

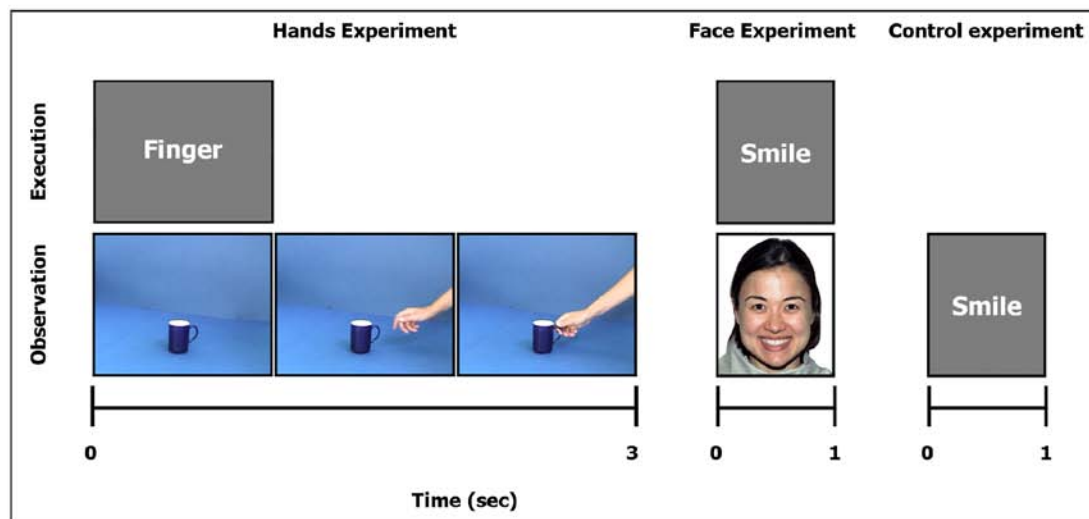
Supplementary online material for

‘Single neuron responses in humans during execution and observation of actions’

Mukamel, Ekstrom, Kaplan, Iacoboni, and Fried

Figure S1

A



B

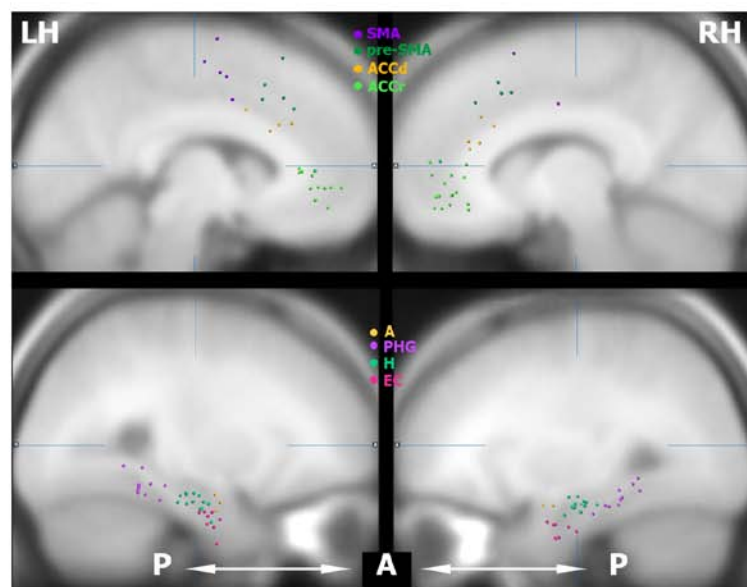


Figure S1 (related to Figure 1): Experimental design. **A)** The experiment was composed of three parts – *Grasp*, *Facial expressions* and *Control*. During *Grasp*, subjects were presented with video clips of a hand grasping a mug and with the words ‘Finger’ or ‘Hand’. They were instructed to grasp a mug with precision grip or whole hand prehension when the words ‘Finger’ or ‘Hand’, respectively, were presented and to simply observe when the video clips were played. During *Facial expressions*, subjects were presented with a picture of a smiling or a frowning face

and with the word ‘Smile’ or ‘Frown’. They were instructed to perform the corresponding action when the words were presented and to simply observe when the pictures were presented. Subjects were also instructed to refrain from making any hand movements or facial gestures during all observation conditions. One experimenter always supervised subject’s compliance during the tasks. In the *Control* task, subjects were presented with the words used as cues in the *Grasp* and *Facial expression* parts of the experiment and were instructed to covertly read the words and refrain from making hand movements or facial gestures. **B)** Anatomical location of electrodes in all 21 patients. Electrode location in each patient was verified by co-registering the post-operative CT scan with the pre-operative structural MRI. Electrode positions were transformed into MNI space and are presented on the MNI 305 brain. Top row shows the electrodes in the medial frontal lobe and bottom row displays the electrodes in the medial temporal lobe. LH – left hemisphere, RH – right hemisphere; A – anterior, P – posterior; SMA – supplementary motor area; ACCd – dorsal aspect of anterior cingulate cortex, ACCr – rostral aspect of anterior cingulate; A – amygdala, PHG – parahippocampal gyrus, H – hippocampus, EC – entorhinal cortex.

Figure S2

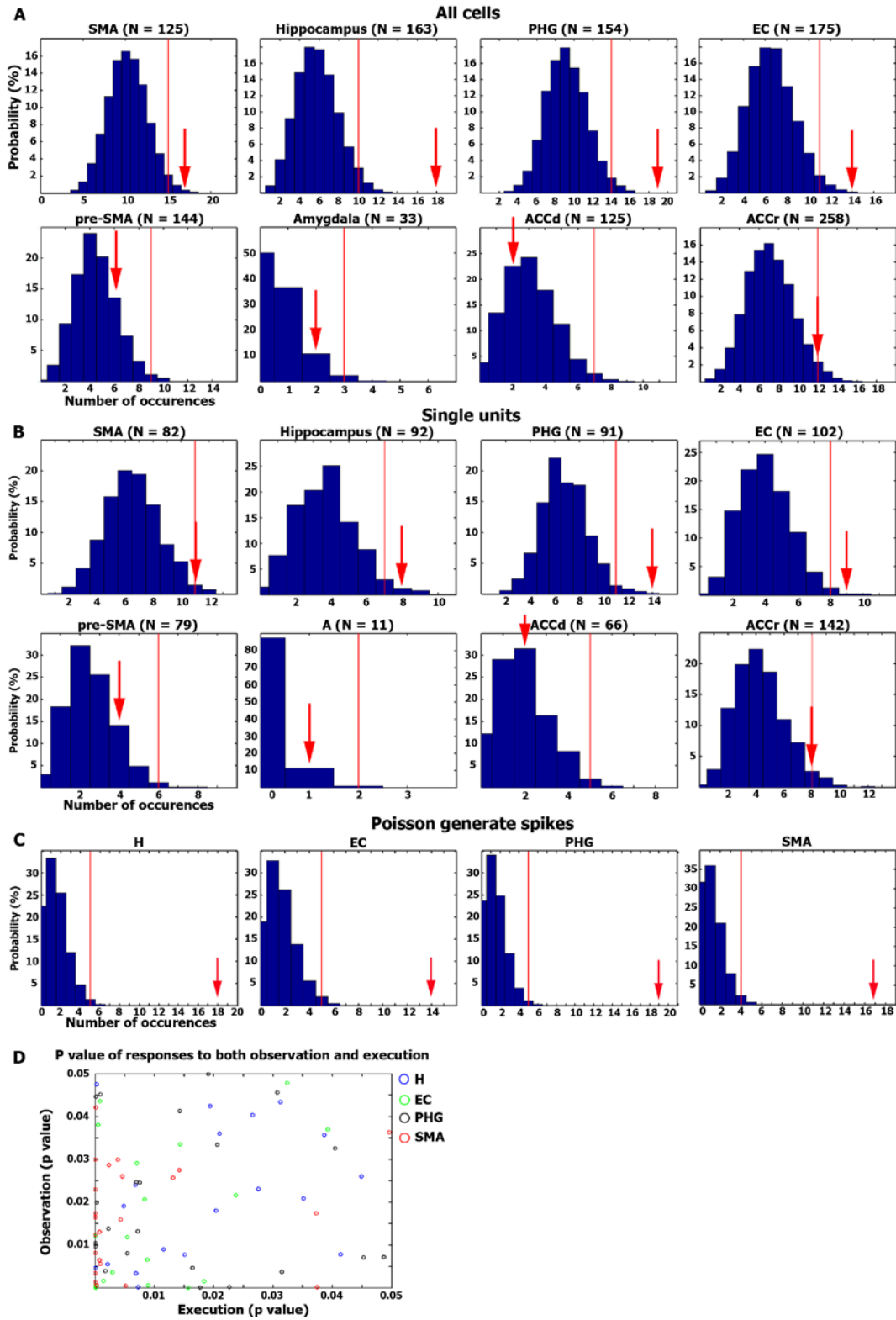


Figure S2 (related to Figure 2): **A)** Bootstrap analysis. In order to assess if the proportion of Action observation/execution matching neurons is significant or not, we compared the actual number of Action observation/execution matching neurons in each region (red arrow) with the null distribution computed over 10,000 iterations (blue bars; see methods). The vertical red line represents the 5% chance level. Note that the number of cells in SMA, hippocampus, parahippocampal gyrus, and entorhinal cortex was significantly higher than expected by chance. **B)** Same analysis as described for panel A but only using data recorded from single units (as opposed to single and multi units used in panel A). Again, the number of Action observation/execution matching cells in SMA, H, PHG and EC was higher than in the shuffled data at significance level of $p < 0.05$. **C)** Number of Action observation/execution matching neurons compared with Poisson generated spike trains with similar firing rates. For each recorded neuron, we calculated the average firing rate and generated surrogate spike trains with Poisson distributed inter-spike intervals and similar firing rate. Next, we assessed whether the neuron with the surrogate spike trains would be considered an Action observation/Execution matching cell. This was performed for each neuron in the population and the number of pseudo action observation/execution matching cells was counted. The blue columns show the distribution of number of action observation/execution matching neurons in the surrogate data after 10,000 iterations. The red arrow points to the actual number of action observation/execution matching cells in the real data. The red vertical line represents 1% chance level. **D)** P-value of response during action-execution (x axis) and action-observation (y-axis) for all action observation/execution matching cells. Acronyms for anatomic regions as in Figure S1.

Figure S3

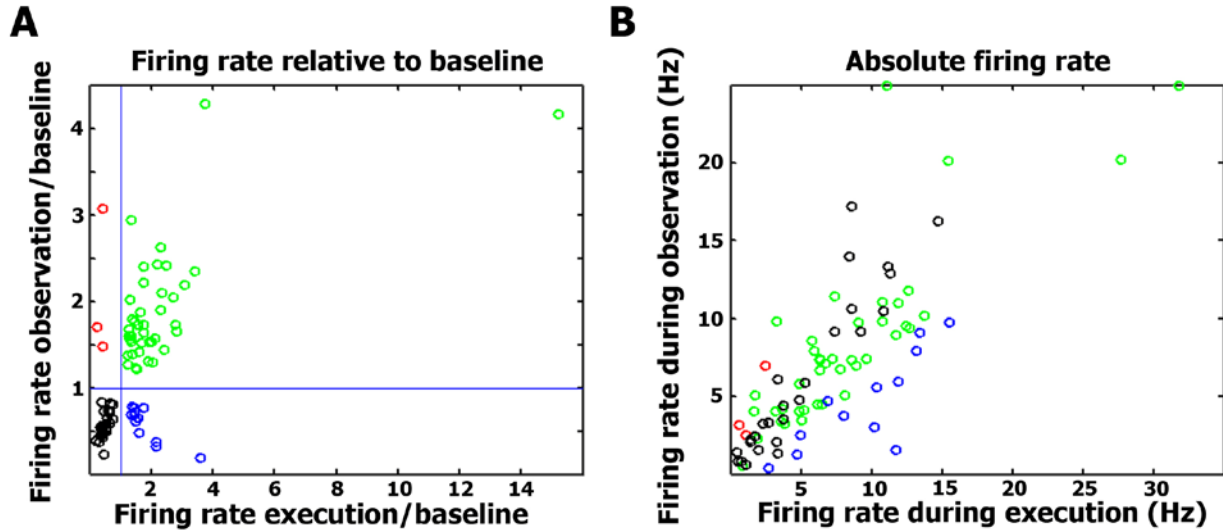


Figure S3 (related to Figure 3): Scatter plots of response amplitude during action-observation and action-execution. (A) For each neuron, the firing rate during action-execution was divided by the firing rate during baseline (x-axis). Similarly, the firing rate during action-observation was divided by the firing rate during baseline (y-axis). Green circles – cells exhibiting excitation to both conditions; Black circles – cells exhibiting inhibition during both conditions; Blue circles – cells exhibiting excitation during action execution and inhibition during action observation; Red circles – cells exhibiting excitation during action-observation and inhibition during action-execution. (B) Absolute firing rates of the same cells shown in (A).

Table S1: regional distribution of responses during action execution/observation**A**

Execution	A	H	EC	PHG	ACCr	ACCd	Pre-SMA	SMA
Face	3	10	13	16	18	13	10	14
Hand	6	16	16	16	26	11	18	20
Both	2	10	8	7	6	4	6	7

B

Observation	A	H	EC	PHG	ACCr	ACCd	Pre-SMA	SMA
Face	2	16	15	10	27	18	9	8
Hand	1	11	11	21	16	7	9	4
Both	1	2	6	4	2	1	1	1

C

	H	EC	PHG	SMA
Smile	3	5	2	3
Frown	6	2	6	3
Precision	3	1	3	3
Wholehand	5	2	5	2
Other	1	4	3	6

Table S1 (related to Table 1): Distribution of cells responding during action-execution (A) and action-observation (B) in the different anatomical regions. Face – number of cells responding during execution (observation) of a facial gesture (smile or frown); Hand – number of cells responding during execution (observation) of a hand grip (precision grip or wholehand prehension); Both – number of cells responding during execution (observation) of a facial gesture and also a hand grip. Within the population of cells responding during action-observation, the proportion of cells responding to observation of hand grasps in PHG was significantly larger than those responding to observation of facial gestures ($\chi^2(1) = 3.9$, $p = 0.04$). The proportion of cells responding to observation of facial gestures in ACCd was significantly

larger than the proportion of cells responding to observation of hand grasps ($\chi^2(1) = 4.8$, $p = 0.02$). C) Anatomical distribution of Action observation/execution matching cells for the different conditions (Smile, Frown, Precision grip, and Wholehand prehension). Other, refers to 14 cells matching more than one condition. Six of those cells matched both facial gestures (Smile and Frown; One cell in H, one in PHG, one in EC and three in SMA). Four cells matched both hand grasps (Precision and Wholehand; One in SMA, one in EC, and two in PHG). The remaining four cells matched one facial gesture, and one hand grasp (two cells in EC and one in SMA matched 'Frown' and 'Precision'; one cell in SMA matched 'Smile' and 'Wholehand').

Table S2: response profile of action observation/execution matching neurons

Cell #	Region	Execution				Observation				Control	Response type	Congruency	Unit type
1	LH	W	P*	S*	F	W	P	F	S		I	M	MU
2	LH	S*	F	W-	P-	F	S	P-	W-		E, E	B	SU
3	LH	S*	F	W	P	S	F	W	P-		I	M	MU
4	LH	F*	S*	W	P-	F	P	W	S		E	M	SU
5	LH	F	S	W-	P-	S	F	P	W		E	M	MU
6	LH	P	W	S	F	W	P	S	F		B	M	SU
7	LH	W*	P*	F-	S-	W	P	S-	F-	F	E	B	SU
8	LH	F	P	W	S	S	F	W	P-	P	E	M	MU
9	RH	P*	W	S-	F*-	W	P	F	S-		E	M	SU
10	RH	S	W	P	F	S	F	W	P	P, W	I	M	SU
11	RH	P	W	S	F	S	P	F	W-		E	M	MU
12	RH	W*	P	S	F	S	W	P	F		B	M	MU
13	RH	W*	P	F	S-	W	P	F-	S-		E	B	SU
14	RH	P	W	S-	F-	P	W	F-	S-		I	B	MU
15	RH	S	F	W	P*	S	F	W	P		B	M	MU
16	RH	F	S	P-	W*-	F	S	P-	W-		E	B	MU
17	RH	S	F	P	W*	S	F	P	W		E	M	MU
18	RH	W	F	P	S	P	W	S	F	W	E	M	SU
19	LEC	P*	W*	F*	S-	W	P	F-	S-		E	M	MU
20	LEC	S	F	W	P	S	P	F	W		B	B	SU
21	LEC	P*	F*	W	S-	P	S	W	F		I	M	SU
22	LEC	S*	F*	P-	W-	F	S	W-	P-	F	E	B	SU
23	LEC	P*	S	W	F	S	F	P	W		B, B	M	SU
24	LEC	F	P*	W	S-	W	F	S	P		E	M	MU
25	LEC	W*	P*	S*	F	S	W	F	P-	P	E	M	SU
26	LEC	F*	S*-	P*-	W-	F	S-	P-	W-		E, E	M	MU
27	REC	S	P	F	W	S	F	W-	P-		E	M	MU
28	REC	F	P	W	S	F	P	S	W		I, I	M	MU
29	REC	W*	F*	P	S*	F	S	W	P	P, W	B	M	SU
30	REC	W	P	S	F	W	S	P	F		B	M	SU
31	REC	P	F	W	S	S	F	P	W		I	M	SU
32	REC	F*	S	W*-	P*-	S	F	P-	W-		E, E	B	SU
33	LPHG	W	P	S-	F-	W	P	F	S-		E	M	SU
34	LPHG	F	S	P	W-	W-	P	S-	F-	S, F	E	B	SU

35	LPHG	S	W	F*	P	F	S	W-	P-		I	M	MU
36	LPHG	P	S	W	F-	F	W	P	S		I, I	M	SU
37	LPHG	F	S	W	P	F	W	S	P		B	M	SU
38	LPHG	F	S	P	W	F	S	W	P		I	M	MU
39	LPHG	W*	P	S	F	S	F	P	W	W	B	M	MU
40	LPHG	W	P	S	F	F	S	P	W	P	B	M	SU
41	LPHG	F	P	W	S	F	S	W	P		I	M	MU
42	LPHG	W	P	F	S	S	W	F	P		I	M	SU
43	RPHG	W*	F	P	S-	S	F	W-	P-	S	E	M	MU
44	RPHG	W	S	F	P	S	P	W	F		I	M	SU
45	RPHG	W	P	S-	F-	W	P	S-	F-		E	B	SU
46	RPHG	W	P	F-	S-	W	F	P	S-		I	M	SU
47	RPHG	S	P	W*	F	P	F	S	W		E	M	SU
48	RPHG	W	S	P	F-	W	P	F	S-		E, E	M	SU
49	RPHG	W*	P	F-	S-	P	F	W-	S-		E, E	M	SU
50	RPHG	P	W	S-	F-	S	F	P	W	F, W	E	M	MU
51	RPHG	P*	W*	S	F	S	F	P-	W-	S	E	M	SU
52	LSMA	F	S*	P	W*	F	S	P	W		B	B	MU
53	LSMA	P	W	F	S-	W	F	S	P-		E	M	SU
54	LSMA	W*	P	F	S	W	P	S	F		B	M	MU
55	LSMA	F	W	S*	P	S	F	P-	W-		I, I	M	SU
56	LSMA	F*	S*	W-	P*-	S	W	F-	P-		I, I	M	SU
57	LSMA	S	W	P	F-	S	F	W	P	F, P	I	M	MU
58	LSMA	F	S	W	P-	F	S	P-	W-		I, I	M	SU
59	LSMA	S*	F*-	W-	P-	F	W	P	S-		E	M	SU
60	LSMA	S	P	F	W-	S	F	W	P	P	I	M	SU
61	LSMA	S*	F*	P*-	W-	F	S	P-	W-		E	B	MU
62	LSMA	W*	P*	S-	F-	W	P	F-	S-		E, E	B	SU
63	LSMA	W	F	P	S-	F	S	P-	W-		I	M	SU
64	LSMA	P*	W*	F*	S-	P	W	S	F-	W	I	M	MU
65	LSMA	P	W	F*	S	S	F	P	W-		E, E	M	SU
66	LSMA	W*	P	S-	F-	W	P	F-	S-		E, E	B	MU
67	RSMA	W*	P	S	F	W	F	S	P		B	M	SU
68	RSMA	W	S	F	P	S	F	W	P		B	M	MU

Table S2 (related to Table 1): Response details of all cells matching execution/observation. Column 1: Serial number of cell. Column 2 (Region): first letter (L or R) corresponds to the hemisphere from which the cell was recorded (Left or Right). Columns 3 – 6 (Execution): letters correspond to the different conditions (S = ‘Smile’, F = ‘Frown’, P = ‘Precision’, and W = ‘Wholehand’). Conditions are in descending order of response magnitude (firing rate). Thus for

excitatory responses, firing rate in column 3 > firing rate in column 6 and for inhibitory responses firing rate in column 3 < firing rate in column 6. Red letters correspond to significant difference in firing rate relative to baseline and asterisks denote significant difference in firing rate relative to the corresponding control condition. Minus signs denote significant difference of response in columns 4, 5, and 6 relative to condition in column 3. Columns 7 – 10 (Observation): same as columns 3 – 6 but for the observation condition. Column 11 (Control): significant responses (if any) to control conditions (letters same as in columns 3 – 10). Column 12 (Response type): Excitatory (E), Inhibitory (I), or Both (B representing excitation to execution and inhibition to observation and **B** excitation to observation and inhibition to execution). Column 13 (Congruency): Broadly congruent cells (B) and Matching cells (M) – see supplemental experimental procedures for definition). Column 14 (Unit type): Single unit (SU) vs. Multi unit (MU).

Table S3: Excitation vs. Inhibition**A**

	H	EC	PHG	SMA
Excitation	11 (6, 5)	7 (3, 4)	9 (7, 2)	6 (4, 2)
Inhibition	4 (1, 3)	3 (2, 1)	7 (4, 3)	7 (5, 2)
Both	3 (1, 2)	3 (3, 0)	1 (0, 1)	4 (1, 3)

B

	PHG (N = 11)	EC (N = 9)	H (N = 12)	SMA (N = 9)
Observation	518±256	777±307	741±253	677±263
Execution	700±272	855±269	883±248	633±244

C

	PHG (N = 8)	EC (N = 4)	H (N = 4)	SMA (N = 10)
Observation	612±216	850±238	650±100	540±309
Execution	762±324	725±50	625±221	660±171

Table S3 (related to Figure 3): **A)** Anatomical distribution of responses of Observation/Execution matching cells. Top row, number of cells (single, multi units) responding with excitation during both action-execution and action-observation. Middle row, number of cells responding with inhibition to both conditions. Bottom row, cells responding with excitation during action-execution and inhibition during action-observation. Two additional cells in PHG responded with excitation during action-observation and inhibition during action-execution. One more cell in EC responded to two conditions (inhibition to frown-execution and excitation to frown-observation; excitation to precision-execution and inhibition to precision-observation). Regional acronyms as in Table 1. **B, C)** Latencies (in ms) of excitatory (**B**) and inhibitory (**C**) action observation/execution matching cells. Latencies were computed as the time between stimulus onset and the first time bin at which neural response reached maximum/minimum value (bin size 100ms). Within each region, no statistical difference between observation and execution latencies was found (two-tailed, paired t-test across all cells). We also compared the latency of SMA neurons with all other temporal lobe regions. The SMA excitatory responses had shorter latencies during action-execution compared with the hippocampus ($p = 0.03$, two-sample equal variance t-test).

Supplemental Experimental Procedures

Patients and experimental setup

We recorded extracellular single and multi unit activity from patients with pharmacologically intractable epilepsy, implanted with intracranial depth electrodes to identify seizure foci for potential surgical treatment. Data was acquired from 21 patients in 43 sessions (range 1 to 5 sessions per patient; median = 2). Mean patient age was 31 (range 18 – 54); 12 males; 15 right handers. Electrode location was based solely on clinical criteria. Each electrode terminated in a set of nine 40- μm platinum-iridium microwires and the signals from eight micro-electrodes were referenced to the ninth, lower impedance micro-electrode [1]. Data was recorded at 28kHz using a 64-channel acquisition system (Neuralynx, Tucson, AZ) and the signals were band-pass filtered between 1Hz and 9kHz. The beginning and end of each experimental trial was marked by electrical triggers sent from the laptop to the recording device. Patients provided written informed consent to participate in the experiments. The study conformed to the guidelines and was approved by the Medical Institutional Review Board at UCLA.

Experiment design

The entire experiment was composed of three parts – ‘*Facial expressions*’, ‘*Grasping*’ and a ‘*Control*’ experiment. Order of experimental parts was randomized across subjects. Stimuli were presented on a standard laptop at the patient’s bed. In the case of the grasping part of the experiment, a mug was placed next to the laptop.

Facial expressions

Patients were presented with pictures of smiling or frowning faces, or with the written word ‘Smile’ or ‘Frown’. The patients were instructed to simply observe the picture and avoid making any facial gestures, but to perform the facial gesture when the written word smile or frown was presented. The experiment started with 6 seconds of a blank grey screen. The pictures/text instructions were presented for one second and followed by a blank grey screen which lasted

either 5 or 6 seconds randomly. Pictures of 16 different individual faces were presented (8 male and 8 female faces) in either a smiling or frowning configuration (total of 32 different images). Each individual image was presented once thus there were 16 trials for smile observation and 16 trials for frown observation. Similarly there were 32 facial gesture execution trials (16 'smile', and 16 'frown'). The order of trials was counter balanced. Total duration for this part was 7:08 minutes.

Grasp

The patients were presented either with 3 second video clips depicting a hand grasping a mug or with the written word 'Finger' or 'Hand'. They were instructed to observe the video clip and refrain from making any hand movements. The video clips depicted a hand grasping a mug with either precision grip or whole-hand grasp. When written words were presented, the patients performed a precision grip (for the word 'Finger') or a wholehand grasp (for the word 'Hand') on a mug placed next to the laptop. The patients performed 36 observation trials (18 trials of precision grasps and 18 trials of whole hand grasps) and 36 execution trials (18 'Finger' and 18 'Hand'). Each trial was followed by a blank grey screen lasting either 5 or 6 seconds. The order of trials was counterbalanced. This part of the experiment lasted 9:12 minutes.

Control

The patients were presented with a written word for one second (either 'Smile', 'Frown', 'Finger' or 'Hand'). These words were the ones used as cues for action-execution in the previously described parts of the experiment (*Facial expression* and *Grasp*). Each word presentation was followed by a blank grey screen lasting either 5 or 6 seconds randomly. Patients were instructed to covertly read the words and refrain from making hand movements/facial gestures. This part of the experiment, lasting 3:40 minutes, was composed of 32 trials (8 trials for each word) presented in a counterbalanced fashion.

The first five patients performed a variation of the task. In the execution conditions of the *Facial expression* and *Grasp* parts of the experiment, instead of a word appearing on the screen to cue the patient to perform the appropriate action, a 100 millisecond auditory tone was used. A low tone (250Hz) cued the patient to frown or perform a whole hand prehension in the *Facial expression* and *Grasp* experiments respectively. Similarly, a high tone (1000Hz) indicated to smile or perform a precision grasp. In the control experiment the same tones were played but the patients were explicitly instructed to simply listen to the tones and avoid making any facial gestures or hand movements during the experiment. There were a total of 24 hand execution trials (12 precision grasp, 12 wholehand prehension), and similarly another 24 hand observation trials. In the *Facial expression* experiment there were 32 observation trials (16 smiling, 16 frowning) and 24 execution conditions (12 smiling, 12 frowning). The execution and observation conditions were separated into blocks and the patients were notified in advance if it was an observation block or an execution block. In the *Control* experiment, each beep was sounded 5 times in a counterbalanced fashion. In order to simplify the task for the patients, the remaining 16 patients were explicitly cued for action execution using written words as described above. Our analyses did not reveal differences between the responses of the first five patients and of the remaining 16 patients. Thus, data from both groups are collapsed here.

Data analysis

Anatomical Localizations

In order to verify the position of implanted electrodes, CT scans following electrode implantation were co-registered to the preoperative MRI using Vitrea® (Vital Images Inc.). In the frontal lobe, we recorded from 16 different patients in rostral ACC, 7 patients in dorsal ACC, 6 patients in pre-SMA, and 5 patients in SMA. In the temporal lobe, we recorded from 4 patients in the Amygdala, 15 patients in the Hippocampus, 7 patients in entorhinal cortex, and 12 patients in the parahippocampal gyrus.

Spiking activity

The raw signal was band-pass filtered between 300 and 3000 Hz and a threshold of five standard deviations above the median of the filtered signal was used to detect suspected action potentials. The suspected action potentials were then clustered and manually sorted as spikes or electrical noise [2]. Similar to [3], the classification between single unit and multi-unit was done visually based on the following: (1) Average spike shape and its variance; (2) the ratio between the spike peak value and the noise level; (3) the inter-spike interval distribution of each cluster; and (4) the presence of a refractory period for the single units (that is, less than 1% of spikes within less than 3ms inter-spike interval).

Statistical analysis

For each neuron, and each condition, we assessed responsiveness by comparing the firing rate during baseline (-1000ms to 0ms relative to stimulus onset) and firing rate during the experimental condition (+200ms to +1200ms relative to stimulus onset) on a trial by trial basis using a two-tailed paired t-test. The statistical significance threshold for the paired t-test across trials was set at 0.05.

Bootstrap analysis (figures S3, S4)

In order to assess whether or not the number of Action Observation/Execution matching neurons in each region is significant, we did the following. For each neuron recorded in a given region (regardless of responsiveness), we shuffled the spike trains from each trial across the different conditions. Thus, in the new shuffled data-set each individual spike-train is real but it is assigned randomly to the different conditions (e.g. a spike-train originally recorded during smile-execution will be assigned to precision-grip observation in the shuffled data-set). Next, we assessed whether or not the neuron would be considered an Action Observation/Execution matching neuron using the same criteria we used for the original data. This was performed for all recorded cells in a given anatomical region and the proportion of pseudo Action Observation/Execution matching neurons was computed. In order to calculate the null distribution, this procedure was repeated for 10,000 iterations. Figures S3, and S4 display the

distribution of proportions of Action Observation/Execution matching neurons across all iterations.

Average response profile (Figure 3)

To calculate the average response profile of cells, we first determined if the response to a given condition was inhibitory or excitatory by using a paired t-test across trials comparing the response to baseline. Subsequently, to average across response profiles of neurons with different firing rates we normalized the PSTH of each neuron, in a fashion similar to Fogassi and colleagues [4]. Excitatory responses were normalized by subtracting the average response during baseline (-1000 to 0ms relative to trial onset), and dividing by the maximum firing rate of the response (bin size = 200ms). Inhibitory responses were normalized by removing the average response during baseline and dividing by the absolute value of the minimum of the response. In this manner, the excitatory response of each neuron ranges between 0 and +1, whereas the inhibitory response of each neuron ranges between 0 and -1. Significant differences between the temporal response profile during action-execution and action-observation were assessed using a two-tailed t-test and a significance level of 0.05 (asterisks in figure). In the case of the control condition, we assessed significant difference from zero.

Congruency of response (supplementary Table 2)

All action observation/execution matching neurons displayed significant deviation from baseline firing rate during the observation and execution of the matched action (and not during the corresponding control condition). However, a more stringent criterion of response matching selectivity, is that the response amplitude for the matched action (during both observation and execution) is also statistically different than the response amplitude for other actions. We defined Broadly congruent cells (B) as cells whose response amplitude to the matched action with one effector (e.g. hand) was statistically higher than the response amplitude of both actions of the other effector (e.g. face) during both observation and execution. Matching cells (M) were defined

as cells having significant response amplitude relative to baseline but not compared to actions with the other effector during both execution and observation.

Supplemental references

1. Fried, I., Wilson, C.L., Maidment, N.T., Engel, J., Jr., Behnke, E., Fields, T.A., MacDonald, K.A., Morrow, J.W., and Ackerson, L. (1999). Cerebral microdialysis combined with single-neuron and electroencephalographic recording in neurosurgical patients. Technical note. *J Neurosurg* *91*, 697-705.
2. Quiroga, R.Q., Nadasdy, Z., and Ben-Shaul, Y. (2004). Unsupervised spike detection and sorting with wavelets and superparamagnetic clustering. *Neural Comput* *16*, 1661-1687.
3. Quiroga, R.Q., Reddy, L., Kreiman, G., Koch, C., and Fried, I. (2005). Invariant visual representation by single neurons in the human brain. *Nature* *435*, 1102-1107.
4. Fogassi, L., Ferrari, P.F., Gesierich, B., Rozzi, S., Chersi, F., and Rizzolatti, G. (2005). Parietal lobe: from action organization to intention understanding. *Science* *308*, 662-667.