Supplemental Figures

Figure 1. Mapping of motor output driven by Arch and electrical stimulation of superficial cerebellar cortex. A) Schematics of the results of Arch stimulation over a grid of sites in lobule simplex (Sim) and Crus I for all five mice. Colors denote outcome of stimulation. B) Same as A for three mice tested with electrical microstimulation.

Figure 2. Retention and extinguishment of learning for cue prediction regime. A) Average reaction times relative to cue for the first (left) and second (right) sessions where animals performed 0.5 s cue prediction on consecutive training days (n=17 sessions, 5 mice). Error bars are ±SEM across sessions. B) Same as A for 2 s cue prediction sessions (n=10 sessions, 6 mice). C) *Left,* example cumulative distributions of reaction times for cue reaction sessions immediately surrounding a cue prediction session (session 0). Trials from the cue prediction session include only the last 1/3 of the session when the animal was anticipating the cue timing. *Right,* summary of distribution skewness across cue reaction sessions immediately surrounding cue prediction sessions (n=43 sessions). Error bars are SEM across sessions. D) Same as A for pairs of 0.5 s prediction sessions separated by 3 or more cue reaction sessions (n=20 sessions, 10 mice).

Figure 3. Performance in the cue prediction condition in control and NBQX sessions. A-E) Summary across experiments of mean percent correct lever releases, mean baseline reaction time from the visual cue, mean reaction time variance, mean duration to press, mean number of trials performed, and the mean duration of behavioral sessions. n=44 sessions, 0.5 s; n=9, 1.0 s; n=6, 1.5 s; n=30, 2.0 s; n=26, Δt; n=14, 0.5 s NBQX. Error bars are SEM across sessions. For each cue delay, the minimum and maximum (x,y) number of trials performed in single sessions was: 0.5 s: (309, 700) 1 s: (252, 900) 1.5 s: (202, 775) 2 s: (278,1185) Δt: (313, 772) 0.5 s NBQX: (227, 429). F_i) Individual (top) and average (bottom) lever trajectories aligned to lever press (left) and release(right) segregated by duration into quartiles (n=112 trials / quartile) for a representative 0.5 s cue prediction session. F_{ii}) Same as F_i for a representative 0.5 s cue prediction session in NBQX. (n=142 trials / quartile). G) Summary of press (left) and release (right) quartiles across experiments for control (top; n=15 sessions) and NBQX (bottom; n=14 sessions) experiments. H) Summary of the ratio between press and release ranges (difference between 1st and 4th quartile) for control (black) and NBQX (blue) experiments.

Figure 4. Pharmacological modulation of complex spiking in lobule simplex. A) Single unit recordings during NBQX application to LS. Top, summary of complex spike rates in control and after application of NBQX for individual cells (black; n=9 cells, 5 mice) and the average of all cells (red). Error bars are ±SEM across cells. Bottom, same as top for simple spike rates. Note that NBQX strongly reduces complex but not simple spike rates. B) (top) Raw epifluorescence images from parasagittal cerebellar sections showing fluorescein labeling after surface application to lobule simplex *in vivo* (methods, 3 mice). (bottom) Pixel masks from the same images above. Pixels were thresholded at 30% of maximum value to visualize and quantify fluorescein labeling. C) Summary of mediolateral fluorescein spread across sections. (n=3 animals). Error bars are SEM across animals.

Figure 5. Segmentation of PC dendrites from imaging experiments. A) Example single photon imaging field of view (left) is segmented to mask individual dendrites (right) used for subsequent single cell analysis. Scale bar = 100 μm. B) Example timecourse of raw fluorescence (top) from the circled dendrite in A showing individual calcium transients identified (gray circles) according to peaks in the first derivative of the fluorescence trace (bottom). C) Same as A for an example two photon field of view. Note that while PC somata are visible (round, top right), image segmentation does not extract activity from these structures. Scale bar = 100 μm. D) Same as B for the experiment in C. E) Top, example time course of a single unit PC recording from an awake mouse illustrating detection of complex spikes (gray circles). Bottom, overlay of individual simple (left) and complex (right) spike waveforms (gray) and the average waveform (black). F) Summary of mean spike rates ± SEM across all two photon dendrite imaging sessions (blue; n=1146 dendrites) and acute single unit recordings (n=11 units).

Figure 6. Single unit PC recordings of complex spiking during the cue prediction condition. A) Top, Schematic of the trial structure schematic, where a constant cue delay of 500ms was imposed on each trial. Middle top, Average simple (SS, left) and complex (CS, right) spike waveforms from an example single unit. Middle bottom, Histogram of SS firing rate aligned to complex spike time revealing post-CS pause, confirming isolation of single PC. Bottom, raster of single trial complex spikes and session PSTH aligned to lever release. B) Single trial voltage traces (complex spikes- open blue circle) from the cell in A aligned to lever release (correct- black circle; early- red circle). C) Expansion of 5 consecutive traces from B at time of lever release. D) Mean normalized complex spike rate (methods) from single unit recordings (31 PCs, 8 mice) aligned to lever release for correct (black) and early (red) trials during the cue prediction

condition. Shaded error is SEM across PCs. E). Mean normalized complex spike rates aligned to cue presentation for the first 1/3 of trials (black) and the last 1/3 of trials (orange) across cue prediction sessions for correct (left) and early (right) trials (31 PCs, 8 mice). Shaded error is SEM across PCs. F) Summary of the S.D. of spike times when aligned to either lever release or visual cue for each PC (n=31). P-values reflect paired t-tests.

Figure 7. Differential complex spiking on correct and early release trials is not driven by motor signals due to licking. A) Example imaging data from an inter-trial interval (ITI) showing a lick responsive dendrite (top, dark blue) and a lick unresponsive dendrite (bottom, black) in the same field of view. B) Averaged lick triggered calcium transients for the lick responsive dendrite (dark blue) and a lick unresponsive dendrite (black) in A). n=28 licks, Error bars are SEM across lick-triggered events. C) Mean lick triggered calcium transient across all lick responsive (dark blue) and lick unresponsive (black) dendrites. Shaded error is SEM across dendrites. D) Summary of the Spearman's correlation between lick rate and amplitude of the calcium transient in each frame across trials for lick responsive (dark blue) and lick unresponsive dendrites (grey). Red points indicate the mean $\pm$ SEM across dendrites (447 dendrites, 6 animals and 10 sessions). E) Summary comparing the amplitude of calcium transients for lick responsive neurons only in response to either a single lick (blue) or a lick bout (3 or more licks with $\leq$ 300 ms between licks; cyan). Note that additional licking does not produce larger responses. F) Summary comparing the amplitude of calcium transients on correct (black) and early (red) lever releases for lick unresponsive dendrites only. G) Same as F, for lick responsive dendrites. P-values reflect paired t-tests.

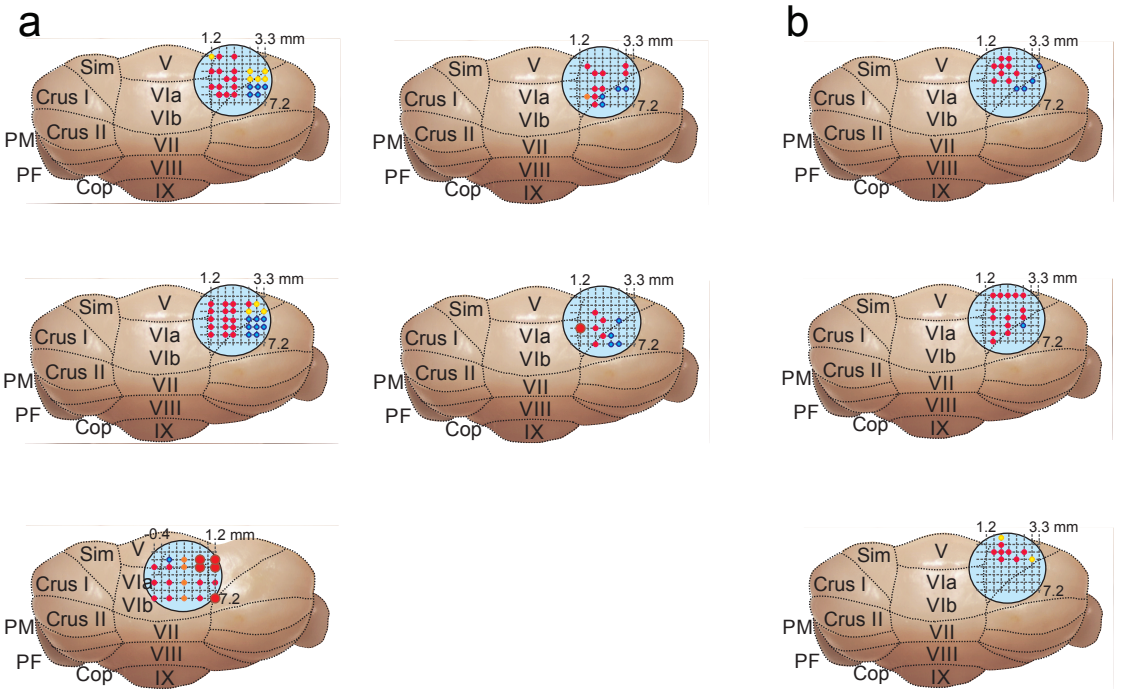
Figure 8. Example single trial and session average calcium transients from 2-photon imaging. A) Example single trial data from an individual dendrite showing the normalized calcium transient across an early release trial from before lever press to after the lever release. B) Same as A) but for a correct trial, including the time of cue presentation. C) Mean calcium transients across trials (n= 186) for the dendrite in A and B. Error bars are SEM across trials. D) Single trial example, aligned to lever release, showing all dendrites from the experiment illustrated in Fig. 5C on a correct trial over 1 second surrounding release (n=115 dendrites). Lever press is off scale. E) Same as D, but for an early release trial from the same imaging session.

Figure 9. Unexpected reward drives complex spiking in lick unresponsive and responsive dendrites. A-B) Top, Average calcium transient in response to unexpected reward (green,

aligned to first lick) and correct lever releases (black, aligned to release). Shaded error is SEM across dendrites. Bottom, Average lick rate for unexpected reward and correct lever releases. Error is SEM across experiments (n=3). Lick responsiveness was defined according to significant responses in the lick triggered averaged taken from the inter-trial interval (Supp. Fig. 7). C). Summary of peak calcium transients in the same neurons for correct lever releases vs unexpected reward trials. Note that responses are proportional, and response amplitude is determined by response probability (Fig. 5,6). P-value reflects paired t-test.

Figure 10. Complex spiking scales with lever hold time for early releases in single photon imaging experiments. Summary of peak calcium transients in a 500 ms window at the time of release (methods) across all single photon experiments for correct (black) and early (red) release trials binned according to hold time (250ms bins). Linear fits were applied to data from each trial type (n=10 animals, 17 sessions). Error bars are SEM across sessions.

Figure 11. Mice have full control over the lever trajectory. A) Individual (top, n=217) and average (bottom, 54 per quartile) lever trajectories aligned to lever press (left) and release (right) for manual lever depression followed by unperturbed gravity return. Shaded error is SEM across trials. B) Summary of press (left) and release (right) quartiles across experiments for manual lever experiments (n=4). Note the lack of variability across quartiles. C) Summary scatterplot comparing 20-80% rise times of lever press and release for experiments where the lever was controlled by mice (black, n=15) and manual press (green, n=4) D) Summary comparison measuring the average difference from the mean of the slowest release quartile and the fastest release quartile for experiments where the lever was controlled by mice (black) and manual press (green). Lever trajectories were sorted according to duration in a window from 200ms before threshold crossing to the time of a threshold crossing half way between the top and bottom of the total lever displacement. The slowest and fastest 25 trials were extracted for comparison, and their average rise time was normalized by subtracting that of the mean trajectory of the whole session. Error bars are SEM across sessions. These results show that mice produce movements both slower and faster than the mean, indicating full control over the lever trajectory.



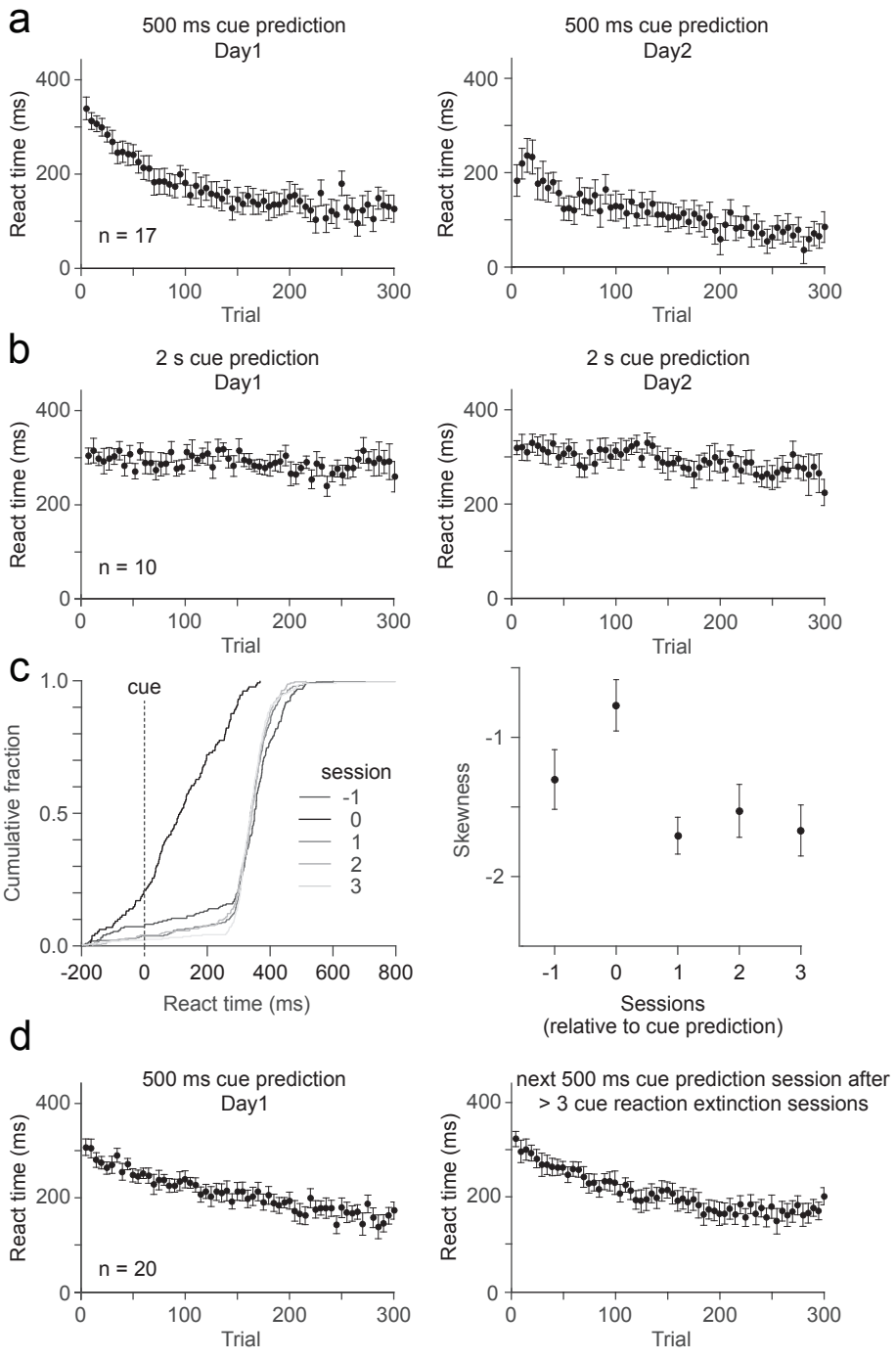
Arch: 5 mice
10-50mW, one Step pulse for 1sec

Right arm
Right & Left arms
Very weak arm movements
Whiskers or mouth or nose
nothing

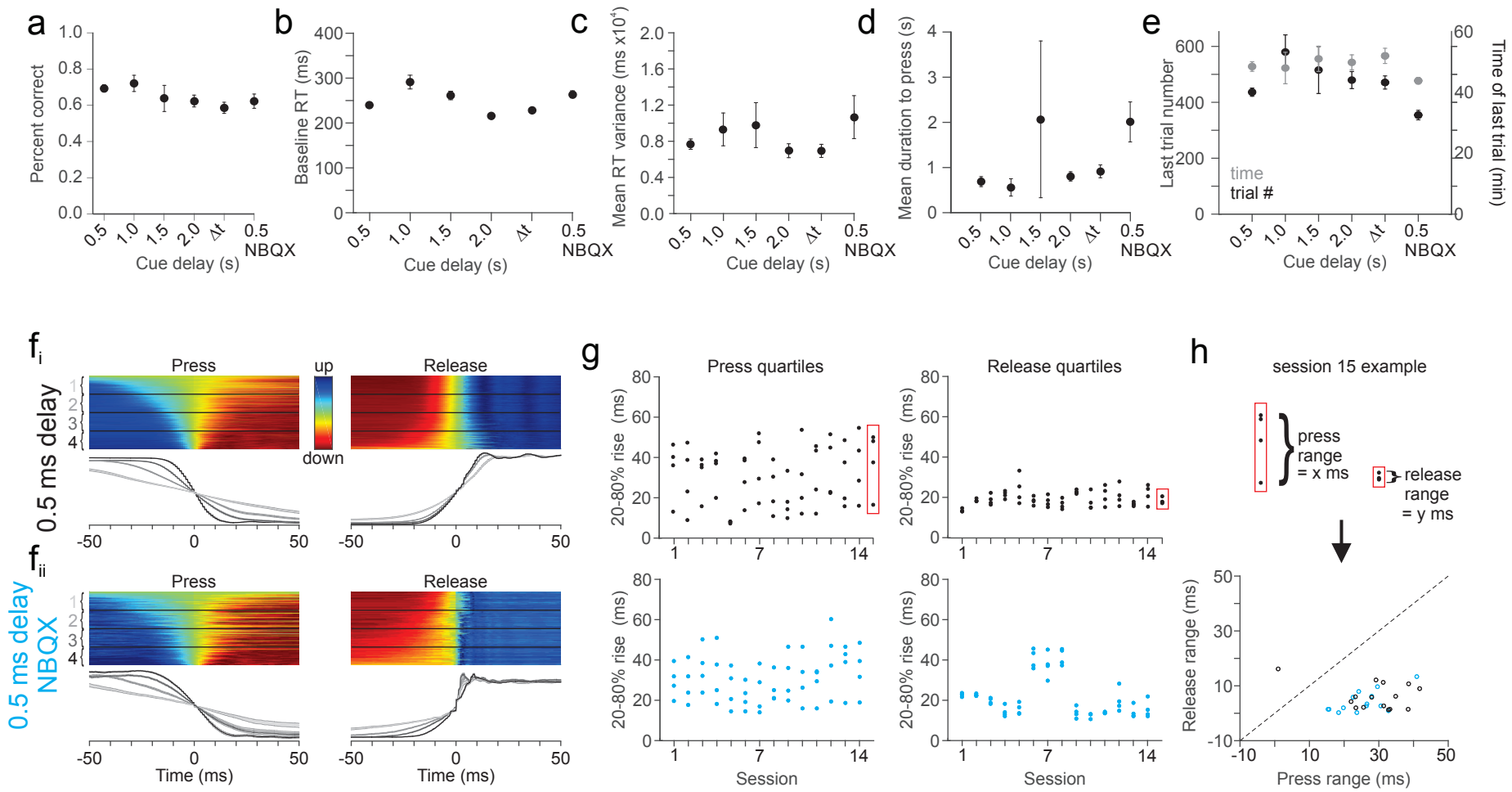
Electrical Stim. 3 mice
0.1mA-1mA
50Hz for 500ms or 1sec

Right arm
Right & Left arms
Very weak arm movements
Whiskers or mouth or nose
nothing

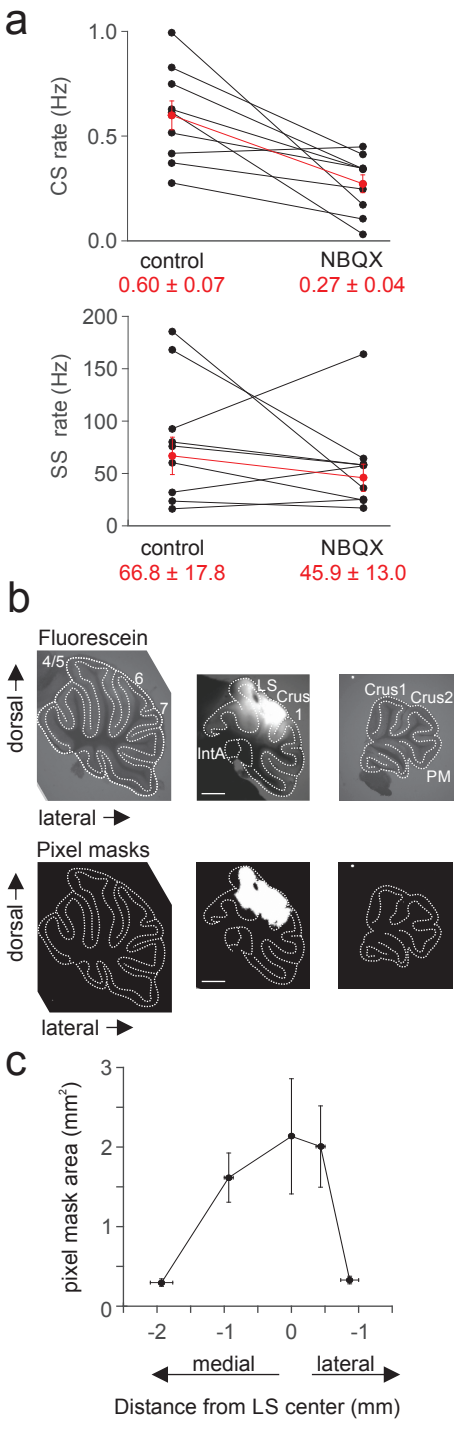
Supplemental Figure 1



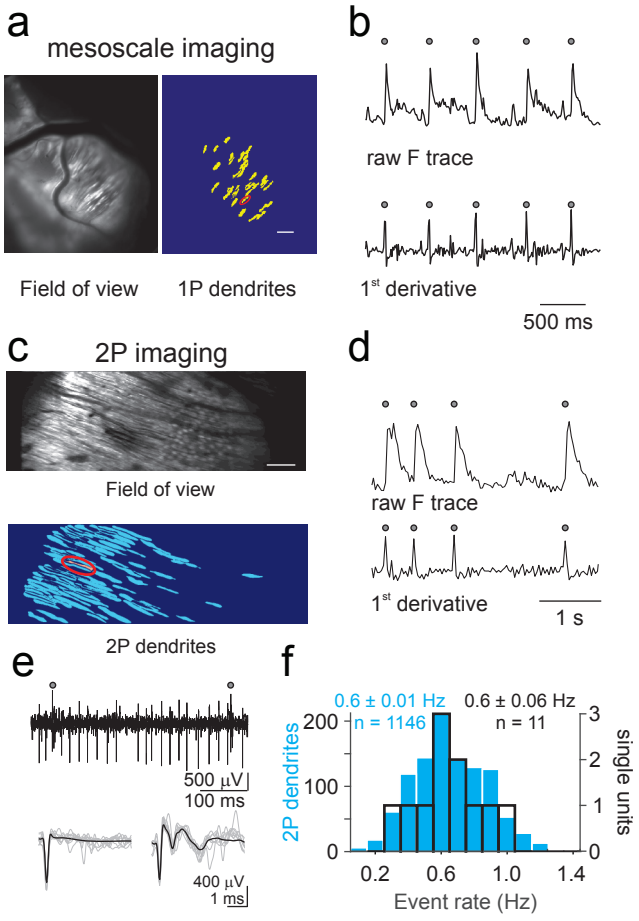
Supplemental Figure 2



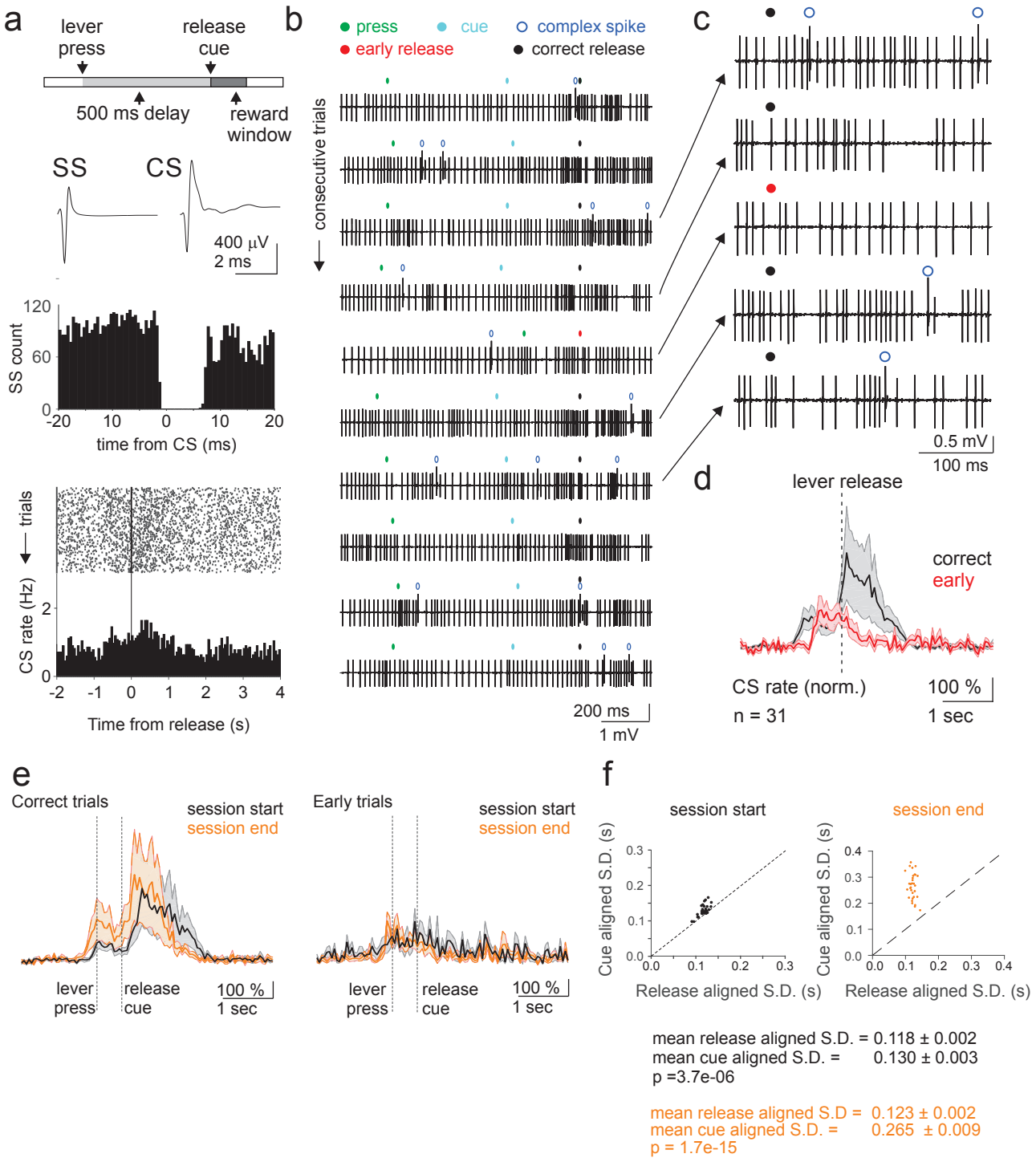
Supplemental Figure 3



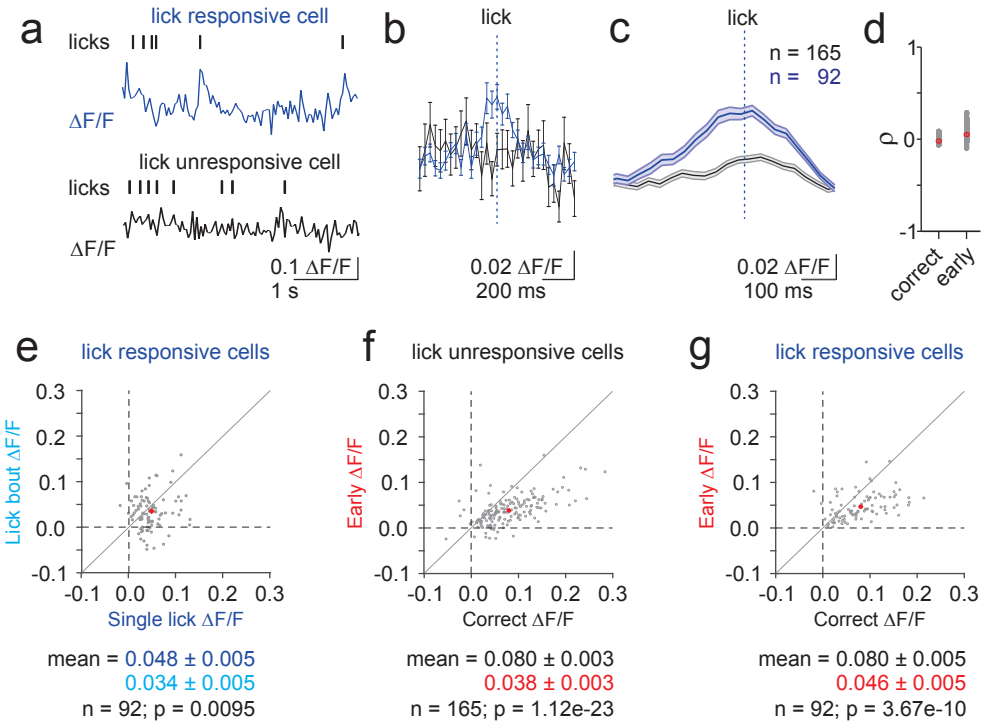
Supplemental Figure 4



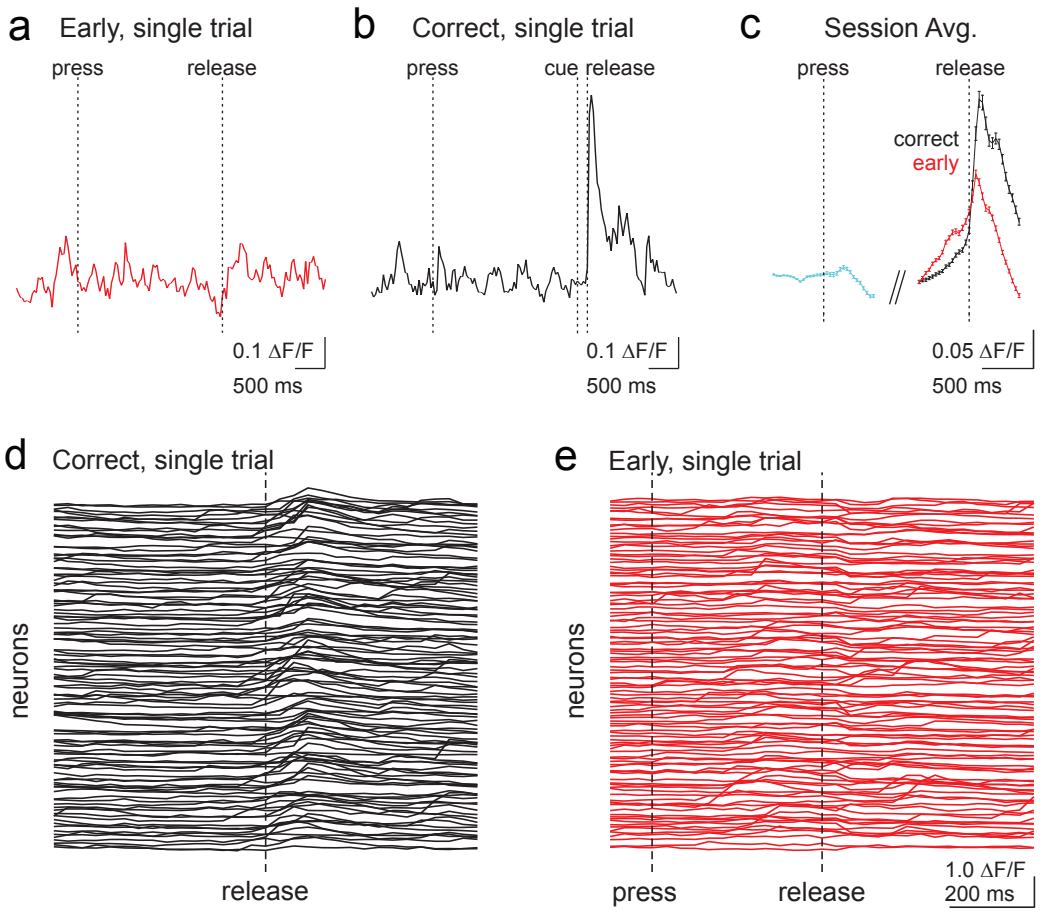
Supplemental Figure 5



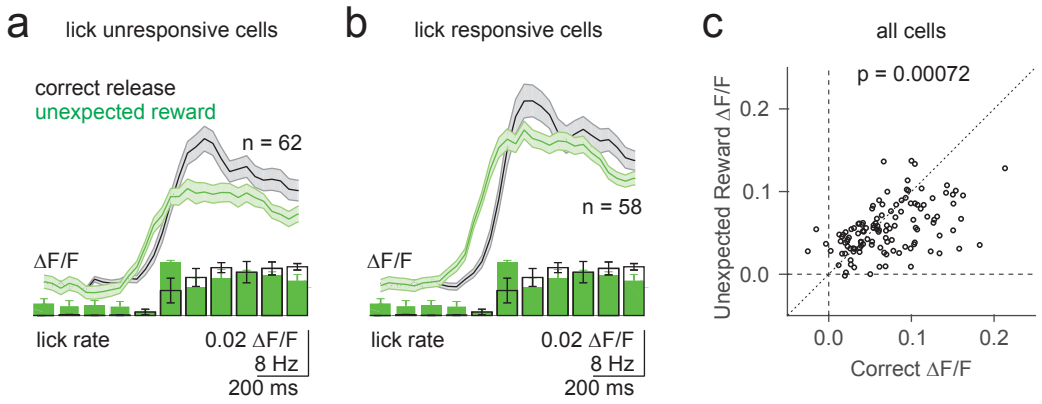
Supplemental Figure 6



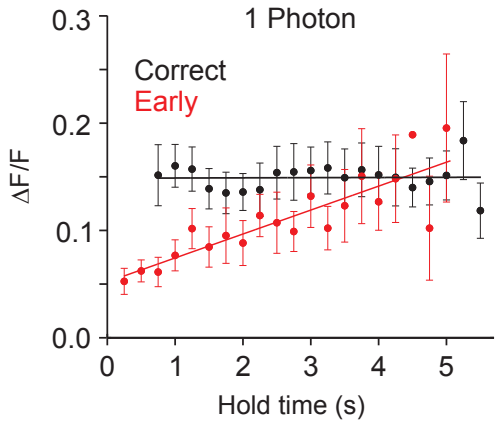
Supplemental Figure 7



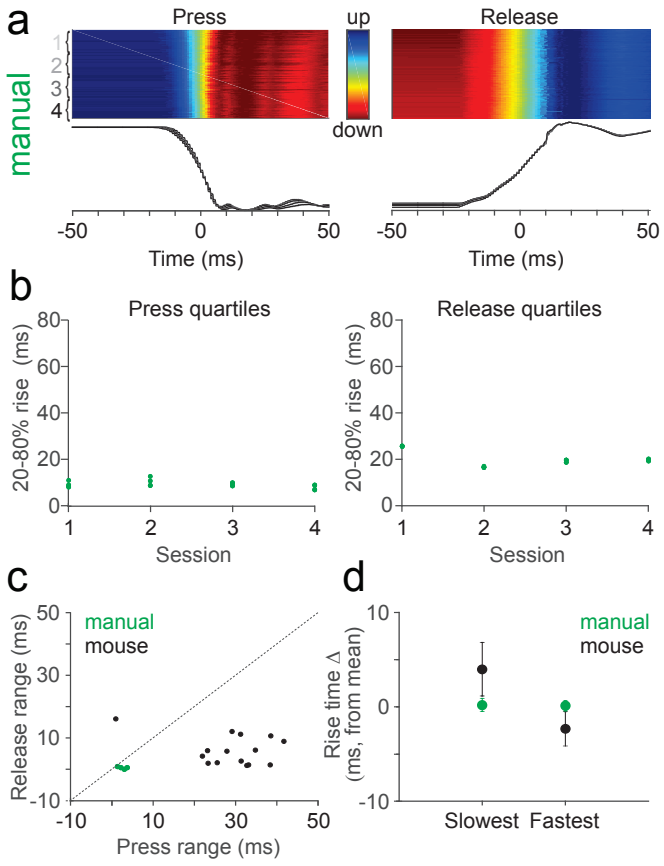
Supplemental Figure 8



Supplemental Figure 9



Supplemental Figure 10



Supplemental Figure 11