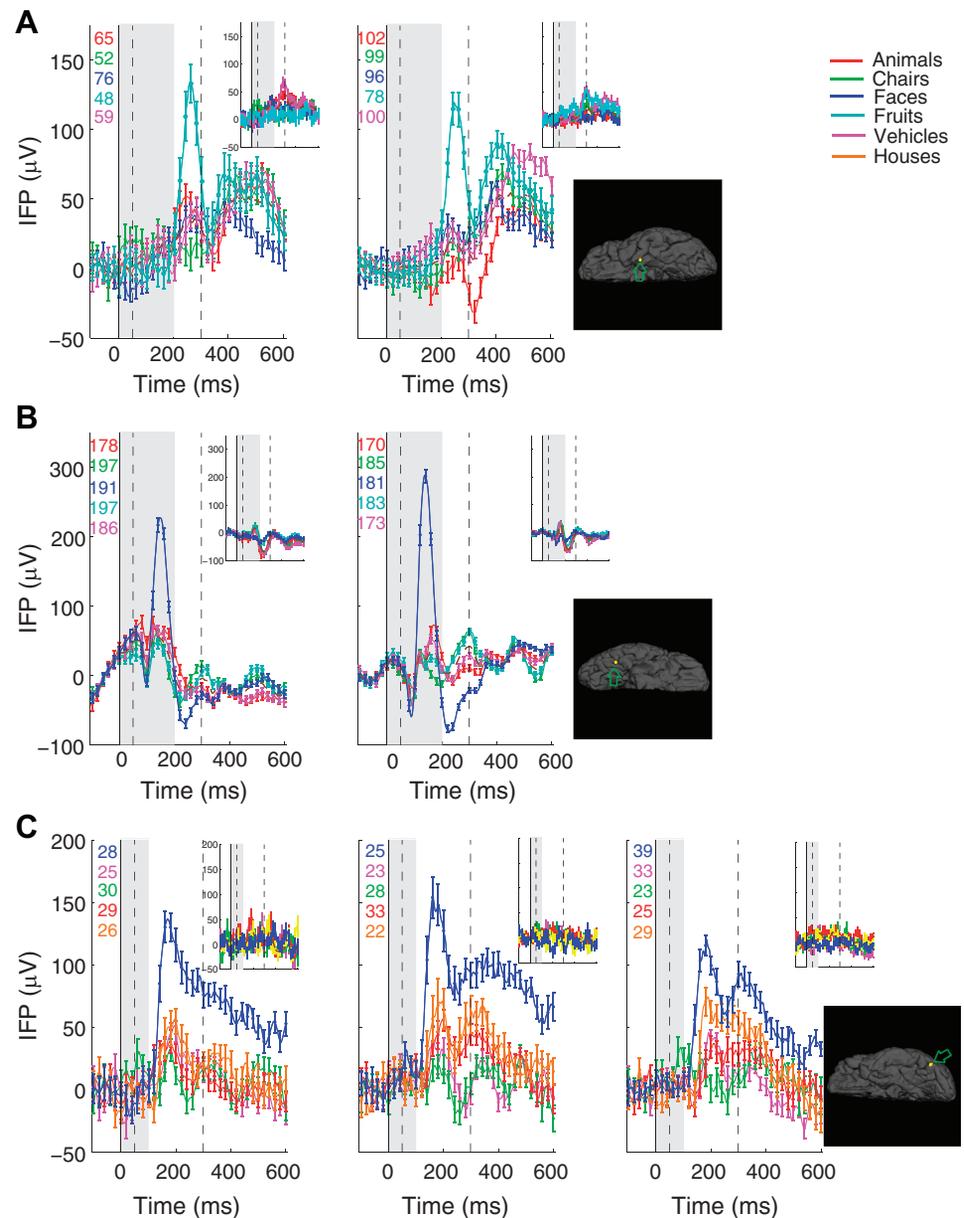


Fig. 1. Examples of visual category selectivity across sessions for 3 electrodes. Each row indicates a separate electrode from a different subject; the examples in *A* and *B* were from *task 1*, and the example in *C* was from *task 2* (METHODS and Table 1). Each column indicates a different session; the times between sessions were 21 h (*A*), 70 h (*B*), and 46 and 24 h (*C*). Each curve shows the average intracranial field potential (IFP) response to all images from a given category (red, animals; green, chairs; blue, faces; cyan, fruits; purple, vehicles; orange, houses). Error bars indicate \pm SE, and the numbers of trials for each category contributing to the mean waveform are indicated in each subplot. The gray rectangle denotes the image presentation time. Vertical dashed lines denote the time period used for the analyses in the text (50 to 300 ms after stimulus onset). Each of these electrodes showed a statistically significant selective response ($P < 0.01$, 1-way ANOVA; see METHODS). The *inset* waveforms depict the IFP responses of an adjacent (*A* and *C*) or nearby electrode (*B*) that did not demonstrate the same selectivity as seen in the featured electrodes. The *inset* brain images depict the electrode locations.

there was any overall trend when plotting each of the IFP response features as a function of the time elapsed between sessions (Fig. 5, *B–J*). For this purpose, we fitted a line to each of these plots (dashed lines in Fig. 5, *B–J*), and the slopes of these lines were typically close to 0 (Fig. 5, see legend). There was a weak but significant ($P < 0.01$) attenuation with time in the IFP amplitude (slope = $-0.044/h$; Fig. 5*B*) and power in the <4 -Hz (slope = $-0.226/h$; Fig. 5*E*), 4- to 8-Hz (slope = $-0.116/h$; Fig. 5*F*), and 8- to 12-Hz bands (slope = $-0.133/h$; Fig. 5*F*) as well as in the 35- to 50-Hz bands (slope = $-0.102/h$; Fig. 5*H*). For the other frequency bands, the slope was not significantly different from 0. Overall, these results show that several different features of the IFP responses for visually selective electrodes remain stable over periods of 1 to 200 h. To a good first approximation, this stability did not depend on the specifics of the task.

Neural signals elicited by a visual stimulus during a given trial need to be read out by target brain regions. If the neural signals representing visual information change over time, then

the decoding process also needs to adapt to those transformations. Following up on previous efforts (see reviews and discussions in Kriegeskorte and Kreiman 2011), we built a machine learning classifier to decode the IFP signals from each electrode in individual trials. We considered multiple different features of the IFP waveform, including principal components or power in different frequency bands (Fig. 6). We used a linear discriminant analysis classifier to decode the presence or absence of the preferred category (or preferred exemplar in *task 3*) for each selective electrode in single trials (see METHODS). A decoding performance of 0.5 corresponds to chance levels and a decoding performance of 1.0 would indicate perfect classification in each single trial. We first asked whether the performance of such a classifier would be consistent in multiple recording sessions. We determined the within-session decoding performance for each session (using subsampling for cross-validation to avoid overfitting; see METHODS). To evaluate the statistical significance of the decoding performance values that we obtained, we compared the actual performance with

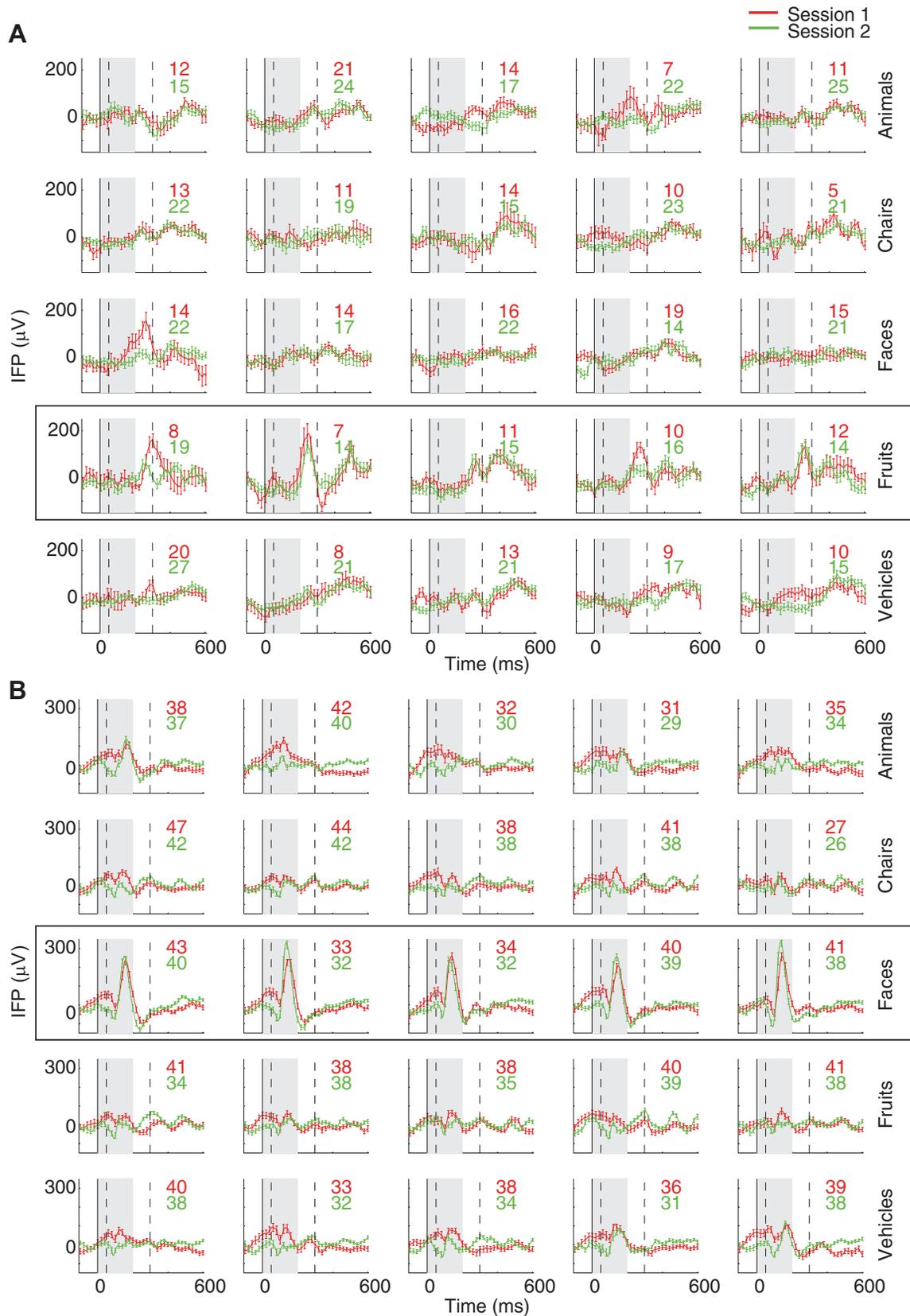


Fig. 2. Responses to each exemplar for the electrodes in Fig. 1, A and B. Each subplot shows the mean IFP response (\pm SE) to a particular exemplar object during the first (red) or second recording session (green). The number of trials averaged is shown in each subplot. The 5 subplots in each row correspond to the 5 exemplars that comprised a visual category. Other formats and conventions are similar to those of Fig. 1. Boxes highlight the category that elicited the strongest responses for each electrode.

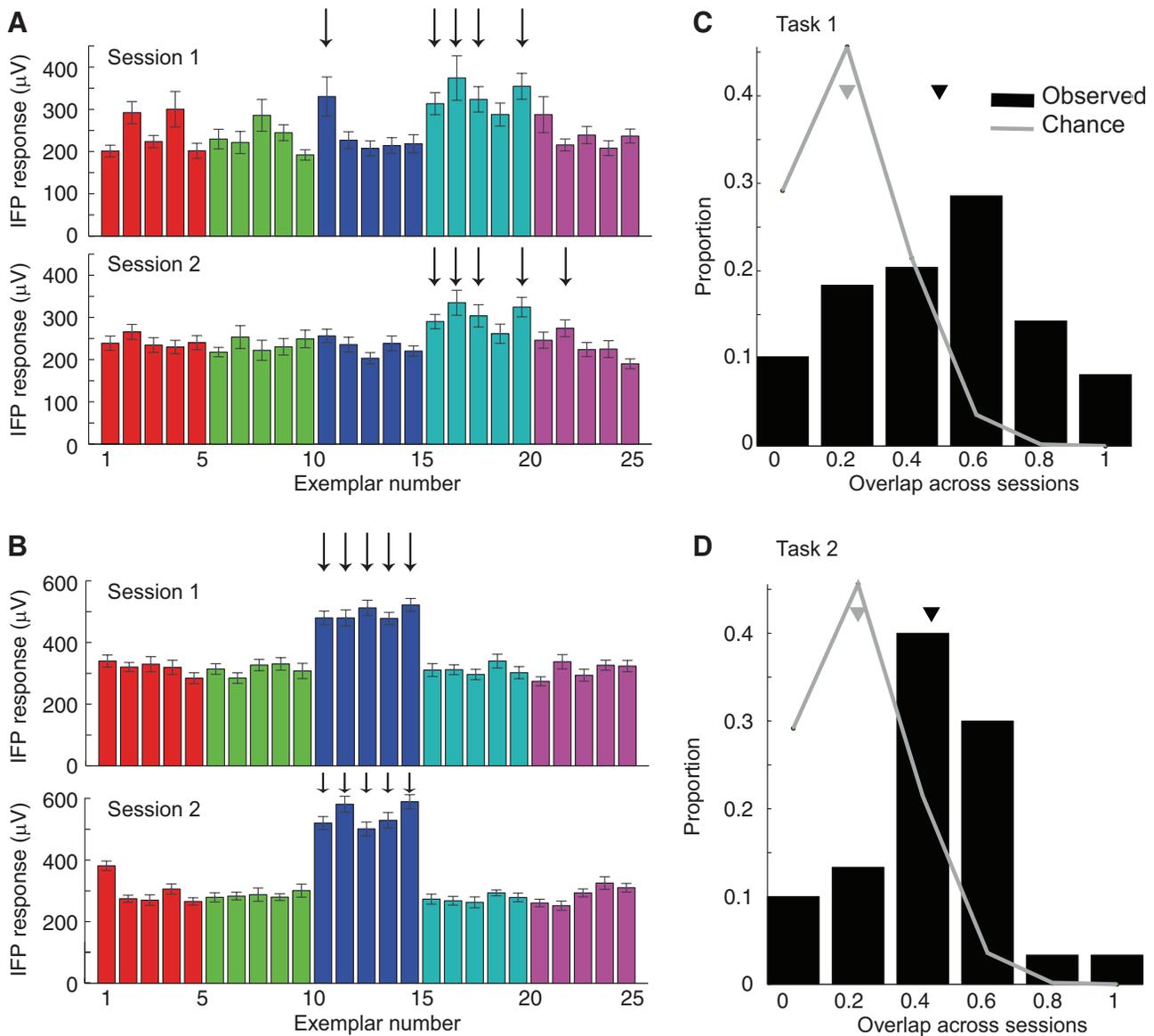


Fig. 3. Preferred exemplar objects for each electrode were consistent across sessions. *A* and *B*: IFP amplitude [max(IFP) – min(IFP) between 50 and 300 ms after stimulus onset, mean \pm SE] for the 2 electrodes in Fig. 2, *A* and *B*, in response to each of the 25 exemplars during *session 1* (top) or *session 2* (bottom). The color of each bar denotes the object category (shown only for comparison with Fig. 1; the category information is not used in the analyses depicted in this figure). Arrows mark the 5 exemplars that elicited the largest IFP amplitude in each session. The degree of overlap was 0.8 (*A*) and 1.0 (*B*). *C* and *D*: we computed the fraction of the top 5 exemplars that overlapped across 2 consecutive sessions (ranging from 0 to 1 in steps of 0.2). Plots show the overlap across sessions for *task 1* (*C*) and *task 2* (*D*). The mean overlap across sessions (black inverted triangle) was 0.49 (SD 0.29; *C*) and 0.43 (SD 0.23; *D*). The gray line shows the expected distribution of overlap values in a shuffle control where 5 exemplars were picked randomly in each session (gray inverted triangle indicates the mean for this shuffle condition). In both tasks, the observed distribution was significantly different from what would be expected by chance (chi-squared goodness-of-fit test, $P < 10^{-6}$). We did not perform this analysis for *task 3* because there were only 5 exemplar objects.

the distribution of decoding performance values obtained after shuffling the image labels (METHODS). Figure 6*A* shows the within-session decoding performance values using a set of robust features consisting of the principal components corresponding to the top eigenvalues that accounted for 95% of the variance. The within-session decoding performance values ranged from 0.44 to 0.91 and were similar in the three different tasks (compare the different colors in Fig. 6). In this analysis, the within-session decoding performance values were independently evaluated in each recording session. We observed that there was a strong correlation in the within-session decoding performance values between each recording session and the subsequent recording session (Fig. 6*A*, $R^2 = 0.55$).

The second best within-session decoding performance values were obtained using the total power in the gamma frequency band, between 70 and 100 Hz (Fig. 6*G*, $R^2 = 0.52$). Other frequency bands yielded weaker within-session decoding performance values, but there was still a strong correlation between sessions (Fig. 6, *B–F*; see Fig. 6 legend for R^2 and slope values). When using a multiclass classifier for decoding instead of a binary classifier, we also observed a strong correlation in the within-session decoding performance values between each recording session and its subsequent session (Fig. 6*H*, $R^2 = 0.61$). In summary, the within-session decoding performance values were consistent across consecutive recording sessions.

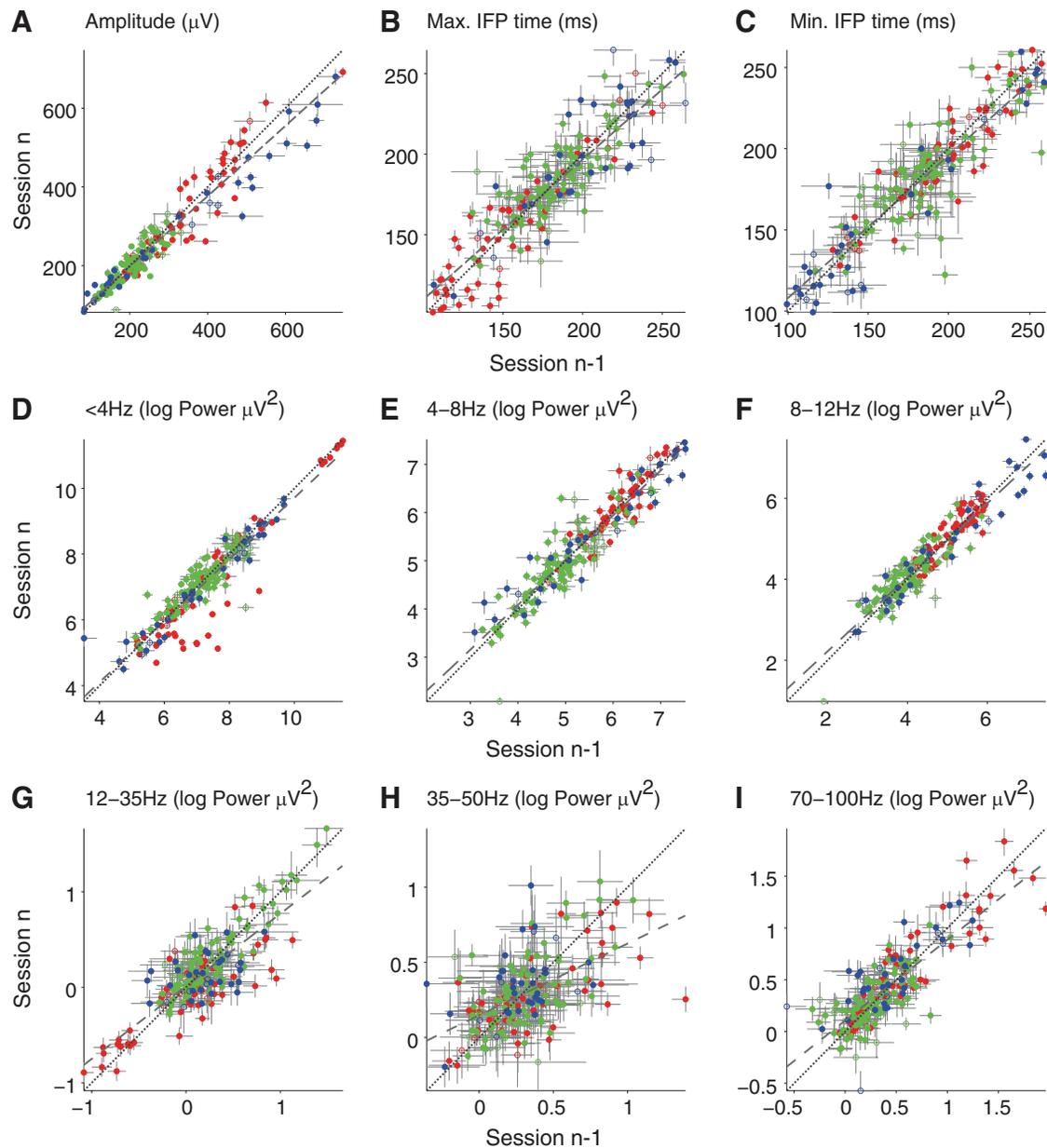


Fig. 4. IFP visual response features were stable over consecutive sessions. Plots show comparison of different response features extracted from the IFP response (for the [50,300 ms] window after stimulus onset) across 2 consecutive sessions ($n - 1$, n). When there were more than 2 sessions, each pair of consecutive sessions contributes 1 point in these plots. All electrodes that showed visual selectivity during at least 1 session were included in this plot [see METHODS; *task 1*: $n = 55$ comparisons, 33 electrodes (red); *task 2*: $n = 36$ comparisons, 18 electrodes (blue); *task 3*: $n = 102$ comparisons, 40 electrodes (green)]. The different features considered were IFP response amplitude [A; $R^2 = 0.93$, slope = 0.85, root mean square error (RMSE) = 48], time of the maximum IFP response (B; $R^2 = 0.76$, slope = 0.84, RMSE = 18), time of the minimum IFP response (C; $R^2 = 0.79$, slope = 0.86, RMSE = 17), and log power in 6 nonoverlapping frequency bands: <4 Hz (D; $R^2 = 0.85$, slope = 0.93, RMSE = 0.53), 4–8 Hz (E; $R^2 = 0.89$, slope = 0.92, RMSE = 0.35), 8–12 Hz (F; $R^2 = 0.9$, slope = 0.9, RMSE = 0.32), 12–35 Hz (G; $R^2 = 0.69$, slope = 0.8, RMSE = 0.25), 35–50 Hz (H; $R^2 = 0.44$, slope = 0.53, RMSE = 0.25), and 70–100 Hz (I; $R^2 = 0.7$, slope = 0.79, RMSE = 0.22). Gray lines indicate \pm SE along each axis. Dashed lines plot the best-fit linear regression, and dotted lines show the diagonal (perfect stability). Based on the criteria described in METHODS, each point reflecting the comparison across consecutive sessions was labeled stationary (filled circles; $n = 166$ comparisons, 81 electrodes) or nonstationary (open circles; $n = 27$ comparisons, 10 electrodes).

In the analyses in Fig. 6, the classifier was trained independently in each recording session. To begin to evaluate whether target brain regions could decode information without changing the synaptic weights in each recording session, we compared the within-session decoding performance with across-session decoding performance values (Fig. 7). In the across-session analyses, the classifier was trained with data from the first recording session and the decoding performance was evaluated by using the data from the second recording session.

Because classifier performance is often very sensitive to the number of training examples, we randomly subsampled the data to ensure that the number of training examples was the same for within-session and across-session analyses (on average, 83.2 ± 57.8 trials per electrode per category per session for training and testing were used). As expected, the within-session decoding performance was consistently above the across-session performance (compare circles vs. triangles in Fig. 7). Yet, in most cases, we obtained a significant across-session decoding

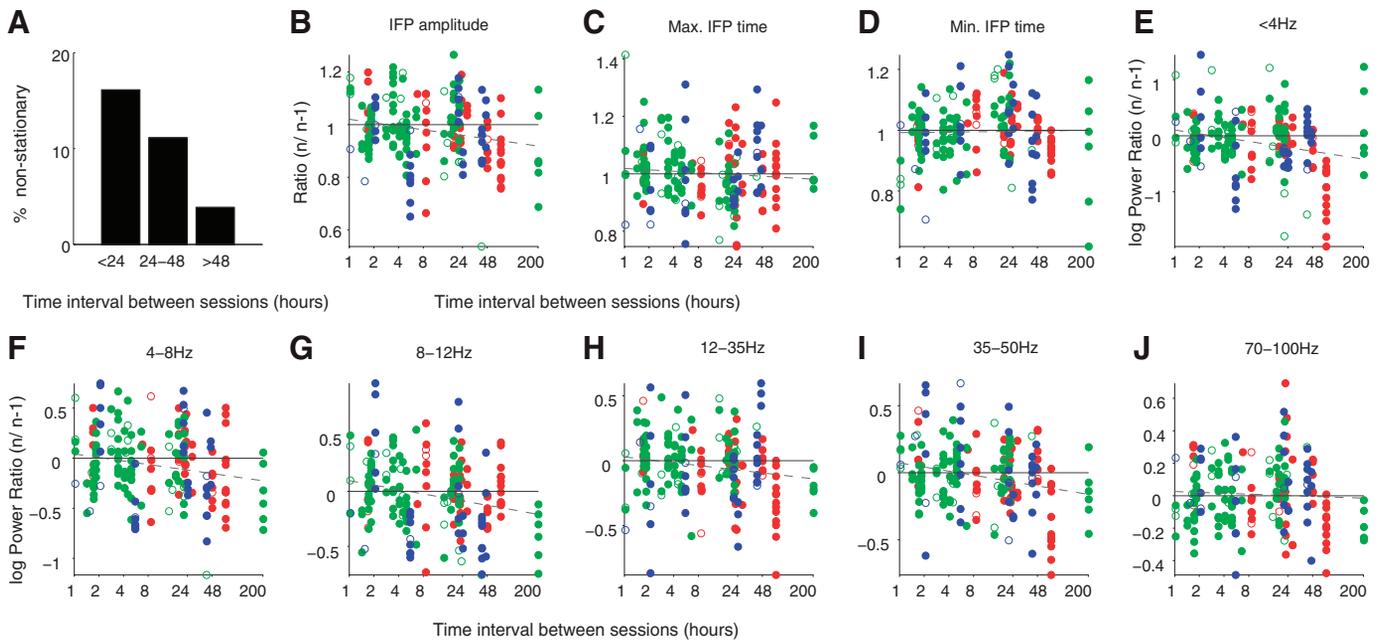


Fig. 5. IFP responses were stable over several days. *A*: fraction of session pairs that were nonstationary as a function of time between sessions. Because of the small number of nonstationary electrodes, we grouped the *x*-axis into 3 possible bins (the number of electrodes in each bin was 9, 2, and 1, respectively). The fraction in the *y*-axis was computed with respect to the total number of session pairs that showed selective responses (149, 18, and 26, respectively). *B–J*: for each of the IFP response features illustrated in Fig. 4, we computed the ratio (or log ratio in *E–J*) between a session and the preceding session. This ratio is plotted as a function of the time between sessions (in log scale). Color and other conventions are as in Fig. 4. Dashed lines show the linear fit for all points not considered outliers (outliers were points outside 4 SD of the mean ratio; the number of outliers was between 0 and 2 for each subplot). The slope, intercept, RMSE, and number of outliers (where non-zero) for each subplot were as follows: *B*, slope = $-0.044/h$ ($P < 0.01$ vs. $0/h$), intercept = 1.02, RMSE = 0.12; *C*, slope = $-0.016/h$, intercept = 1.02, RMSE = 0.10; *D*, slope = $0.003/h$, intercept = 0.99, RMSE = 0.10; *E*, slope = $-0.226/h$ ($P < 0.01$ vs. $0/h$), intercept = 0.10, RMSE = 0.53; *F*, slope = $-0.116/h$ ($P < 0.01$ vs. $0/h$), intercept = 0.04, RMSE = 0.35; *G*, slope = $-0.133/h$ ($P < 0.01$ vs. $0/h$), intercept = 0.10, RMSE = 0.32; *H*, slope = $-0.073/h$, intercept = 0.03, RMSE = 0.25; *I*, slope = $-0.102/h$ ($P < 0.01$ vs. $0/h$), intercept = 0.07, RMSE = 0.24, outliers = 1; *J*, slope = $-0.018/h$, intercept = 0.03, RMSE = 0.19, outliers = 2. Solid lines show the expected ratio assuming perfect stability.

performance, demonstrating that it is possible to read out information without retraining the classifier. Furthermore, there was a strong correlation between the within-session decoding performance and the across-session decoding performance for all three tasks (Fig. 7, solid black lines).

DISCUSSION

In this study we characterized the temporal stability of visually selective intracranial field potentials elicited by presentation of flashes of visual stimuli. We considered timescales of minutes (short-term within-session comparisons) to hours to days. Our study was based on a large data set of 2,195 electrodes recording IFP activity from 27 subjects performing 3 different visual recognition tasks. We focused on those electrodes that revealed strong visually selective responses, because this constitutes the most rigorous assessment of stability. Overall, we found a strong degree of stability in the visually elicited responses. This stability was manifested in multiple different analyses including consideration of the object preferences (Figs. 1–3), examination of various features of the IFP waveform (Figs. 4 and 5), and decoding performance in single trials (Figs. 6 and 7).

Field potential recordings can reflect contributions from volume conduction (Kajikawa and Schroeder 2011). A recent study argued that volume conduction as measured using IFP electrodes is insignificant at distances >20 mm (Vidal et al. 2012). To minimize potential effects of volume conduction, we subtracted the mean signals across all electrodes. In Fig. 1, *insets*, we showed that the activity near the selective electrodes

(within ~ 1 – 2 cm) was clearly distinct, suggesting that the selective responses that we considered here do not spread over more than ~ 1 – 2 cm of cortex. The notion that the IFP signals represent approximately local activity is also consistent with receptive field mappings in early visual cortex with this same type of electrode (Yoshor et al. 2007).

It should be noted that the biophysical origin of the IFP signals studied here is not well understood. IFPs may comprise a weighted average of excitatory and inhibitory postsynaptic potentials across a large number of neurons (Buzsaki et al. 2012; Helmchen et al. 1999; Mitzdorf 1987). The spatial scale of IFP signals is larger than that of local field potentials (LFP) measured with intracortical microwires, due to the 2.3-mm vs. ~ 50 - μ m electrode diameter and the low impedance of the IFP electrodes. Several studies have shown that there is a significant correlation between LFPs recorded from high-impedance microwires and spiking activity in sensory and motor cortices (Bansal et al. 2011, 2012; Belitski et al. 2008; Burns et al. 2010; Katzner et al. 2009; Liu and Newsome 2006; Nir et al. 2007; Rasch et al. 2008; Rasch et al. 2009; Ray et al. 2008; Xing et al. 2009), but the two signals are clearly distinct and LFPs do not directly reflect firing rates (Belitski et al. 2008; Ekstrom et al. 2007; Kajikawa and Schroeder 2011; Kreiman et al. 2006; Logothetis 2003; Nielsen et al. 2006; Norena and Eggermont 2002). It has been suggested that LFPs represent a combination of local dendritic processing and afferent input as opposed to the output spike rate (Logothetis 2003). Relatively little is known about the relationship between IFPs and spiking activity.

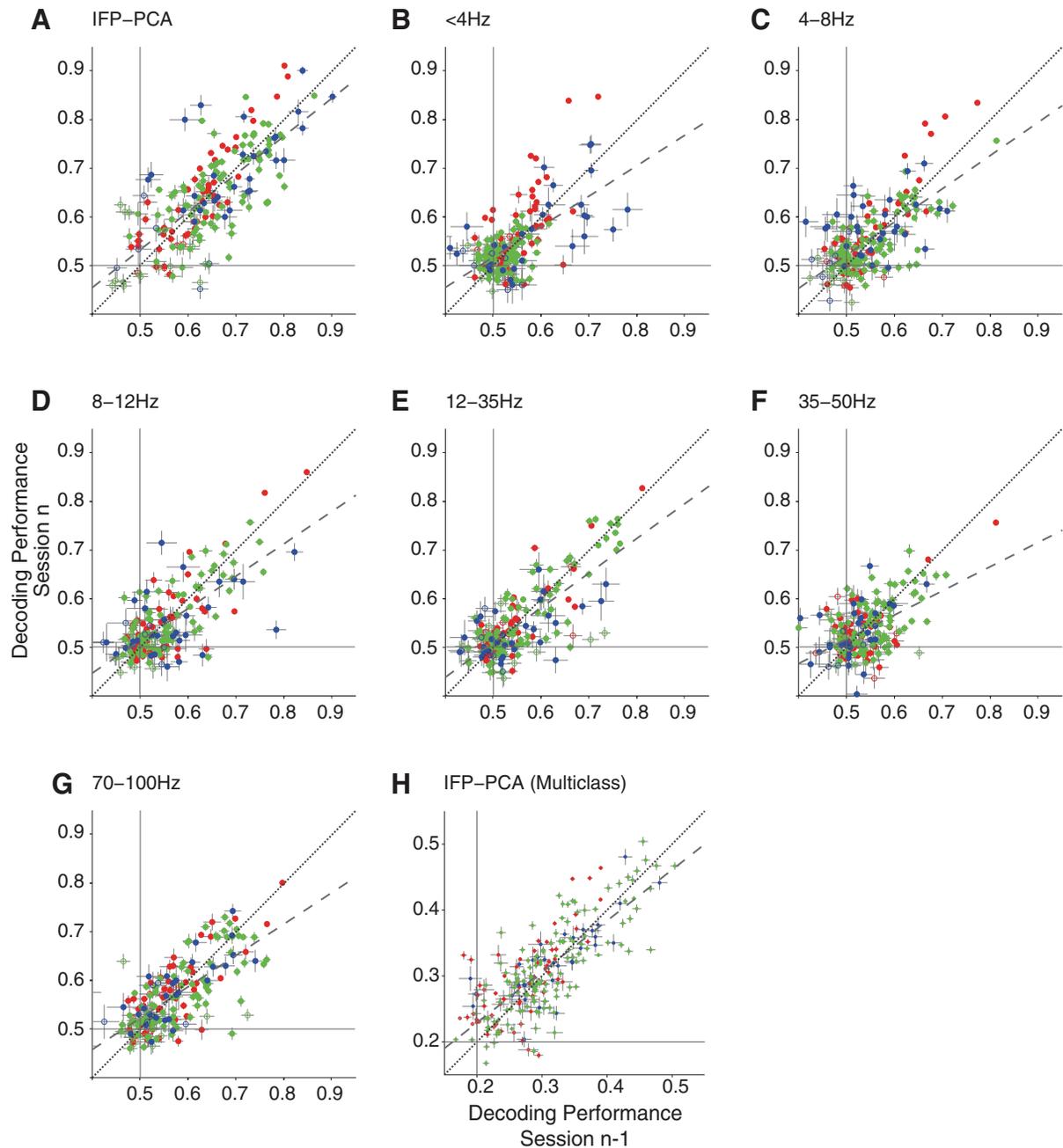


Fig. 6. Decoding performance in single trials was stable over consecutive sessions. We used a linear classifier to decode the preferred stimulus for each electrode in single trials (see METHODS). The decoding performance values reported correspond to using cross-validation test data. In A–G, we used a binary classifier: a decoding performance value of 0.5 indicates chance level (thin horizontal and vertical gray lines), and a decoding performance of 1 indicates perfect classification in each trial. In H, we used a multiclass classifier: a decoding performance value of 0.2 indicates chance level (thin horizontal and vertical gray lines). Here we compare the decoding performance in a given session (*x*-axis) with the decoding performance on the next session (*y*-axis). Classifiers were trained independently in each session (in contrast to Fig. 7). Different features of the IFP response were used to train the classifier (see METHODS): the top principal components accounting for 95% of the variance of the IFP response (A; $R^2 = 0.55$, slope = 0.78, RMSE = 0.069) and log power in 6 nonoverlapping frequency bands: <4 Hz (B; $R^2 = 0.35$, slope = 0.63, RMSE = 0.057), 4–8 Hz (C; $R^2 = 0.46$, slope = 0.68, RMSE = 0.055), 8–12 Hz (D; $R^2 = 0.50$, slope = 0.67, RMSE = 0.056), 12–35 Hz (E; $R^2 = 0.59$, slope = 0.72, RMSE = 0.053), 35–50 Hz (F; $R^2 = 0.28$, slope = 0.50, RMSE = 0.053), and 70–100 Hz (G; $R^2 = 0.52$, slope = 0.65, RMSE = 0.056). In H, we considered a multiclass classifier using the top principal components accounting for 95% of the variance of the IFP response ($R^2 = 0.61$, slope = 0.77, RMSE = 0.045). Gray bars indicate \pm SE along each axis. Dashed lines plot the best linear fit, and the dotted line shows the diagonal. Other colors, numbers, and conventions are as in Fig. 4. PCA, principal components analysis.

Given our incomplete knowledge about the relationship between neuronal responses and IFPs, caution should be taken in the interpretation of our results in terms of the temporal stability of the underlying neuronal signals. IFP signals could remain stable even if there are significant changes at the

neuronal level. For example, half the cells could increase their firing rates, and the other half could decrease their firing rates (e.g., Woloszyn and Sheinberg 2012) in such a way that might not change the net IFP signal. Most single-neuron studies have focused on neuronal preferences over short time scales of

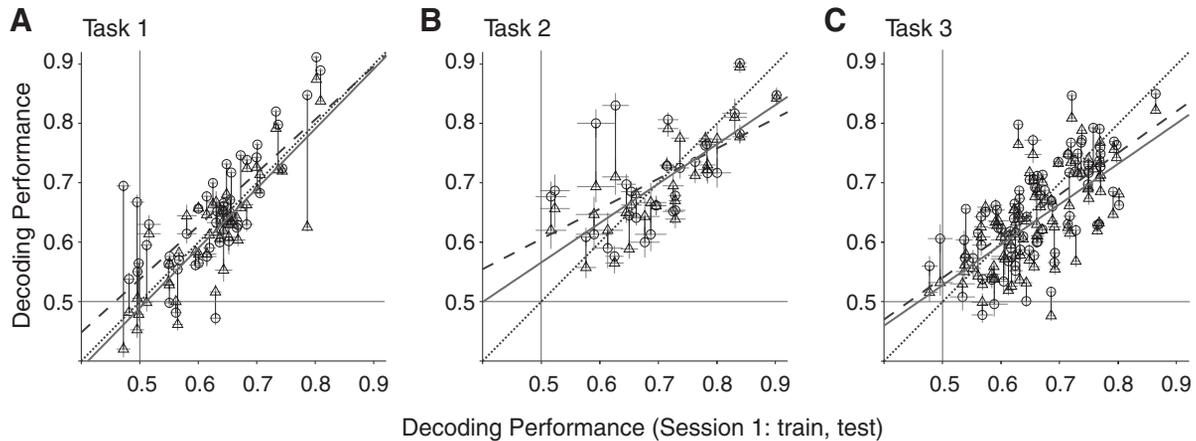


Fig. 7. A machine learning classifier can extrapolate across sessions. Plots show comparisons between across-session decoding performance and within-session decoding performance for stationary electrodes. The *x*-axis plots the decoding performance when data from *session 1* are used for training and testing (using different trials for cross-validation; see METHODS). On the *y*-axis, circles show the decoding performance when data from within *session 2* are used for training and testing (cross-validated, within-session decoding). Also on the *y*-axis, triangles show the decoding performance when data from *session 1* are used for training and data from *session 2* are used for testing (across-session decoding). The number of trials used for training was the same for the connected circular and triangular markers. Each subplot shows the results using data from a different task (*A: task 1*, $n = 49$ comparisons, 30 electrodes; *B: task 2*, $n = 30$ comparisons, 15 electrodes; *C: task 3*, $n = 87$ comparisons, 36 electrodes). Only electrodes labeled stationary are plotted. The dotted line shows the diagonal. Thin gray lines denote chance performance (0.50). The dashed line shows the linear fit for the circles (within-*session 2* decoding performance vs. within-*session 1* decoding performance), and the solid line shows the linear fit for the triangles (across-session decoding performance vs. within-*session 1* decoding performance).

minutes to a few hours partly due to technical limitations (e.g., neuronal recordings may become unstable, or it may not be trivial to ensure that the same neuron is being monitored over long periods of time). Several studies have examined the properties of neuronal populations over longer time scales but recording from different neurons each day (e.g., Freedman et al. 2005; Woloszyn and Sheinberg 2012). Recently, a few studies have examined the degree of stability of neuronal recordings and/or neuronal preferences over several days (Bondar et al. 2009; Nicolelis et al. 2003; Tolias et al. 2007). Some of these studies point to a remarkable degree of stability in feature preferences at the single-unit or multi-unit level. Our study expands on these observations by providing evidence of temporal stability at temporal scales of hours to days and at a spatial “mesoscale” of several millimeters of cortex.

The degree of stationarity in the responses was approximately similar across the multiple frequency bands of the IFP signals evaluated here (Figs. 4 and 6). The degree of stationarity was weakest for the 35- to 50-Hz band (smallest slope in Figs. 4 and 6). This result is consistent with reports from Belitski et al. (2008) describing that in LFPs recorded from primary visual cortex, visual information was represented mostly in the low and high frequencies, whereas the middle frequency bands were related to neuromodulatory inputs.

A small fraction of our recordings (<11%) did not fulfill our stability criteria and were labeled “nonstationary” (Table 2). Nonstationarity in IFP signals over scales of hours to days could be due to several factors, including small electrode shifts, changes in the underlying neural tissue, interictal discharges, medication, global state changes such as varying levels of attention or motivation, and also plasticity in the neural responses. Some chronic single-unit studies have distinguished between the stability of the equipment and recordings and the stability of the underlying neuronal representation (Bondar et al. 2009; Nicolelis et al. 2003; Tolias et al. 2007). Given our methodology and the observation that there were very few nonstationary electrodes, it was difficult to distinguish among

these different putative changes over time. To a first approximation, the degree of temporal stability did not depend on the task (compare different colors in Figs. 4–7) or the electrode location (Table 3). The types of visual recognition tasks that we considered here did not require any type of explicit visual learning across sessions. The quantitative data presented here provide a baseline for subsequent studies to examine long-term plasticity, long-term adaptation, novelty detection, and other forms of changes over time.

It is tempting to speculate that some brain regions may show more stability, whereas other brain regions may be more plastic or variable (Maimon and Assad 2009; Mandelblat-Cerf et al. 2009; Shinomoto et al. 2009). For example, responses at the level of the retina could show long-term stability, whereas higher areas of the temporal lobe such as anterior inferior temporal cortex and perirhinal cortex could be more plastic (Buffalo et al. 1998; Kourtzi and DiCarlo 2006; Tolias et al. 2007). Within our recordings, we did not find any clear differences in long-term temporal stability among the different areas of the ventral visual cortex to which we had access (Table 3). However, it should be noted that the tasks examined here did not require any long-term learning. Furthermore, the number of electrodes as well as sampling of different parts of the ventral visual cortex in our study is limited.

We decoded signals from single-trial responses (Figs. 6 and 7), but it should be noted that the brain does not directly interpret IFPs as measured in our experiments for computation. The decoding performance values reported here quantify the information contained in single trials in the IFPs. Furthermore, the stability in decoding performance is of interest regarding the possibility of building brain-machine interfaces for prosthetic applications (Hochberg et al. 2006; Musallam et al. 2004; Shenoy et al. 2003; Waldert et al. 2009; Winter et al. 2007). Typically, a machine learning algorithm is trained to read out the activity of ensembles of neurons or electrodes recording field potential signals. For these applications, it is important to establish if and when this type of device needs to

be retrained (Ganguly and Carmena 2009; Simeral et al. 2011). The decoding performance analyses presented here quantify how well such a device could extrapolate across sessions spanning hours to days.

The brain needs to maintain information over long time scales while at the same time adapting, learning, and acquiring novel skills and information. The stationary neural signals that we describe could reflect the stability of recognition performance for familiar stimuli over long temporal scales. How these signals relate to the stability of information at longer time scales (e.g., months to years), at different spatial scales (e.g., neurons or brain regions), or to learning and perceptual changes remains an important theme for future studies.

ACKNOWLEDGMENTS

We thank Julie Blumberg for helping with the initial analyses of the data. The data from *tasks 1* and *2* were collected by Yigal Agam and Hesheng Liu. The data from *task 3* were collected by Jedediah Singer. The analyses for this article were run on the Orchestra cluster supported by the Harvard Medical School Research Information Technology Group.

GRANTS

This work was supported by grants from the National Science Foundation (G. Kreiman) and the National Institutes of Health (G. Kreiman).

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the authors.

AUTHOR CONTRIBUTIONS

A.K.B., J.M.S., and G.K. conception and design of research; A.K.B., J.M.S., and G.K. analyzed data; A.K.B., J.M.S., and G.K. interpreted results of experiments; A.K.B., J.M.S., and G.K. prepared figures; A.K.B., J.M.S., and G.K. drafted manuscript; A.K.B., J.M.S., and G.K. edited and revised manuscript; A.K.B., J.M.S., W.S.A., A.G., J.R.M., and G.K. approved final version of manuscript; J.M.S., W.S.A., A.G., and J.R.M. performed experiments.

REFERENCES

- Agam Y, Liu H, Pappanastassiou A, Buia C, Golby AJ, Madsen JR, Kreiman G. Robust selectivity to two-object images in human visual cortex. *Curr Biol* 20: 872–879, 2010.
- Bansal AK, Truccolo W, Vargas-Irwin CE, Donoghue JP. Decoding 3D reach and grasp from hybrid signals in motor and premotor cortices: spikes, multiunit activity, and local field potentials. *J Neurophysiol* 107: 1337–1355, 2012.
- Bansal AK, Vargas-Irwin CE, Truccolo W, Donoghue JP. Relationships among low-frequency local field potentials, spiking activity, and three-dimensional reach and grasp kinematics in primary motor and ventral premotor cortices. *J Neurophysiol* 105: 1603–1619, 2011.
- Belitski A, Gretton A, Magri C, Murayama Y, Montemurro MA, Logothetis NK, Panzeri S. Low-frequency local field potentials and spikes in primary visual cortex convey independent visual information. *J Neurosci* 28: 5696–5709, 2008.
- Bondar I, Leopold D, Richmond B, Victor J, Logothetis N. Long-term stability of visual pattern selective responses of monkey temporal lobe neurons. *PLoS One* 4: e8222, 2009.
- Buffalo E, Reber P, Squire L. The human perirhinal cortex and recognition memory. *Hippocampus* 8: 330–339, 1998.
- Burns SP, Xing D, Shapley RM. Comparisons of the dynamics of local field potential and multiunit activity signals in macaque visual cortex. *J Neurosci* 30: 13739–13749, 2010.
- Buzsaki G, Anastassiou CA, Koch C. The origin of extracellular fields and currents—EEG, ECoG, LFP and spikes. *Nat Rev Neurosci* 13: 407–420, 2012.
- Chao Z, Nagasaka Y, Fujii N. Long-term asynchronous decoding of arm motion using electrocorticographic signals in monkeys. *Front Neuroeng* 3: 3, 2010.
- Dale AM, Fischl B, Sereno MI. Cortical surface-based analysis. I. Segmentation and surface reconstruction. *Neuroimage* 9: 179–194, 1999.
- Destrieux C, Fischl B, Dale A, Halgren E. Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature. *Neuroimage* 53: 1–15, 2010.
- Ekstrom A, Viskontas I, Kahana M, Jacobs J, Upchurch K, Bookheimer S, Fried I. Contrasting roles of neural firing rate and local field potentials in human memory. *Hippocampus* 17: 606–617, 2007.
- Engel AK, Moll CK, Fried I, Ojemann GA. Invasive recordings from the human brain: clinical insights and beyond. *Nat Rev Neurosci* 6: 35–47, 2005.
- Freedman DJ, Riesenhuber M, Poggio T, Miller EK. Experience-dependent sharpening of visual shape selectivity in inferior temporal cortex. *Cereb Cortex* 16: 1631–1644, 2006.
- Ganguly K, Carmena JM. Emergence of a stable cortical map for neuroprosthetic control. *PLoS Biol* 7: e1000153, 2009.
- Helmchen F, Svoboda K, Denk W, Tank DW. In vivo dendritic calcium dynamics in deep-layer cortical pyramidal neurons. *Nat Neurosci* 2: 989–996, 1999.
- Hochberg LR, Serruya MD, Friehs GM, Mukand JA, Saleh M, Caplan AH, Branner A, Chen D, Penn RD, Donoghue JP. Neuronal ensemble control of prosthetic devices by a human with tetraplegia. *Nature* 442: 164–171, 2006.
- Hung C, Kreiman G, Poggio T, DiCarlo J. Fast read-out of object identity from macaque inferior temporal cortex. *Science* 310: 863–866, 2005.
- Kajikawa Y, Schroeder CE. How local is the local field potential? *Neuron* 72: 847–858, 2011.
- Katzner S, Nauhaus I, Benucci A, Bonin V, Ringach DL, Carandini M. Local origin of field potentials in visual cortex. *Neuron* 61: 35–41, 2009.
- Kourtzi Z, DiCarlo JJ. Learning and neural plasticity in visual object recognition. *Curr Opin Neurobiol* 16: 152–158, 2006.
- Kreiman G. Single neuron approaches to human vision and memories. *Curr Opin Neurobiol* 17: 471–475, 2007.
- Kreiman G, Hung C, Quiñero R, Kraskov A, Poggio T, DiCarlo J. Object selectivity of local field potentials and spikes in the inferior temporal cortex of macaque monkeys. *Neuron* 49: 433–445, 2006.
- Kriegeskorte N, Kreiman G. *Visual Population Codes*. Cambridge, MA: MIT Press, 2011.
- Liu H, Agam Y, Madsen JR, Kreiman G. Timing, timing, timing: fast decoding of object information from intracranial field potentials in human visual cortex. *Neuron* 62: 281–290, 2009.
- Liu J, Newsome WT. Local field potential in cortical area MT: stimulus tuning and behavioral correlations. *J Neurosci* 26: 7779–7790, 2006.
- Logothetis NK. The underpinnings of the BOLD functional magnetic resonance imaging signal. *J Neurosci* 23: 3963–3971, 2003.
- Maimon G, Assad JA. Beyond Poisson: increased spike-time regularity across primate parietal cortex. *Neuron* 62: 426–440, 2009.
- Mandelblat-Cerf Y, Paz R, Vaadia E. Trial-to-trial variability of single cells in motor cortices is dynamically modified during visuomotor adaptation. *J Neurosci* 29: 15053–15062, 2009.
- Mitzdorf U. Properties of the evoked-potential generators: current source-density analysis of visually evoked-potentials in the cat cortex. *Int J Neurosci* 33: 33–59, 1987.
- Musallam S, Corneil B, Greger B, Scherberger H, Andersen R. Cognitive control signals for neural prosthetics. *Science* 305: 258–261, 2004.
- Nicoletis M, Dimitrov D, Carmena J, Crist R, Lehew G, Kralik J, Wise S. Chronic, multisite, multielectrode recordings in macaque monkeys. *Proc Natl Acad Sci USA* 100: 11041–11046, 2003.
- Nielsen K, Logothetis N, Rainer G. Dissociation between LFP and spiking activity in macaque inferior temporal cortex reveals diagnostic parts-based encoding of complex objects. *J Neurosci* 26: 9639–9645, 2006.
- Nir Y, Fisch L, Mukamel R, Gelbard-Sagiv H, Arieli A, Fried I, Malach R. Coupling between neuronal firing rate, gamma LFP, and BOLD fMRI is related to inter-neuronal correlations. *Curr Biol* 17: 1275–1285, 2007.
- Norena A, Eggermont J. Comparison between local field potentials and unit cluster activity in primary auditory cortex and anterior auditory field in the cat. *Hear Res* 166: 202–213, 2002.
- Porada I, Bondar I, Spatz WB, Kruger J. Rabbit and monkey visual cortex: more than a year of recording with up to 64 microelectrodes. *J Neurosci Methods* 95: 13–28, 2000.

- Rasch M, Gretton A, Murayama Y, Maass W, Logothetis NK. Inferring spike trains from local field potentials. *J Neurophysiol* 99: 1461–1476, 2008.
- Rasch M, Logothetis NK, Kreiman G. From neurons to circuits: linear estimation of local field potentials. *J Neurosci* 29: 13785–13796, 2009.
- Ray S, Crone NE, Niebur E, Franaszczuk PJ, Hsiao SS. Neural correlates of high-gamma oscillations (60–200 Hz) in macaque local field potentials and their potential implications in electrocorticography. *J Neurosci* 28: 11526–11536, 2008.
- Shenoy KV, Meeker D, Cao S, Kureshi SA, Pesaran B, Buneo CA, Batista AP, Mitra PP, Burdick JW, Andersen RA. Neural prosthetic control signals from plan activity. *Neuroreport* 14: 591–596, 2003.
- Shinomoto S, Kim H, Shimokawa T, Matsuno N, Funahashi S, Shima K, Fujita I, Tamura H, Doi T, Kawano K, Inaba N, Fukushima K, Kurkin S, Kurata K, Taira M, Tsutsui K, Komatsu H, Ogawa T, Koida K, Tanji J, Toyama K. Relating neuronal firing patterns to functional differentiation of cerebral cortex. *PLoS Comput Biol* 5: e1000433, 2009.
- Simeral JD, Kim SP, Black MJ, Donoghue JP, Hochberg LR. Neural control of cursor trajectory and click by a human with tetraplegia 1000 days after implant of an intracortical microelectrode array. *J Neural Eng* 8: 025027, 2011.
- Suner S, Fellows M, Vargas-Irwin C, Nakata G, Donoghue J. Reliability of signals from a chronically implanted, silicon-based electrode array in non-human primate primary motor cortex. *IEEE Trans Neural Syst Rehabil Eng* 13: 524–541, 2005.
- Tolias A, Ecker A, Siapas A, Hoenselaar A, Keliris G, Logothetis N. Recording chronically from the same neurons in awake, behaving primates. *J Neurophysiol* 98: 3780–3790, 2007.
- Vidal JR, Freyermuth S, Jerbi K, Hamame CM, Ossandon T, Bertrand O, Minotti L, Kahane P, Berthoz A, Lachaux JP. Long-distance amplitude correlations in the high gamma band reveal segregation and integration within the reading network. *J Neurosci* 32: 6421–6434, 2012.
- Waldert S, Pistohl T, Braun C, Ball T, Aertsen A, Mehring C. A review on directional information in neural signals for brain-machine interfaces. *J Physiol (Paris)* 103: 244–254, 2009.
- Winter JO, Cogan SF, Rizzo JF 3rd. Retinal prostheses: current challenges and future outlook. *J Biomater Sci Polym Ed* 18: 1031–1055, 2007.
- Woloszyn L, Sheinberg DL. Effects of long-term visual experience on responses of distinct classes of single units in inferior temporal cortex. *Neuron* 74: 193–205, 2012.
- Xing D, Yeh CI, Shapley RM. Spatial spread of the local field potential and its laminar variation in visual cortex. *J Neurosci* 29: 11540–11549, 2009.
- Yoshor D, Bosking WH, Ghose GM, Maunsell JH. Receptive fields in human visual cortex mapped with surface electrodes. *Cereb Cortex* 17: 2293–2302, 2007.

