Figure S1

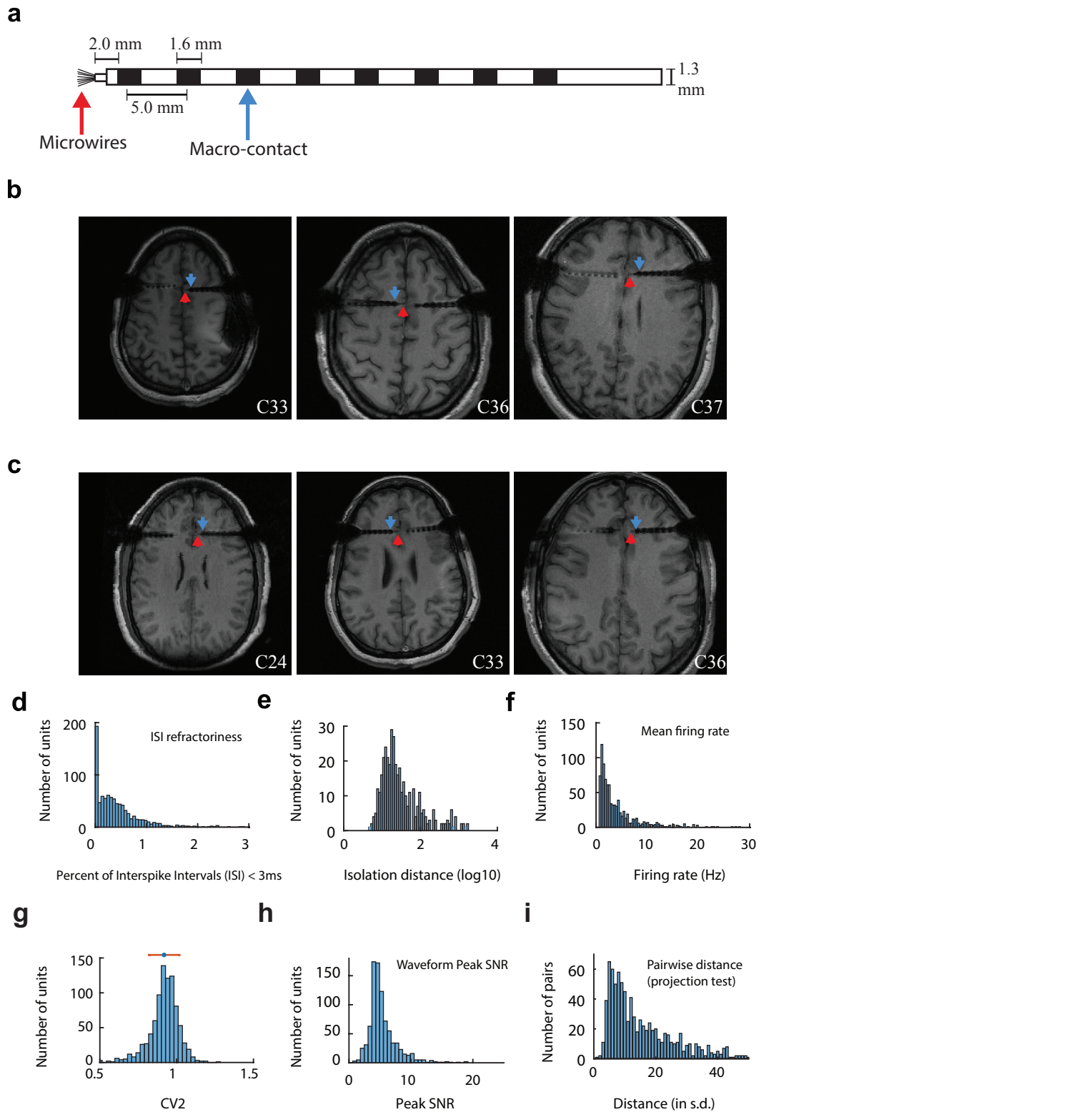

Figure S1. Recording electrode, example post-operative structural MRIs and spike sorting quality. Related to Figure 1-2.

(a) The hybrid macro-micro electrode used. Individual neurons are recorded using high-impedance microwires (red arrow; diameter 40µ, impedance 400-600kΩ). Field potentials are recorded from the low-impedance (<2kΩ) macro-contact most adjacent to the micro-wires (blue).
(b,c) Example axial T1 MRI scans of recording locations used in pre-SMA (b) and dACC (c). Red and blue arrowheads indicate tips of microwires and the macro contacts used, respectively.
(d-i) Spike sorting quality. Metrics quantifying the individual clusters that we used as putative single-units. (d) Histogram of proportion of inter-spike intervals (ISIs) shorter than 3ms. Most of our recorded clusters had less than 0.5% of their ISIs smaller than 3ms. (e) Isolation distance of all units for which this metric was defined (median 21.5). (f) Histogram of mean spike rates. (g) Histogram of coefficient-of-variation (CV2) values of all units. (h) Histogram of the signal to noise ratio (SNR) of the mean waveform peak computed for each unit. (i) Pairwise distance between all possible pairs, calculated using the projection test (see methods), of units on all wires with at least one cluster isolated. Distance is in unit of standard deviation after normalizing the data such that the distribution of waveforms around their mean is equal to one.

Figure S2

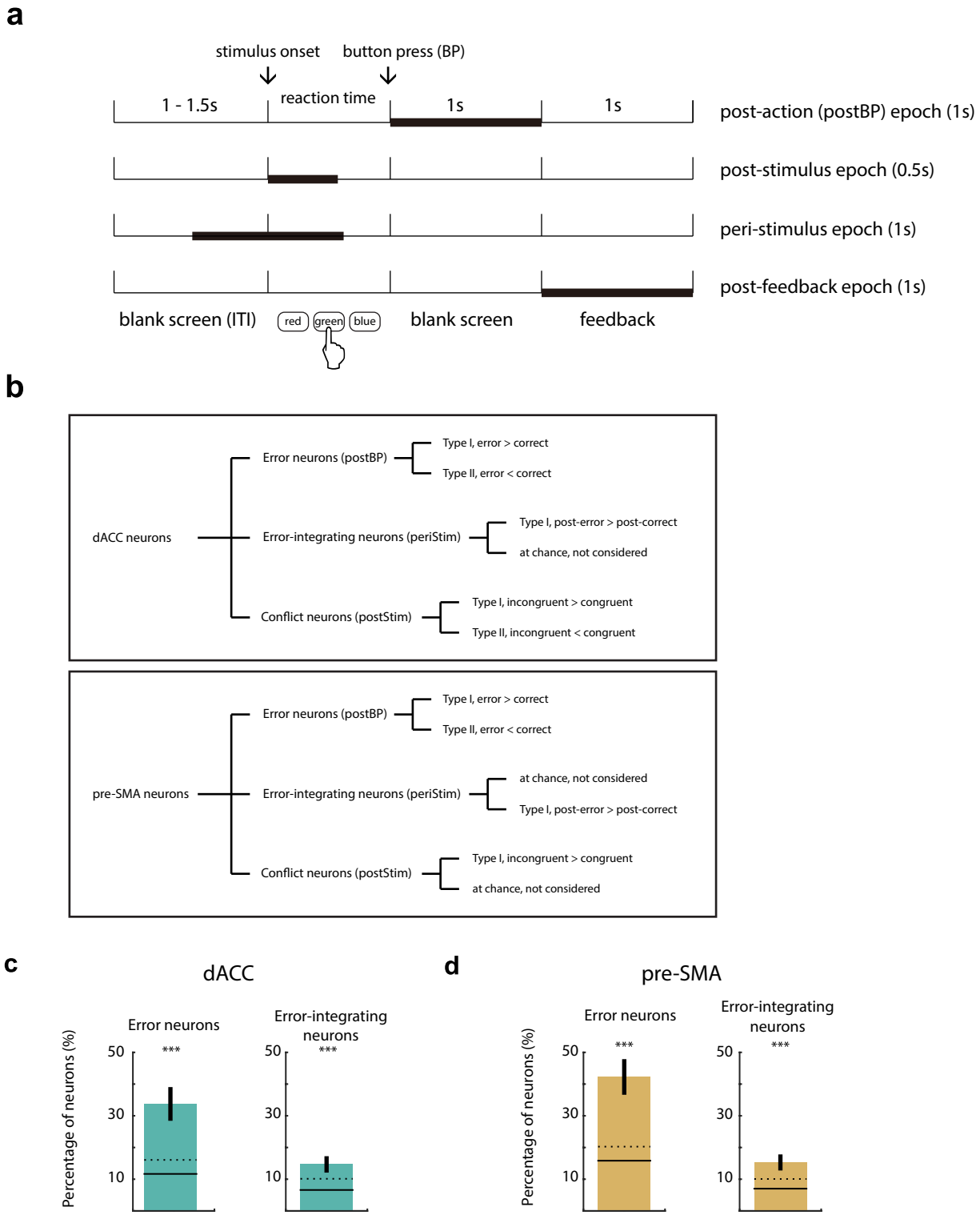

Figure S2. Summary of epochs of interest, neuronal categories and distribution of neuronal categories separately for each session. Related to Figures 2-3.

(a) Epochs used to analyze spiking activity. Thick lines indicate the extent of the time windows. Length of each analysis epoch is indicated in brackets on the right.
(b) Summary of neuronal categories identified in dACC and pre-SMA. The second level lists all neuronal types and the time window (brackets) in which we identified more neurons than expected by chance. The third level lists all sub-types (Type I or II) which were identified at levels higher than expected by chance. The contrasts listed refer to the spike rates during the trial types mentioned (e.g. 'error > correct' means spike rates in the error trials were larger than those in the correct trials for this particular type of neurons).
(c-d) Percentage of significant neurons identified in dACC (c) and pre-SMA (d) that qualified as error or error-integrating neurons across recording sessions. Error bars represent ± s.e.m across sessions, solid and broken horizontal lines, the mean and the 97.5th percentile of the null distribution of the number of neurons expected by chance as estimated using permutation tests.
'***' marks groups of neurons which were observed more than expected by chance with p values ≤ 0.001.

Figure S3

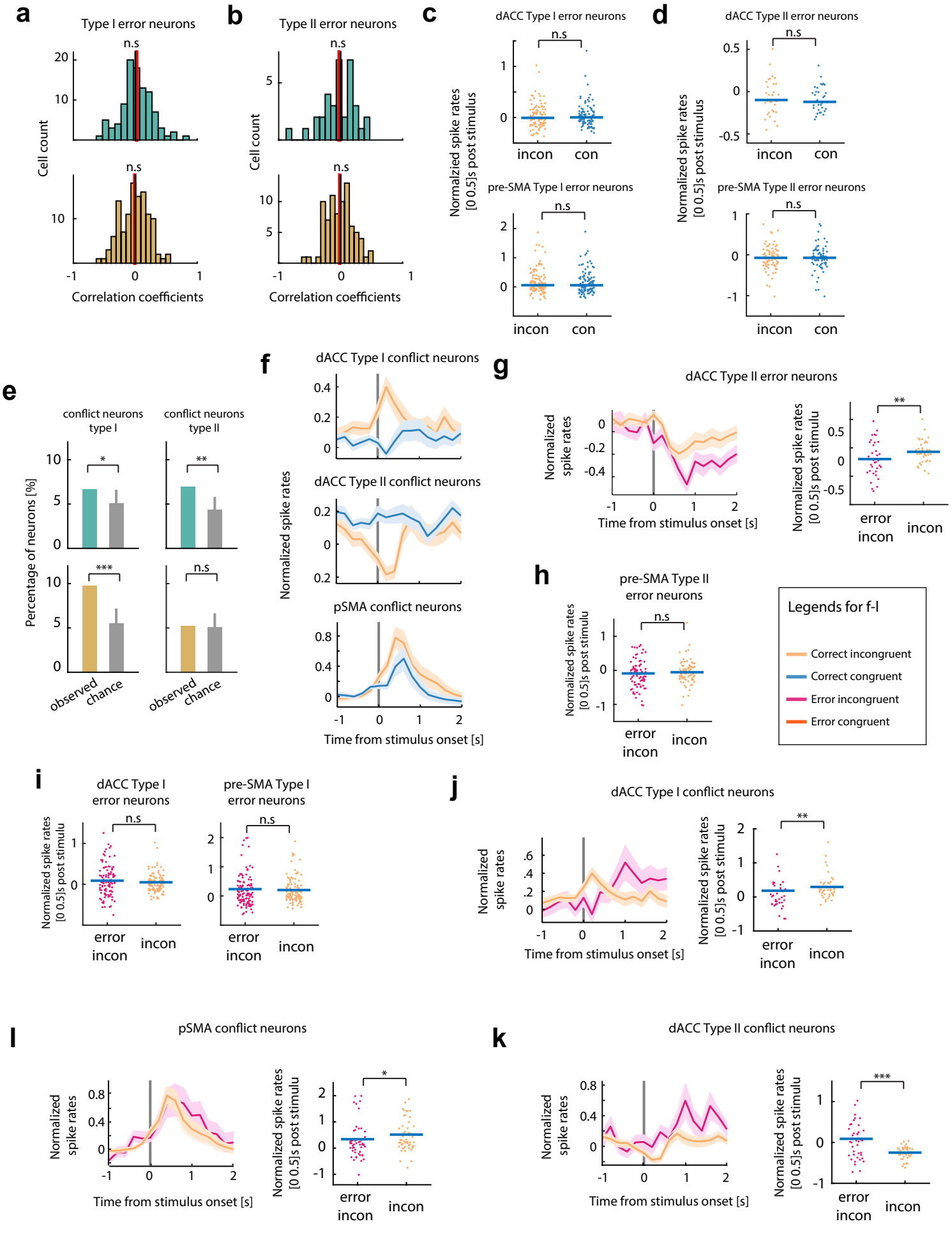
a
Type I error neurons
n.s
Cell count
Correlation coefficients
b
Type II error neurons
c
dACC Type I error neurons
pre-SMA Type I error neurons
Normalzied spike rates [0 0.5]s post stimulus
incon
con
d
dACC Type II error neurons
pre-SMA Type II error neurons
e
conflict neurons type I
conflict neurons type II
Percentage of neurons [%]
observed
chance
f
dACC Type I conflict neurons
dACC Type II conflict neurons
pSMA conflict neurons
Normalized spike rates
Time from stimulus onset [s]
g
dACC Type II error neurons
Normalized spike rates
Time from stimulus onset [s]
error incon
h
pre-SMA Type II error neurons
Legends for f-l
Correct incongruent
Correct congruent
Error incongruent
Error congruent
i
dACC Type I error neurons
pre-SMA Type I error neurons
j
dACC Type I conflict neurons
k
dACC Type II conflict neurons
l
pSMA conflict neurons

Figure S3. Signatures of conflict and control. Related to Figure 3.

(a) Correlation between RT and the spike rate of Type I error neurons identified in dACC (top) and pre-SMA (bottom) on the same error trials. There was no significant correlation ($p > 0.04$, $t(98) = 0.86$, t-test) in either area.
(b) Same as in (a) but for Type II error neurons. There was no significant correlation ($p > 0.05$, $t(117) = -0.41$, t-test) in either area.
(c) Spike rate of Type I error neurons did not differentiate significantly between correct congruent and incongruent trials in both dACC (upper; $p = 0.92$, z value = 0.1) and pre-SMA (lower; $p = 0.18$, z value = 1.33). Each data point shown is one neuron. Spike rates were quantified within a bin of 500ms size starting at stimulus onset and normalized by the baseline spike rates ([-700ms -200ms] relative to stimulus onset). Blue horizontal bars represent median values of the population.
(d) Same as in (c) but for Type II error neurons. Spike rate of Type II error neurons did not differentiate significantly between correct congruent and incongruent trials in both dACC (upper; $p = 0.61$, z value = 0.51) and pre-SMA (lower; $p = 0.91$, z value =0.12).
(e) Number of conflict neurons identified in dACC (green) and pre-SMA (brown). For the definition of Type I and II, see Methods. In dACC, both Type I and Type II conflict neurons have significantly greater number than that is expected by chance ($p = 0.03$ for Type I, $p < 0.001$ for Type II conflict neurons, permutation test). Gray bar shows the mean value of the empirical null distribution. Error bar shows the 95th percentile of the empirical null distribution.
(f) Average spike rates as a function of time for Type I conflict neurons (top), Type II conflict neurons (middle) in dACC and conflict neurons in pre-SMA (bottom). The spike rates were normalized by the baseline ([-700ms -200ms] relative to stimulus onset). Gray bar marks the onset of stimulus.
(g) Signature of control. Average spike rates as a function of time (left) and within the post-stimulus epoch ([0 500ms] relative to stimulus onset) of Type II error neurons in dACC for error incongruent vs. correct incongruent trials. The spike rates within the post-stimulus epoch differentiated error incongruent and correct incongruent trials significantly ($p = 0.0062$, z value = -2.74, Wilcoxon's signed rank test). Spike rates were normalized by the baseline ([-700ms -200ms] relative to stimulus onset). Gray bar marks the onset of stimulus. Blue bars on the scatter represents median of the population.
(h) Same as in (g) but for Type II error neurons in pre-SMA. The spike rates within the post-stimulus epoch did not differentiate error incongruent and correct incongruent trials significantly ($p = 0.26$, z value = -1.12).
(i) Same as in (g) but for Type I error neurons in dACC (left; $p = 0.41$, z value = 0.82, Wilcoxon's signed rank test) and pre-SMA (right; $p = 0.26$, z value = -1.12).
(j) Same as in (g) but for Type I conflict neurons in dACC. The spike rates within the post-stimulus epoch differentiated error incongruent and correct incongruent trials significantly ($p = 0.006$, z value = -2.75).
(k) Same as in (g) but for Type II conflict neurons in dACC. The spike rates within the post-stimulus epoch differentiated error incongruent and correct incongruent trials significantly ($p < 0.001$, z value = 3.54).
(l) Same as in (g) but for conflict neurons in pre-SMA. The spike rates within the post-stimulus epoch differentiated error incongruent and correct incongruent trials significantly ($p = 0.034$, z value = -2.12).
(c,f) Orange represent correct incongruent trials and blue represents correct congruent trials.
(g-l) Orange represents correct incongruent trials and magenta represents error incongruent trials.
'*', '**', '***' mark groups of neurons which were observed in proportions different then in the overall population with p values ≤ 0.05, ≤ 0.01 and ≤ 0.001 respectively (for a-b, t-test; for c,g-l, Wilcoxon's signed rank test). 'n.s' marks not significant ($p > 0.05$).

Figure S4

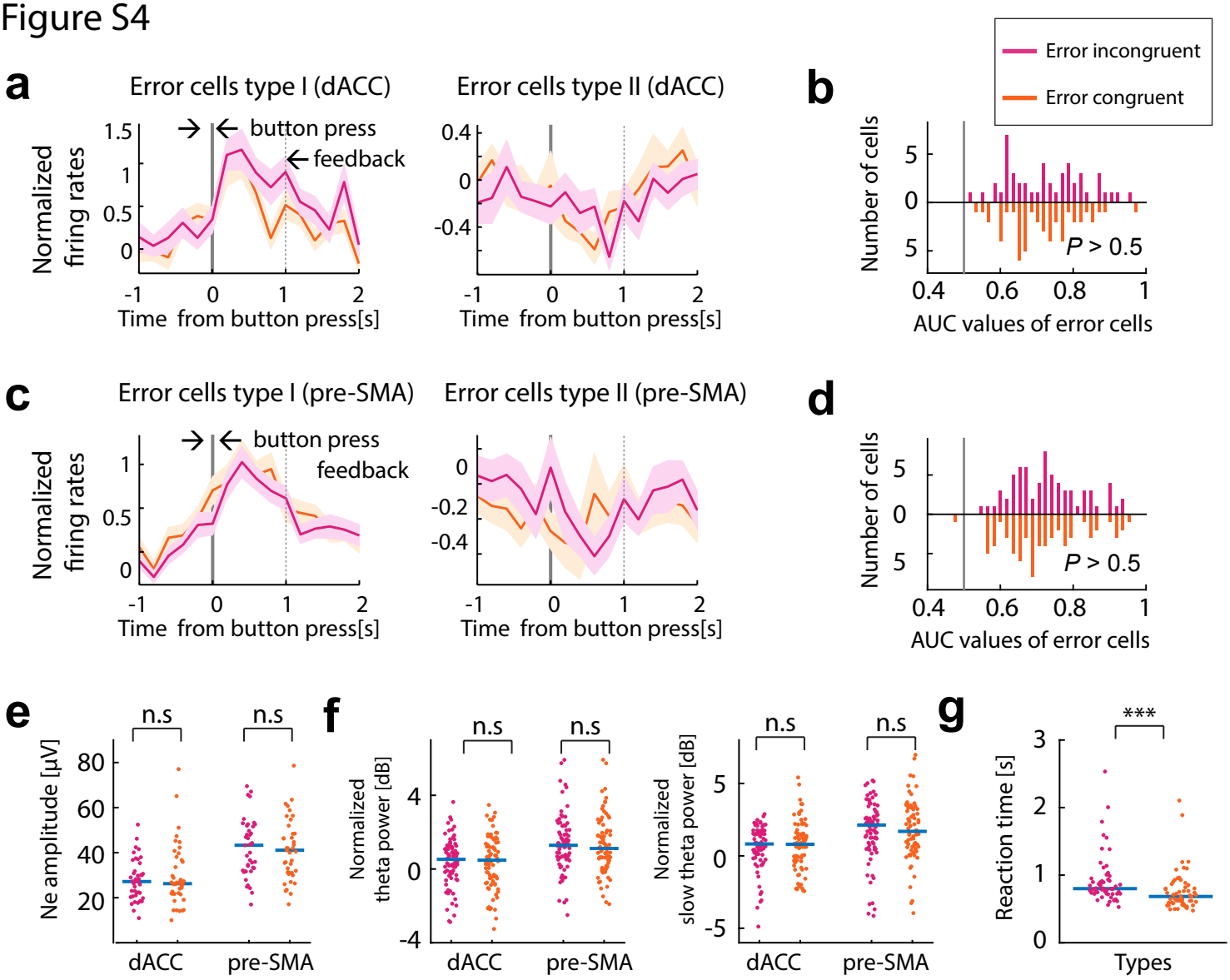

Figure S4. Response of error neurons and the iERN did not differ by error types (congruent error/incongruent error). Related to Figure 3 and 5.

(a) Average spike rates as a function of time for error neurons (Type I and II) in the dACC, normalized by the baseline. Right, Type I error neurons; Left, Type II error neurons. Response is aligned at button press (right). Trials are grouped by congruence (colors; magenta for error incongruent trials and orange for error congruent trials).
(b) Single-neuron ROC analysis of error neurons (both Type I and II) in dACC. The ability of spike rates in the post-button press time window ([0 1]s relative to button press) to differentiate error incongruent and error congruent trials each from correct trials did not differ significantly (AUC values computed from differentiating between correct/congruent error, between correct/incongruent error, 0.59±0.02 vs. 0.58 ±0.02 in dACC, 0.54±0.02 vs. 0.53±0.02 in pre-SMA; $p > 0.5$ for both areas, Wilcoxon rank sum test).
(c) Same as in (a), but for pre-SMA.
(d) Same as in (b), but for pre-SMA.
(e,f) iERN did not differ between incongruent and congruent errors (iERN amplitude comparisons: $p = 0.8$, $z = 0.25$ for dACC, $p = 0.93$, $z = -0.09$ for pre-SMA, signed rank test. Theta power comparisons: $p = 0.72$, $z = -0.35$ for dACC, $p = 0.93$, $z = -0.09$ for pre-SMA, signed rank test; Slow theta power comparisons; $p = 0.49$, $z = -0.68$ for dACC, $p = 0.19$, $z = -1.3$ for pre-SMA, signed rank test). Shown are the ERN amplitudes (e), theta and slow-theta power (f) for dACC and pre-SMA. Color code, same as in (a-d). Each dot shows one session, horizontal line shows the mean.
(g) The Stroop effect was significant on error trials. Average reaction times in the error incongruent trials were significantly longer than in the error congruent trials ($p < 0.001$, sign rank test). Errors thus did not occur due to an absence of conflict processing.
'***' mark statistical comparison with p value ≤ 0.001. 'n.s' marks not significant ($p > 0.05$).

Figure S5

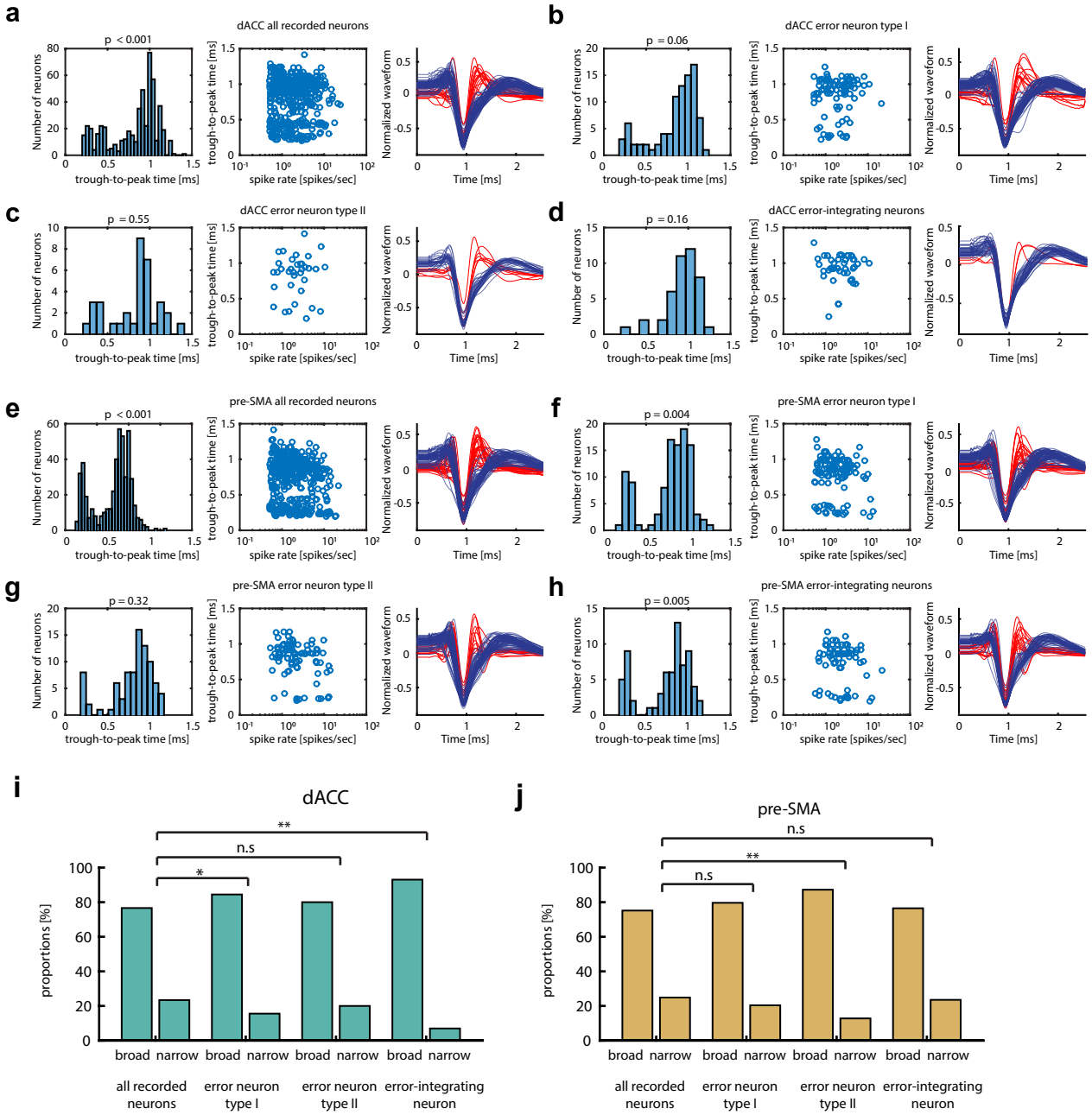

Figure S5. Waveform analyses of error and error-integrating neurons. Related to Figure 2.

(a-h) Distribution of trough-to-peak time (left), trough-to-peak time as a function of spike rates (middle) for each recorded neuron of a given group (described below). The rightmost plot shows the average spike waveforms of all neurons in the group, colored either blue or red depending on whether their trough-to-peak time was longer (blue) or shorter (red) than 0.5ms.
(a) All recorded neurons in dACC. The trough-to-peak distribution is significantly bimodal ($p < 0.001$).
(b) All Type I error neurons in dACC. The distribution of trough-to-peak time is not significantly different from unimodal ($p = 0.05$).
(c) All Type II error neurons in dACC. The distribution of trough-to-peak time is not significantly different from unimodal ($p = 0.55$).
(d) All error-integrating neurons in dACC. The distribution of trough-to-peak time is not significantly different from unimodal ($p = 0.16$).
(e) All recorded neurons in pre-SMA. The distribution of trough-to-peak time is significantly bimodal ($p < 0.001$).
(f) All Type I error neurons in pre-SMA. The distribution of trough-to-peak time is significantly bimodal ($p = 0.004$).
(g) All Type II error neurons in pre-SMA. The distribution of trough-to-peak time is significantly bimodal ($p = 0.32$).
(h) All error-integrating neurons in pre-SMA. The distribution of trough-to-peak time is significantly bimodal ($p = 0.005$).
(i) Proportions of putative pyramidal neurons (trough-to-peak time > 0.5ms) and interneurons (trough-to-peak time < 0.5ms) in dACC. Type I error neurons and error-integrating neurons have a significantly lower proportion of putative inhibitory neurons than the rest of the dACC population (15% and 7% vs 25% in the overall population, $p = 0.05$, odds ratio = 1.81 and $p = 0.0074$, odds ratio = 4.41, respectively; Fisher's exact test).
(j) Same as in (i) but for pre-SMA. Only the Type II error neurons had a significantly lower proportion of putative inhibitory neurons than the rest of the pre-SMA population (12% vs 26% in the overall population, $p = 0.0034$, odds ratio = 2.58, Fisher's exact test).

'*', '**' mark groups of neurons which were observed in proportions different then in the overall population with p values ≤ 0.05, ≤ 0.01, respectively (Hartigan's dip test). 'n.s' marks not significant ($p > 0.05$).

Figure S6

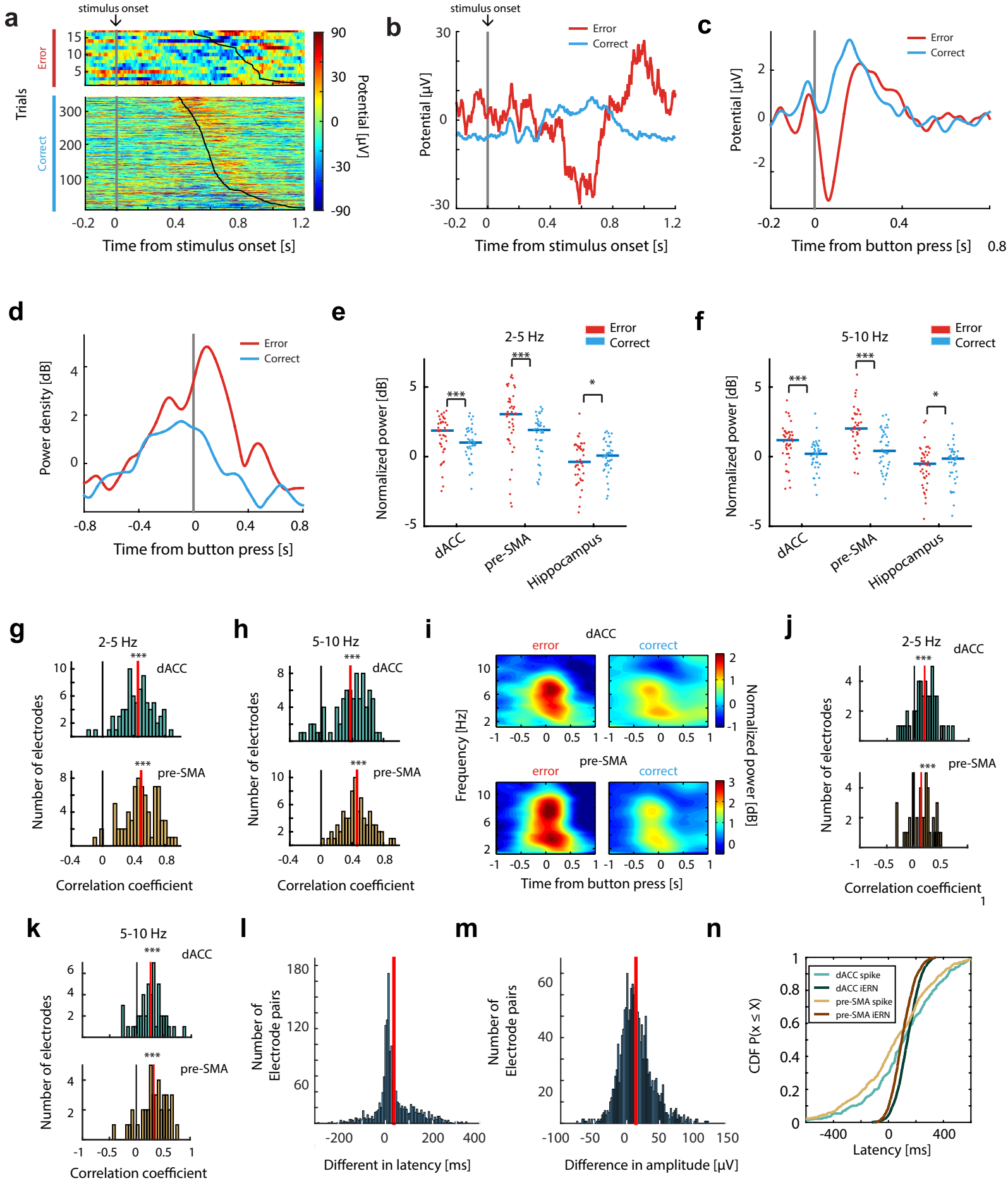
a
stimulus onset
Error
Correct
Trials
Potential [μV]
Time from stimulus onset [s]
b
stimulus onset
Error
Correct
Potential [μV]
Time from stimulus onset [s]
c
Error
Correct
Potential [μV]
Time from button press [s]
d
Error
Correct
Power density [dB]
Time from button press [s]
e
2-5 Hz
Error
Correct
Normalized power [dB]
dACC
pre-SMA
Hippocampus
f
5-10 Hz
Error
Correct
Normalized power [dB]
dACC
pre-SMA
Hippocampus
g
2-5 Hz
dACC
pre-SMA
Number of electrodes
Correlation coefficient
h
5-10 Hz
dACC
pre-SMA
Number of electrodes
Correlation coefficient
i
dACC
error
correct
pre-SMA
error
correct
Frequency [Hz]
Normalized power [dB]
Time from button press [s]
j
2-5 Hz
dACC
pre-SMA
Number of electrodes
Correlation coefficient
k
5-10 Hz
dACC
pre-SMA
Number of electrodes
Correlation coefficient
l
Number of Electrode pairs
Different in latency [ms]
m
Number of Electrode pairs
Difference in amplitude [μV]
n
dACC spike
dACC iERN
pre-SMA spike
pre-SMA iERN
CDF P(x ≤ X)
Latency [ms]

Figure S6. Stimulus-onset aligned intracranial event-related responses, statistics for time-frequency analysis of iERN, and scalp ERN control. Related to Figure 5.

(a) The same example single-trial event-related potential data as shown in Fig. 5a, but aligned at stimulus onset (t=0). The trials were sorted by reaction time (black lines; RT increases from top to bottom) and trial types (upper: error trials, lower: correct trials). Color bar represents ERP amplitude. Note the prominent ERP activities following button press (black line) as well as shortly after stimulus onset (blue). Gray bar represents stimulus onset.
(b) Average of data shown in (a) by trial types (colors; red for error, green for correct), aligned at stimulus onset (t = 0). Note that the sensory-evoked potential did not differ between trial types. Gray bar represents stimulus onset.
(c-d) Scalp-EEG recordings of non-surgical control subjects (N = 12) performing the same task reproduced the classical error-related negativity (ERN) and response-locked theta power (Compare with Fig. 5a-b).
(c) ERN (negative peak following button press, red) is significantly larger in amplitude in error compared to in correct trials (blue, $t(11) = 4.53$, $p < 0.001$). Gray bar represents button press.
(d) Theta power as a function of time. Error-related theta power (red) is significantly larger compared to in correct trials (green) after button press ($t(11) = 6.47$, $p < 0.001$). Gray bar represents button press.
(e) Power in the 2-5Hz band (0 to 500ms following button press) increased significantly more in error trials than in correct trials in both dACC ($p < 10\text{-}5$, z value = 6.17) and pre-SMA ($p < 10\text{-}5$, z value = 7.3). By contrast, hippocampal theta power also differed, but these differences were of opposite sign ($p = 0.01$, $z = -2.55$ for theta power, $p = 0.01$, $z = -2.56$ for slow theta power). Each dot shows one session, horizontal bar shows mean.
(f) Same as in (e) but for power in the 5-10Hz band.
(g) Mean Pearson's correlation coefficients between iERN amplitude and slow theta (2-5Hz) power are significantly larger than zero over all electrodes in dACC (mean correlation= 0.44, $p < 10\text{-}10$, $t(78) = 19.2$) and pre-SMA (mean correlation = 0.48, $p < 10\text{-}10$, $t(79) = 19.4$, $t(79) = 11.52$, t-test versus 0) Red vertical bars show population means.
(h) Same as in (g) but for correlations between iERN amplitude and theta (5-10Hz) power in dACC (mean correlation = 0.25, $p < 10\text{-}10$, $t(78) = 9.7$) and pre-SMA (mean correlation = 0.33, $p < 10\text{-}10$, $t(79) = 11.52$, t-test versus 0).
(i) Induced theta power, calculated after subtracting the ERP for each condition separately. Spectrograms shown are averaged across all sessions, see panel h for single-session statistics.
(j) Induced power was significantly correlated with the iERN amplitude in the slow theta (2-5Hz) band in both dACC (mean correlation = 0.21, $p < 10\text{-}10$, $t(78) = 8.41$) and pre-SMA (mean correlation = 0.24, $p < 10\text{-}10$, $t(79) = 9.02$, t-test versus 0) in pre-SMA.
(k) Induced power was significantly correlated with the iERN amplitude in the theta (5-10Hz) band in both dACC (mean correlation = 0.25, $p < 10\text{-}10$, $t(78) = 9.7$) and pre-SMA (mean correlation = 0.33, $p < 10\text{-}10$, $t(79) = 11.52$).
(l) Latency difference between pairs of iERNs recorded simultaneously in dACC and pre-SMA. The median latency difference of 18ms is significantly different from zero ($p < 10\text{-}5$, z value = 19.27, Wilcoxon's signed rank test).
(m) Amplitude difference between pairs of iERNs recorded simultaneously in dACC and pre-SMA. The median latency difference of 11 μV is significantly different from zero ($p < 10\text{-}5$, z value = 20.14, Wilcoxon's signed rank test).
(n) Comparisons of spike latencies and iERN latencies (replotting of data shown in main figure on different scale).
'*', '**', '***' mark statistical significance for $p \leq 0.05$, ≤ 0.01 and ≤ 0.001 respectively.

Figure S7

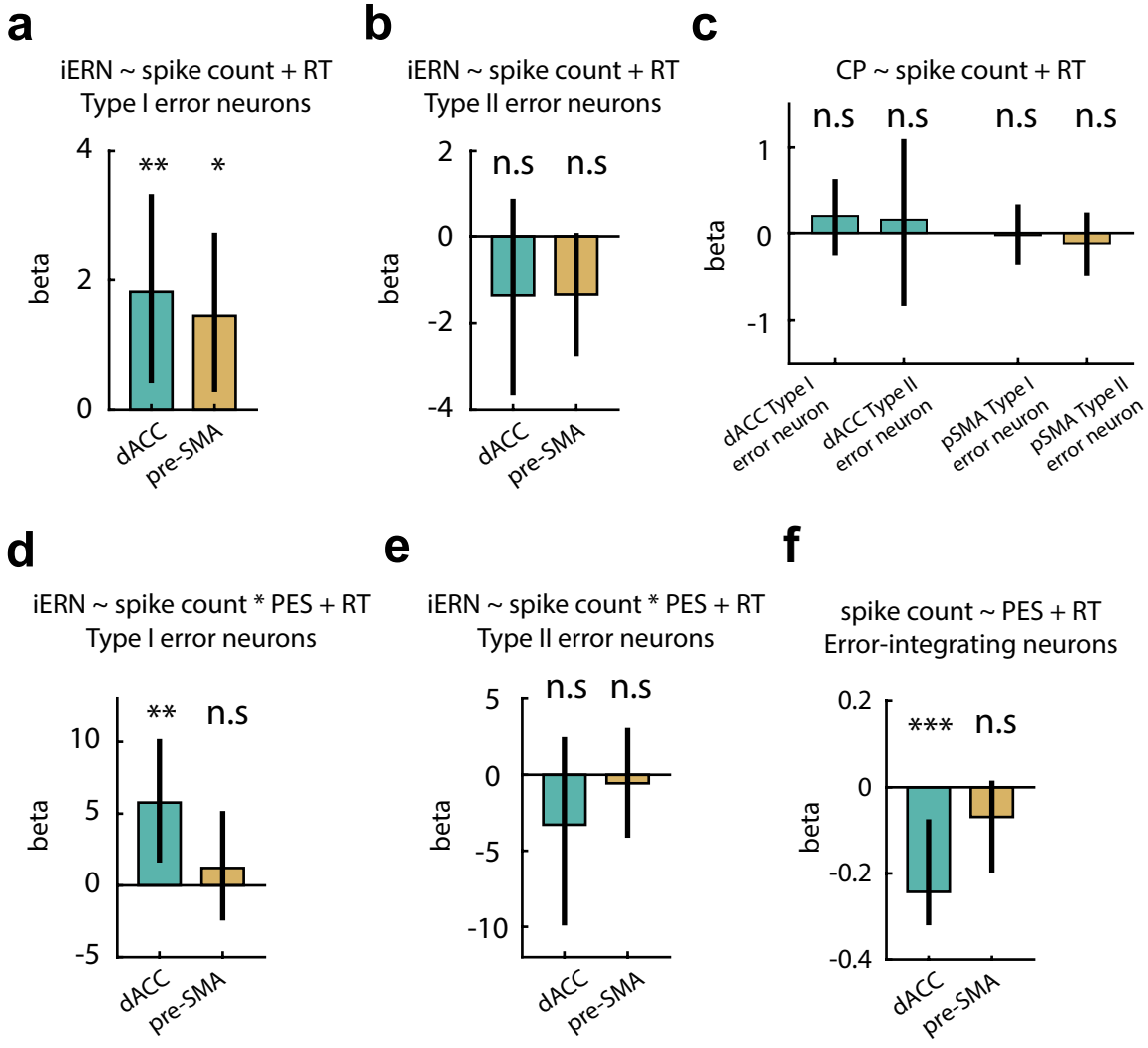

Figure S7. Regression coefficients of GLM models. Related to Figure 6-7.

(a) Regression coefficients for the fixed effect of spike rates (within [0 1s] after button press). The iERN amplitude was significantly correlated with spike of Type I error neurons in both dACC ($p = 0.008$, likelihood ratio = 6.56, likelihood ratio test) and pre-SMA ($p = 0.012$, likelihood ratio = 5.81, likelihood ratio test).
(b) Same as in (a) but for Type II error neurons. The iERN amplitude did not correlate significantly with spike rates (within [0 1s] after button press) in either dACC ($p = 0.19$, likelihood ratio = 1.64) nor pre-SMA ($p = 0.07$, likelihood ratio = 3.36, likelihood ratio test)
(c) Regression coefficients for the fixed effect of spike rates (within [0 1s] after button press). The CP ('correct potential') was not correlated with the spike rates of any types of error neurons in either dACC (for Type I error neurons, $p = 0.34$, likelihood ratio = 0.92; for Type II error neurons, $p = 0.74$, likelihood ratio = 0.11, likelihood ratio tests) or pre-SMA (for Type I error neurons, $p = 0.88$, likelihood ratio = 0.023; for Type II error neurons, $p = 0.48$, likelihood ratio = 0.49, likelihood ratio test).
(d) Regression coefficients of the interaction term between the spike rate (within [0 1s] after button press) of the Type I error neurons and PES levels. The correlation between iERN amplitude and the spike rates of the Type I error neurons in dACC was stronger when PES was larger ($p = 0.009$, likelihood ratio = 6.56, likelihood ratio test). The same relationship did not hold significantly for Type I error neurons in pre-SMA ($p = 0.5$, likelihood ratio = 0.44, likelihood ratio test).
(e) Same as in (d) but for Type II error neurons. The strength of correlation between iERN amplitude and the spike rates of the Type II error neurons did not vary significantly between PES levels in either dACC ($p = 0.67$, likelihood ratio = 0.18, likelihood ratio test) or pre-SMA ($p = 0.48$, likelihood ratio = 0.5, likelihood ratio test).
(f) Regression coefficients of the fixed effect of PES levels. The spike rates of the error-integrating neurons were strongly correlated with PES in dACC ($p < 0.001$, likelihood ratio = 15.76, likelihood ratio test), but only marginally so in pre-SMA ($p = 0.07$, likelihood ratio = 3.31, likelihood ratio test).
Error bars represent 95% confidence interval obtained from parametric bootstrapping. '*', '**', '***' mark statistical significance with p values ≤ 0.05, ≤ 0.01 and ≤ 0.001 respectively using the likelihood ratio test. 'n.s' marks not significant ($p > 0.05$).

Supplementary Tables

Table S1. Subjects recorded. Related to Figure 1.

List of all subjects recorded.

ID	Sex	Age	Epi Diagnosis	Macro recording performed	Sessions performed
H11	M	16	right lateral frontal	N	2
H14	M	31	Bilateral indep. Temporal	N	2
H16	F	34	right frontal	N	2
H17	M	19	left inferior frontal	N	3
H18	M	40	Right temporal	N	1
H19	M	34	Left frontal	N	1
H21	M	20	Not localized	N	2
H28	M	23	Right mesial temporal	N	1
H31	M	30	Right temporal	N	1
H41	M	19	Right posterior temporal	N	1
H42	M	29	Not localized	N	1
H49	F	54	Right amygdala and hippocampus	N	2
C24	F	47	Not localized	N	2
C25	F	36	Bilateral indep. Temporal	N	2
C26	F	56	Right temporal	N	1
C27	M	45	Left temporal	N	1
C29	M	19	Left temporal neocortical	N	4
C31	M	31	Left temporal neocortical	N	3
C32	M	19	Not localized	N	1
C33	F	44	Right temporal	N	4
C34	M	70	Bilateral temporal	N	5
C35	M	63	Left temporal neocortical	Y	6
C36	M	45	Right Hippocampus	Y	6
C37	F	33	Right Hippocampus	Y	11

C39	M	26	Right insula	Y	6
C40	M	25	Right motor cortex	Y	4
C42	F	25	Not localized	Y	5
C47	M	33	Right mesial temporal	Y	2
C48	F	32	Left medial temporal	Y	1

Table S2. Percentage and average spike rate of neurons. Related to Figures 2-3.

Summary of percentages and average spike rates (±s.d.) of neuronal categories. Neurons of the types marked as "NA" were found not more than expected by chance.

dACC	**Error neurons**		**Error-integrating neurons**	
Types	**I**	**II**	**I**	**II**
Spike rate (Hz, ±s.d)	2.61±2.8	3.24±2.2	2.77±2.1	NA
Percentage (%)	24.8	8.8	11.9	NA

pre-SMA	**Error neurons**		**Error-integrating neurons**	
Types	**I**	**II**	**I**	**II**
Spike rate (Hz, ±s.d)	2.47±2.3	3.6±3.3	NA	3.47±3.6
Percentage (%)	27.4	18.6	NA	13.6

Table S3. Percentage of overlap between neuron types. Related to Figures 3, S3

Summary of proportions of neurons in the overlap between error neurons and conflict neurons. The overlap was tested with Fisher's exact test. Example of how to read this table: the first entry below shows that "In dACC, out of all conflict neurons type I, 24.4% are also error neurons Type I".

dACC	Percentage in: Conflict neuron Type I	Percentage in: Conflict neuron Type II
Error neuron Type I	24.4%	23.3%

Error neuron Type II	8.6%	7%

dACC	Percentage in: Error neuron Type I	Percentage in: Error neuron Type II
Conflict neuron Type I	10%	7.3%
Conflict neuron Type II	10.1%	8.6%

dACC	Conflict neuron Type I	Conflict neuron Type II
Error neuron Type I	p = 0.18, odds ratio = 1.77	p = 0.19, odds ratio = 1.65
Error neuron Type II	p = 0.50, odds ratio = 1.34	p = 0.73, odds ratio = 1.27

Pre-SMA	Percentage in: Conflict neuron Type I
Error neuron Type I	27.8%
Error neuron Type II	20.4%

Pre-SMA	Percentage in: Error neuron Type I	Percentage in: Error neuron Type II
Conflict neuron Type I	12.7%	13.8%

Pre-SMA	Conflict neuron Type I
Error neuron Type I	p = 0.224, odds ratio = 1.48
Error neuron Type II	p = 0.22, odds ratio = 1.59