Mice preferentially use increases in cerebral cortex spike rates to detect changes in visual stimuli

Jackson J. Cone[1], Morgan L. Bade[1], Nicolas Y. Masse[1], Elizabeth A. Page[1], David J. Freedman[1], and John H.R. Maunsell[1*]

[1]Department of Neurobiology and Grossman Institute for Neuroscience, Quantitative Biology and Human Behavior, University of Chicago, Chicago IL, 60637 USA

***Correspondence:**
maunsell@uchicago.edu (JHRM)

Summary

When the retinal image changes, some neurons in primary visual cortex (V1) increase their rate of spiking, while other decrease their rate of spiking. How the brain decodes changes in visual stimuli from this set of neuronal responses remains largely unknown. We trained mice to respond to increases or decreases of visual contrast by releasing a lever. We used optogenetic approaches to increase or decrease V1 neuron responses evoked by contrast changes. The ability of mice to detect increases and decreases in visual contrast were both enhanced by increasing V1 spiking, whereas detection was always impaired by suppressing V1 spiking, even when the change was a decrease in contrast. The results suggest a strong asymmetry in the readout of signals from V1 insomuch as decrements in neuronal spiking contribute markedly less to perceptual reports.

Introduction

Artificial electrical activation of neurons in cerebral cortex can produce robust percepts in the absence of natural sensory input (Penfield and Rasmussen, 1950). Such direct activation of cerebral cortex with electrical microstimulation in alert subjects has greatly advanced our understanding of the functional organization of localized subsets of neurons and their contributions to perception (Bak et al., 1990; DeAngelis and Newsome, 2004; Doty, 1965; Gu et al., 2012; Kim et al., 2015; Romo et al., 2000; Salzman et al., 1992; Schmidt et al., 1996; Verhoef et al., 2012). With practice, subjects can learn to detect stimulation of virtually any part of their cerebral cortex (Doty, 1965, 1969; Murphey and Maunsell, 2007, 2008; Murphey et al., 2009), suggesting that changes in neuronal spiking arising anywhere in cerebral cortex can be used to guide behaviors.

However, the effects of electrical microstimulation on population responses, and therefore perception and behavior, are not straightforward. Electrical microstimulation increases the rate of action potentials initially, although strong stimulation can be quickly overwhelmed by long-lasting inhibition of spiking (Logothetis et al., 2010). Two-photon imaging of single-cell activity has shown that the population of activated cells is sparse and can extend for hundreds of microns (Histed et al., 2009). This is likely because axons passing near a microelectrode will be preferentially excited by electrical microstimulation compared to soma (Rattay, 1999). Furthermore, microstimulation can produce excitation as well as inhibition both at the stimulated site (Ferster and Chung, 1998) and in downstream areas (Logothetis et al., 2010). More recently, observations based on electrical stimulation have been confirmed and expanded in experiments using transcranial magnetic stimulation (TMS; Marg and Rudiak, 1994; Merabet et al., 2003) and, in animal subjects, with optogenetic stimulation (Histed and Maunsell, 2014; Huber et al., 2008; Jazayeri et al., 2012; O'Connor et al., 2013). Unlike electrical microstimulation or TMS, which can directly excite all neuronal cell types, optogenetic approaches enable inhibitory signals to be inserted into sensory populations either via direct inhibition of pyramidal neurons or by activating the inhibitory cortical network (Wiegert et al., 2017). Synchronous inhibition of sensory cortex impairs perception across a range of modalities and stimuli (Cone et al., 2019; Glickfeld et al., 2013; Guo et al., 2014; Resulaj et al., 2018). That electrical microstimulation can produce percepts, whereas optogenetic inhibition of cortical spiking suppresses perceptual reports suggests there may be an asymmetry in the ability of increments and decrements of cortical spiking to be used for guiding behavior.

Equivalent increments and decrements in neuronal spiking can in principle provide just as much information to downstream brain regions. However, we hypothesize that spike rate increments might dominate in terms of downstream readout, primarily because cortical neurons have relatively low baseline spike rates. Spikes are energetically expensive (Attwell and Laughlin, 2001; Lennie, 2003), which may favor neural encoding that uses fewer spikes. Such constraints might have driven the development of the retinal ON/OFF pathways, where increments and decrements in luminance are converted into parallel channels that increase firing to either change (Schiller, 1992). It

is therefore possible that many brain circuits work preferentially with spike increments because of the signal processing advantages conferred by such rectification. ON/OFF pathways use fewer spikes to convey the same amount of information compared to an ON only system (Gjorgjieva et al., 2014). Moreover, rectification is widely employed in engineered circuits; where it can greatly simplify signal processing (e.g., detection of amplitude-modulated radio signals), eliminate the need to track a zero-point, and speed decrement detection in systems with sparse signal rates.

To understand how changes in population spiking yield usable signals for downstream circuits, it is critical to identify the mechanisms that enable and limit the readout from cortex. Optogenetic methods provide an approach for producing controlled increments or decrements in neuronal firing. However, a direct comparison of the effects of spiking increments and decrements requires that spike rate changes of comparable magnitude be produced, which effectively means that such changes must be applied when neurons are responding to a stimulus. Additionally, subjects need to be encouraged to respond to both increments and decrements in signals. We have examined how increments and decrements in V1 spiking affect the ability of mice to detect bidirectional changes in visual contrast. We report here that even when increments and decrements in V1 output are put on an equal footing, decrements in spiking do not appear to contribute to perceptual reports.

Results

Mice were surgically implanted with a headpost and a cranial window to give stable optical access to V1 (Goldey et al., 2014). We used transgenic mouse lines that expressed Cre-recombinase selectively in one of three major subclasses of cortical neurons: excitatory neurons (Emx1, Emx), parvalbumin expressing neurons, PV, and somatostatin-expressing neurons, SST (Gorski et al., 2002; Hippenmeyer et al., 2005; Taniguchi et al., 2011). These strains allow selective targeting of excitatory opsins to the neurons of interest with >95% specificity (Madisen et al., 2012; Pfeffer et al., 2013). We targeted injections of Cre-dependent viruses containing ChR2-tdTomato to monocular V1 (described in subsequent sections). Electrophysiological and behavioral experiments were conducted following stable ChR2 expression (≥1 month post-injection).

Changes in visual contrast evoke increments and decrements in V1 spiking.

To document how V1 neurons respond to changes in visual contrast, we performed electrophysiological recordings from V1 in awake, head-fixed mice (n=8; 4 Emx, 4 PV). Two PV mice were first used in the behavioral experiments that are described later, while the others were prepared only for electrophysiological recordings.

A full-screen 50% contrast sinusoidal grating stimulus (0.1 cycles/degree, static, vertically oriented) was always present on the visual display except during contrast changes. We presented a range of randomly interleaved 500 ms contrast increments and decrements. Spikes were sorted offline and responses to different presentations of each contrast change were averaged. For each stimulus, we calculated the change in

spike rate (Δspikes/s) by subtracting the pre-stimulus firing rate (50-250 ms before stimulus onset) from the firing rate during the stimulus epoch (50-250 ms after stimulus onset).

V1 units had diverse responses to contrast changes. Some units increased their firing rates for either increases or decreases in contrast (Figure 1A), while others decreased their firing rates for contrast changes (Figure 1B). As our primary goal was to assess the total stimulus-evoked signal present in V1, we classified units as excited or inhibited if the average firing rate increased or decreased by 10% relative to baseline when the initial change in contrast either halved or doubled. Using this criterion, many units were excited by contrast decrements (117/250, 47%; Figure 1C) and many by contrast increments (109/250, 44%; Figure 1D). Smaller proportions of units were inhibited by decrements (62/250, 25%; Figure 1E) and by increments (81/250, 32%; Figure 1F). In addition to being less prevalent, inhibitory responses were also weaker compared to excitatory responses, especially for contrast decrements. The mixture of selectivity for increases and decreases in visual contrast is to be expected given the juxtaposition of ON/OFF receptive fields in early visual areas.

Given that we presented gratings of a single orientation and spatial frequency, the visual stimulus was suboptimal for most V1 neurons recorded, but the average population responses to both contrast decrements and increments were positive changes in firing rate (Figure 1G). The response was larger for contrast decrements than increments (Figure 1G). This observation is consistent with the relative contribution of the weak decrease in spiking observed in the subpopulation of units that were inhibited by contrast decrements in Figure 1E,F. Overall, these recordings show that changes in the contrast of a sustained stimulus drive a diverse set of responses in V1 units that could support contrast change detection, and that V1 is configured so that increases and decreases in stimulus contrast both produce a net increase in spiking (Figure 1G, S1).

Figure 1

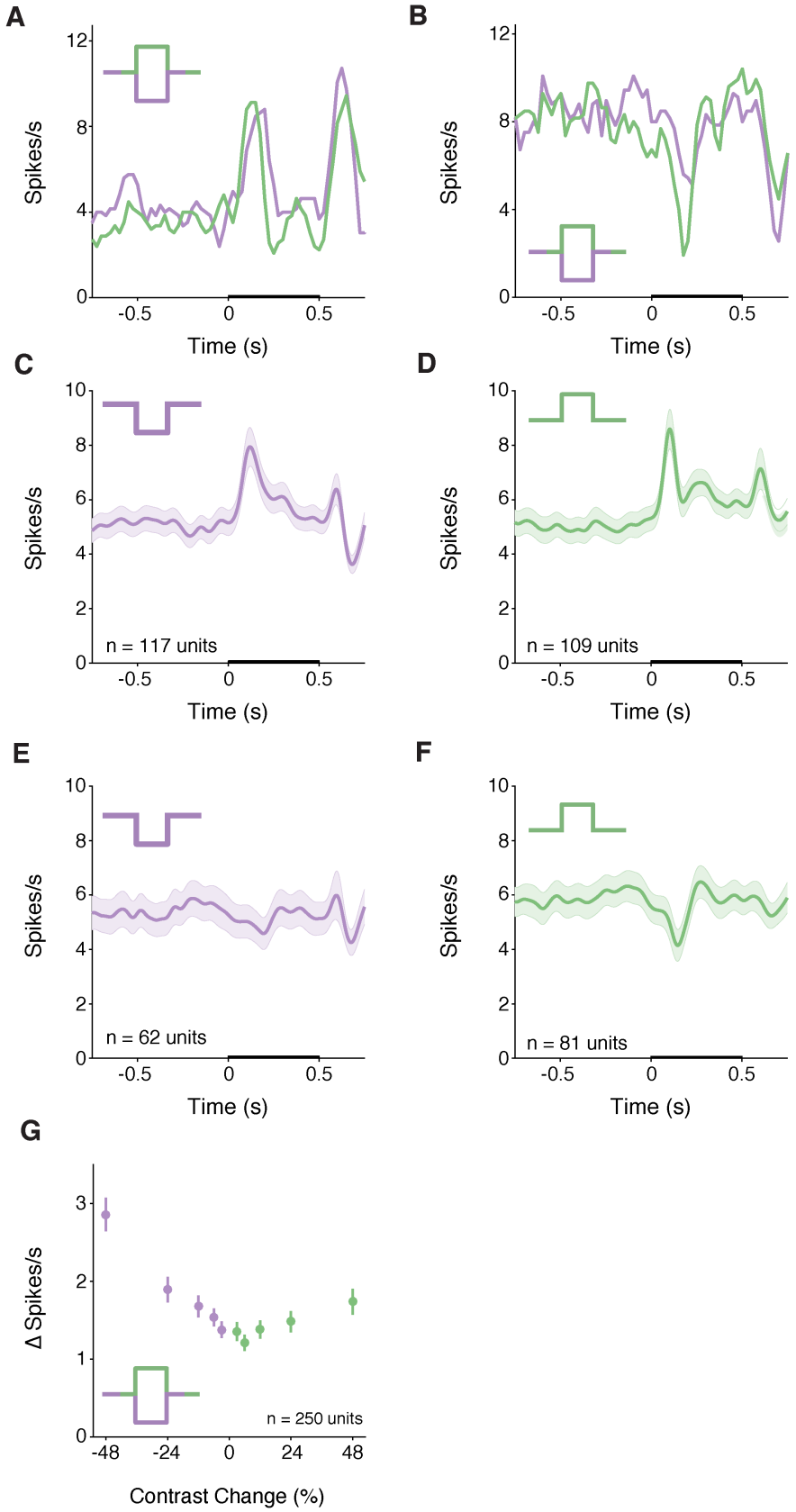
A
B
C
D
E
F
G
Spikes/s
Time (s)
n = 117 units
n = 109 units
n = 62 units
n = 81 units
Δ Spikes/s
n = 250 units
Contrast Change (%)

Figure 1. Changes in visual contrast evoke diverse responses in V1 units in awake, passively viewing mice. *A) Average contrast change responses from a representative unit that was excited by halving (decrements, purple) and doubling (increments; green) the stimulus contrast. Legend depicts the contrast profile. Visual stimulus duration is indicated by the thickening of the x-axis in all PSTHs. Bin size = 25 ms, smoothed. B) Same as in A except for a unit that was inhibited by contrast changes. C) Gaussian filtered (σ = 25 ms) PSTH for units that were excited by contrast decrements (n = 117 units). Shaded region = SEM. D) Same as in C except for units excited by contrast increments (n = 109 units). E-F) Same as in C, D except for units that were inhibited by contrast decrements (E, n = 62 units) or contrast increments (F, n = 81 units). G) Evoked change in firing rate relative to baseline across the population for all contrast changes. See also Figure S1.*

Visual and optogenetic stimuli

We compared how much V1 spike rate increases and decreases influence behavioral detection by using optogenetic methods to perturb V1 spiking. Mice were surgically prepared as described above. Following implantation, we mapped retinotopy in V1 using intrinsic signal imaging (Figure 2A). Imaging data was used to target injections of Cre-dependent viruses containing ChR2-tdTomato (Figure 2B,C; Nagel et al., 2003).

Mice were trained to perform a contrast change detection task while head fixed. In this task (Figure 2D), the mouse faced a video display containing static, monochromatic, and vertically oriented 50% contrast Gabor (centered at 20-25° azimuth, -15-15° elevation, 5-7° SD; 0.1 cycles/degree; odd-symmetric) on a mid-level gray background. To start a trial, the mouse depressed and held a lever through a randomly varying delay period (600-3000 ms) after which the contrast of the Gabor increased or decreased (randomly interleaved; 700 ms). The mouse had to release the lever within a 700 ms response window to receive a reward. We randomly varied the magnitude of the contrast change between trials using a range that spanned psychophysical detection threshold. An optical fiber was attached to the headpost to deliver optogenetic stimulation to a consistent cortical location each day (Figure 2B). The optical fiber was aligned with the retinotopic location of the visual stimulus representation in V1 by comparing images of virus expression and intrinsic signal imaging data (Figure 2A,B).

During each session, we delivered optogenetic stimulation on a random half of presentations of a single, near threshold, increment or decrement in contrast. The optogenetic stimulus was delayed by 35 ms relative to the onset of the visual stimulus to account for the arrival of spikes in V1. We restricted optogenetic perturbations to the stimulus epoch so as to only augment V1 spiking responses evoked by the visual stimulus. Moreover, delivering optogenetic stimulation throughout a trial can augment the signal to noise ratio against which sensory responses are processed (Aizenberg et al., 2015). Both the visual and optogenetic stimulus remained on until the end of the trial so as to prevent stimulus offsets from producing an additional signal that could drive behavioral responses.

Figure 2

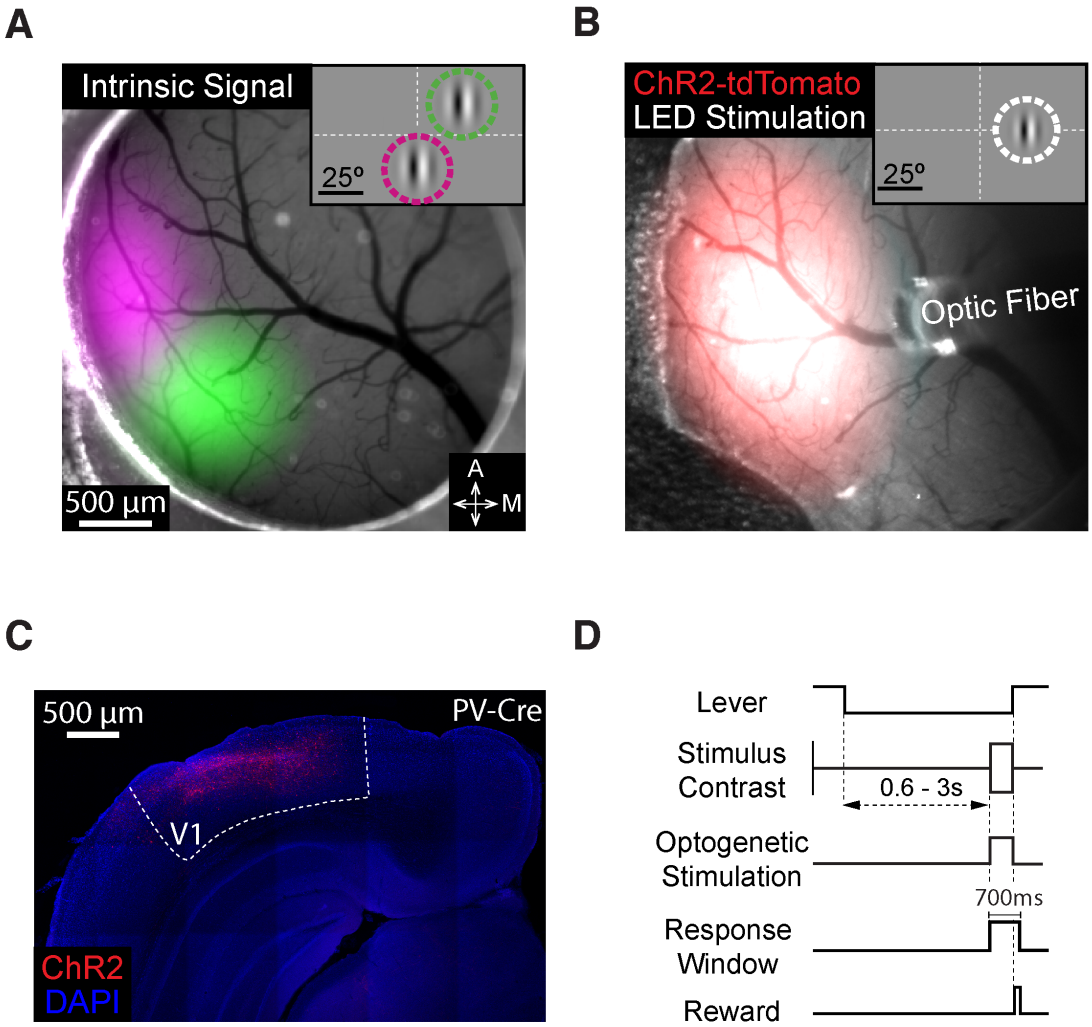

Figure 2. Targeting ChR2 to retinotopically defined areas of visual cortex. *A) Pseudo-colored intrinsic autofluorescence responses to visual stimuli presented in two locations in a PV-Cre mouse. Magenta and green features represent 2D-Gaussian fits of responses to stimuli at visual field locations depicted in the inset (magenta: 0° azimuth, -20° elevation; green: 25° azimuth, +20° elevation; Gabor SD = 10°). Dashed lines represent horizontal and vertical meridians. A: anterior; M: medial. B) ChR2-tdTomato fluorescence (2D-Gaussian fit) from the same cortical region shown in A. Area of LED illumination (2D-Gaussian fit) through the optic fiber positioned above ChR2-expressing V1. The retinotopic location corresponding to maximal expression was used in all behavioral sessions (shown in inset; 25° azimuth, 0° elevation; Gabor SD = 6.75°). C) Representative confocal image of ChR2-tdTomato expression in the visual cortex of a different PV-Cre mouse. D). Trial schematic of the contrast change detection task. A single increment and decrement in contrast was selected for stimulation (±15% contrast change for all mice).*

Increased spiking in pyramidal neurons facilitates detection of both contrast increments and decrements.

Emx mice detected ±15% contrast changes comparably well when those visual stimuli were presented without optogenetic stimulation (median percent correct, decrements 44%, 3% SEM, versus increments 42%, 2% SEM; 21 sessions in 3 mice; $p > 0.05$, Wilcoxon signed rank test). Optogenetic excitation of pyramidal neurons significantly increased the proportion of trials in which mice detected contrast decrements or increments. Figure 3A shows data from a representative session in which EMX neurons were activated during some trials on which the contrast increased or decreased by 15%. In either case detection was enhanced.

Improvements in detecting contrast increments and decrements were seen in virtually every session (Figure 3B; 21 sessions in 3 mice; decrements median stimulated = 66% [range: 46% – 92%], unstimulated 44% [range: 16%-81%], $p < 10^{-4}$; increments median stimulated = 68% [range: 44% – 97%] versus unstimulated 42% [range: 12%-58%], $p < 10^{-4}$; Wilcoxon signed rank tests). Activation of excitatory neurons significantly increased the proportion of hits in many individual sessions (contrast decrements: 9/21 sessions; contrast increments: 5/21 sessions; $p < 0.05$, both Fisher's exact test). Table S1 summarizes session averages for each mouse. Optogenetic activation of excitatory neurons also shortened reaction times on trials in which the animal correctly detected the stimulus compared to trials with no optogenetic stimulation (Figure S2A,B). In summary, optogenetic stimulation of excitatory neurons enhanced the ability of mice to detect changes in contrast regardless of the sign of the contrast change.

Optogenetic stimulation of pyramidal neurons increases neuronal responses to contrast increments and decrements.

To see how optogenetically-induced changes in detection performance were related to changes in V1 output, we recorded from awake, passively viewing mice expressing ChR2 in Emx-positive V1 neurons. We recorded from 119 units (including multi-units) across 12 V1 locations in four mice. We presented contrast increments and decrements and delivered optogenetic stimulation in conjunction with a random half of the presentations of 12% (contrast increment) and -12% (contrast decrement) stimuli. We used an optogenetic stimulation power that was comparable to the highest powers used in the behavioral experiments (0.3 mW) so as to best approximate an upper limit on the population effects. Visual and optogenetic stimuli were presented for 500 ms.

We quantified visual responses as the difference between the average firing rate for each unit during the 50-500 ms before stimulus onset compared to 50-500 ms after stimulus onset. Here, the analysis window was extended compared to Figure 1 to examine the full duration of the optogenetic stimulus on V1 spiking. Optogenetic effects were measured by comparing the stimulus evoked firing rates for 12% and -12% contrast changes with and without optogenetic stimulation. As expected, optogenetic activation of pyramidal neurons significantly changed the firing rate of many units (31%; 37/119, $p < 0.05$; Wilcoxon's signed rank test). Of the units significantly modulated by

optogenetic stimulation, the vast majority (95%, 35/37) had higher firing rates on trials with optogenetic stimulation compared to trials without stimulation (Figure S3A-D).

Large contrast decrements increased overall firing rates (black, Figure 3C,D; -48%: mean +1.6 spikes/s, 0.2 SEM). Moderate contrast decrements drove a weaker response (gray, Figure 3C,D; -12%: mean +0.7 spikes/s, 0.1 SEM). However, when moderate contrast decrements were paired with optogenetic activation of pyramidal neurons, the average response was comparable in magnitude to response for the largest contrast decrement (aqua, Figure 3C,D; -12%+Emx: mean +1.9 spikes/s, 0.2 SEM). Average responses differed significantly between stimulus conditions (Figure 3D; all comparisons at least $p < 0.05$; Friedman's test with Dunn–Šidák correction). Thus, optogenetic excitation of V1 Emx-positive neurons significantly increased spiking when paired with moderate contrast decrements.

V1 responses to contrast increments were weaker than those for decrements (Figure 3E,F, c.f. Fig 1G). Large (48%) contrast increments evoked an overall increase in firing across the V1 population (black, Figure 3E,F; 48%: mean +1.1 spikes/s, 0.2 SEM), while the moderate 12% contrast increment evoked a weaker response (gray, Figure 3E,F; 12%: mean +0.7 spikes/s, 0.1 SEM). As before, when a 12% contrast change was paired with optogenetic activation of pyramidal neurons, the population response was significantly enhanced (gold, Figure 3E,F; 12%+Emx: mean +1.8 spikes/s, 0.2 SEM; $p < 0.0001$ for 12% change with vs. without optogenetic stimulation; Friedman's test with Dunn–Šidák correction). Thus, as with contrast decrements, optogenetic excitation of V1 Emx neurons increased the population response to contrast increments.

The population response to the 12% contrast increments paired with pyramidal neuron stimulation (gold, Figure 3E,F) was larger than the response to 48% contrast increments (black, Figure 3E,F). During behavioral sessions, animals typically detected large contrast increments with greater frequency than trials with optogenetic stimulation (Figure 3A,B). This difference is likely due to our recordings being conducted in different animals outside of the behavioral task, using stimulation powers at the upper limit of those used during behavior.

Figure 3

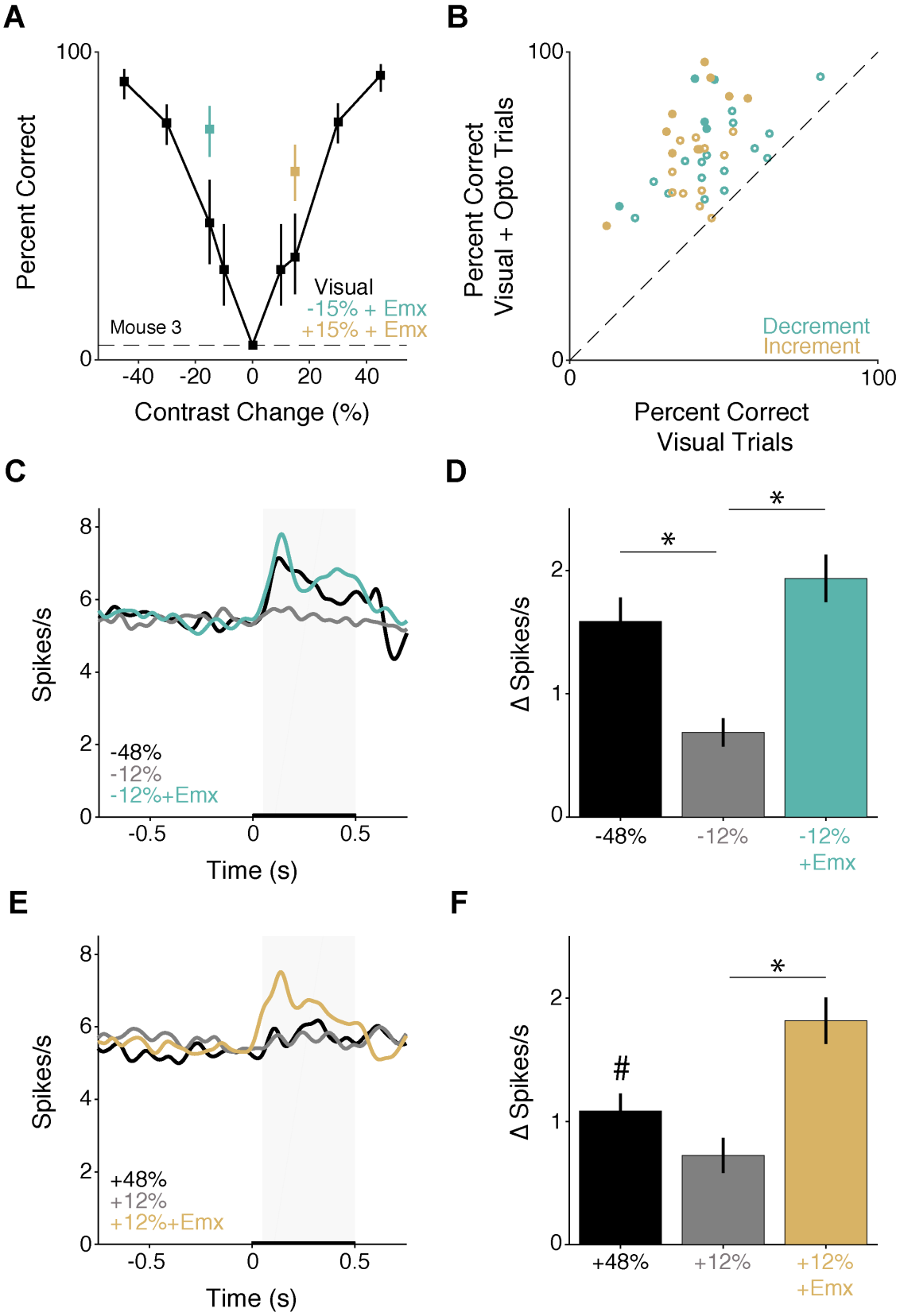
A
100
0
Percent Correct
Mouse 3
Visual
-15% + Emx
+15% + Emx
-40
-20
0
20
40
Contrast Change (%)
B
100
0
Percent Correct
Visual + Opto Trials
Decrement
Increment
0
100
Percent Correct
Visual Trials
C
8
6
4
2
0
Spikes/s
-48%
-12%
-12%+Emx
-0.5
0
0.5
Time (s)
D
*
*
2
1
0
Δ Spikes/s
-48%
-12%
-12%
+Emx
E
8
6
4
2
0
Spikes/s
+48%
+12%
+12%+Emx
-0.5
0
0.5
Time (s)
F
*
#
2
1
0
Δ Spikes/s
+48%
+12%
+12%
+Emx

Figure 3. Optogenetic excitation of visual responses in V1 similarly facilitates detection of increases and decreases in visual contrast. *A) Representative behavioral performance from a single session in an Emx mouse. Points and lines depict the percent correct ± 67% CI for trials without (black) optogenetic stimulation or contrast decrements (aqua) and increments (gold) paired with optogenetic of excitatory neurons. Dashed line = false alarm rate. B) Summary of stimulation effects in Emx mice. Circles depict the percent correct in individual behavioral sessions (3 mice, 21 sessions) with (y-axis) and without (x-axis) optogenetic stimulation, separately for increments (gold) and decrements (aqua). Filled circles = significant change in detection performance (14/42 observations, $p < 0.05$, Fisher's exact test). C) Population Gaussian filtered (σ = 25 ms) PSTH in response to large (black) and moderate contrast decrements without (gray) or with (aqua) optogenetic stimulation of pyramidal neurons in Emx mice (n = 119 units). Thickening of the x-axis represents duration of visual and optogenetic stimuli. Gray box = analysis window (+50 - +500 ms) used for spike rate quantification in D. D) Average change in spike rate compared to the time matched baseline period. $^{*}p < 0.0001$ relative to -12% change; Friedman's test with Dunn–Šidák correction. E) Same as in C except for contrast increments. F) Quantification of spike rate changes evoked by contrast increments with and without optogenetic stimulation. $^{\#}p < 0.1$ for 48% compared to both 12% contrast change without (gray) and with (gold) optogenetic stimulation. $^{*}p < 0.0001$ for 12% change with vs. without optogenetic stimulation; Friedman's test with Dunn–Šidák correction. See also Table S1, Figures S2, S3, S4.*

Optogenetic stimulation of PV interneurons impairs detection of contrast increments and decrements.

The above data show that potentiating spiking responses in V1 facilitates behavioral detection of both decrements and increments in contrast. Previously, we showed that detection of contrast increments is impaired by PV neuron stimulation (Cone et al., 2019), but contrast decrements were not examined in that study. The detection of contrast decrements might be mediated, at least in part, by monitoring the spiking of neurons whose firing rates decrease. We tested this by preparing and training a cohort of PV-Cre mice (n=6) to detect interleaved, bidirectional changes in contrast as above. Prior work has shown that activation of PV interneurons in V1 inhibits visually evoked neuronal responses (Atallah et al., 2012; Glickfeld et al., 2013; Wilson et al., 2012). As above, we delivered optogenetic stimulation concurrently with a single, near threshold, increment or decrement in contrast.

In the absence of optogenetic stimulation, animals had comparable levels of detection performance for ±15% contrast changes without optogenetic stimulation (median across 47 sessions, decrements 54% correct, 2% SEM versus increments 59%, 3% SEM; 47 sessions in 6 mice; $p > 0.05$, Wilcoxon signed rank test). Optogenetically stimulating PV interneurons produced behavioral effects that were opposite to those observed in Emx mice. Figure 4A shows data from a representative session in which PV interneurons were activated during some trials on which the contrast increased or decreased by 15%. In either case detection was impaired.

Impairments in detecting increments and decrements were seen in virtually every session (Figure 4B: decrements median stimulated percent correct = 24% [range: 3% – 56%] versus unstimulated 54% [range: 29%-82%], $p < 10^{-8}$; increments median stimulated = 29% [range: 0 – 66.7%] versus unstimulated 59% [range: 18%-90%], $p < 10^{-8}$; both Wilcoxon signed rank test; Figure 4A,B). PV interneuron stimulation significantly decreased the proportion of hits in most individual sessions (decrements: 30/47 sessions; increments: 27/47 sessions; $p < 0.05$, both Fisher's exact test; Figure 4B). Table S2 summarizes session averages for each mouse. PV stimulation also affected reaction times, generally slowing responses (Figure S2C,D). These data argue strongly against the possibility that the facilitation of detection observed in Emx mice was due to the mice seeing scattered light from the optogenetic stimulus.

To further explore these effects, we performed several additional control experiments. As expected based on prior work (Glickfeld et al., 2013), the effect of PV stimulation depended on retinotopic alignment between the visual stimulus and optogenetic manipulations. Moving the optical stimulation from the center of the representation of the visual stimulus in V1 reduced the change in performance (Figure S4E). Contrast change detection was also impaired when those changes were presented on the background of a counterphase modulated Gabor stimulus (Figure S5; Table S3), suggesting that the main effects did not depend on the adaptation of neuronal responses to the static Gabor. Thus, in stark contrast to the results of stimulation in Emx mice, optogenetic stimulation of PV interneurons consistently impairs contrast change perception regardless of sign.

Optogenetic stimulation of PV interneurons suppresses the population response to contrast increments and decrements.

We recorded V1 responses from awake, passively viewing mice with ChR2 expressed in V1 PV interneurons. We recorded from 131 units (including multi-units) from 12 sites in four mice. Consistent with the strong effect of PV neurons on cortical spiking, most recorded units (65%; 86/131) were significantly modulated by optogenetic stimulation ($p < 0.05$; Wilcoxon's signed rank test comparing evoked responses with and without optogenetic stimulation). Of these units, almost all (93%, 80/86) had lower firing rates on trials with optogenetic stimulation compared to trials without stimulation, as expected for activation of inhibitory interneurons (Figure S3E-H).

As before, we compared the change in firing rate evoked by large contrast changes (±48%), moderate contrast changes (±12%), and moderate contrast changes with concurrent PV stimulation (±12% + PV; Figure 4C-F). Large contrast decrements evoked a robust increase in firing rate across the population (black, Figure 4C,D; -48%: mean +1.9 spikes/s, 0.2 SEM), whereas moderate decrements evoked a weaker response (gray, Figure 4C,D; -12%: mean +1.2 spikes/s, 0.1 SEM). Pairing moderate decrements with optogenetic activation of PV interneurons robustly decreased V1 output compared to pre-stimulus firing rates (aqua, Figure 4C,D; -12% + PV: mean -2.2 spikes/s, 0.5 SEM). Average responses differed significantly between stimulus conditions (Figure 4D; all comparisons $p < 0.05$; Friedman's test with Dunn–Šidák

correction). Thus, PV stimulation produced a change in V1 output that was comparable in magnitude, but opposite in sign, to the response to large contrast changes.

Responses to contrast increments were weaker compared to decrements, though still above baseline firing rates. Large contrast increments evoked an increase in V1 output (black, Figure 4E,F; 48%: mean +1.3 spikes/s, 0.2 SEM) and a moderate change evoked a smaller response (gray, Figure 4E,F; 12%: mean +0.7 spikes/s, 0.1 SEM). Pairing a moderate contrast increment with optogenetic activation of PV interneurons strongly and significantly suppressed V1 output compared to pre-stimulus firing rates (gold, Figure 4E,F; 12% + PV: mean -2.4 spikes/s, 0.5 SEM; 12% change with optogenetic stimulation versus other conditions; both $p < 10^{-8}$; Friedman's test with Dunn–Šidák correction).

We obtained similar results when we stimulated PV interneurons while mice viewed contrast changes presented on a counterphase modulated Gabor (Figure S5E,F). In additional experiments, we tested different optogenetic stimulation powers and found that increasing power monotonically increased effects on both V1 spiking and behavioral performance (Figure S4A-D).

Figure 4

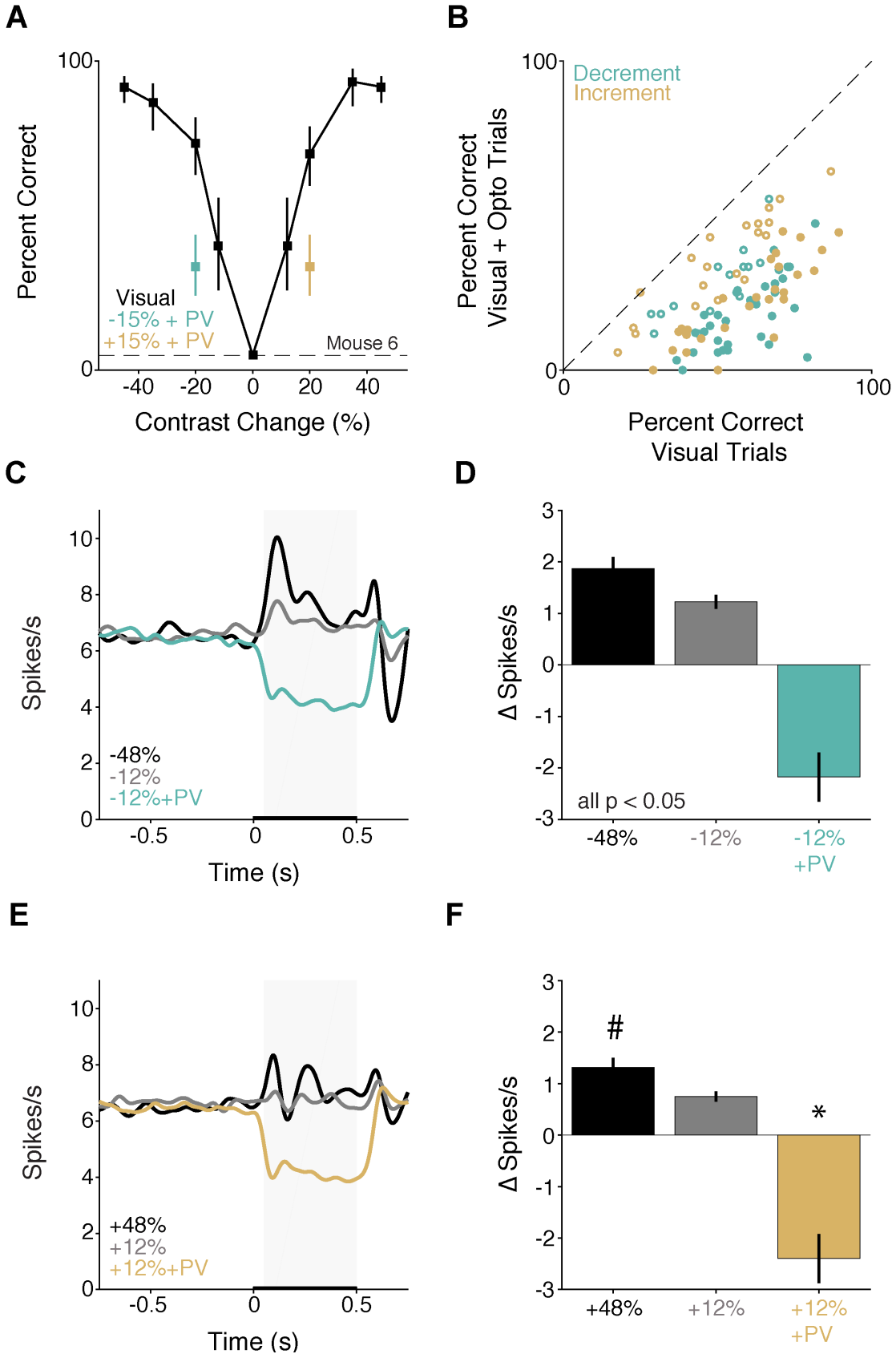
A
100
0
Percent Correct
Visual
-15% + PV
+15% + PV
Mouse 6
-40
-20
0
20
40
Contrast Change (%)
B
100
0
Decrement
Increment
Percent Correct
Visual + Opto Trials
0
100
Percent Correct
Visual Trials
C
Spikes/s
10
8
6
4
2
0
-48%
-12%
-12%+PV
-0.5
0
0.5
Time (s)
D
Δ Spikes/s
3
2
1
0
-1
-2
-3
all p < 0.05
-48%
-12%
-12%
+PV
E
Spikes/s
10
8
6
4
2
0
+48%
+12%
+12%+PV
-0.5
0
0.5
Time (s)
F
Δ Spikes/s
3
2
1
0
-1
-2
-3
#
*
+48%
+12%
+12%
+PV

Figure 4. Optogenetic suppression of visual responses in V1 similarly impairs detection of increases and decreases in visual contrast. *A) Representative behavioral performance (± 67% CI) from a single session in a PV mouse. Conventions are the same as in Figure 3. B) Summary of stimulation effects in PV mice. Circles depict the percent correct observed in individual behavioral sessions (6 mice, 47 sessions) with (y-axis) and without (x-axis) optogenetic stimulation, separately for increments (gold) and decrements (aqua). Filled circles indicate a significant change in detection performance (57/92 observations, both increments and decrements; $p < 0.05$, Fisher's exact test). C) Population Gaussian filtered (σ = 25 ms) PSTH in response to large (black) and moderate contrast decrements without (gray) or with (aqua) optogenetic stimulation of PV interneurons in PV mice (n = 131 units). Thickening of the x-axis represents the duration of the visual and optogenetic stimuli. Gray box = analysis window (+50 - +500 ms) used for spike rate quantification in D. D) Average change in spike rate compared to the time matched baseline period (All comparisons at least $p < 0.05$; Friedman's test with Dunn–Šidák correction). E) Same as in C except for contrast increments. F) Average change in spike rate evoked by contrast increments with and without optogenetic stimulation of PV interneurons. $^{\#}p = 0.07$ for 48% compared to 12% contrast change without (gray) optogenetic stimulation. $^{*}p < 10^{-8}$ for +12% change with optogenetic stimulation (gold) compared to large (+48%) and moderate (+12%) contrast increments. See also Tables S2, S3 and Figures S2, S3, S4.*

PV or SST interneuron stimulation impairs detection of brief contrast decrements.

Our electrophysiological recordings demonstrate that suppression of firing in response to visual contrast changes is typically transient (~50-200 ms, Figure 1B,F), whereas in the optogenetic stimulus used in our main behavioral experiments (above) persisted for 700 ms. We wanted to test whether brief activation of inhibitory neurons could facilitate performance using activation that more closely mirrors visually evoked changes in neuronal spiking. In these experiments, we made trials with optogenetic stimulation far more frequent to encourage the mice to exploit the decrements in V1 spiking in detecting stimulation. Additionally, we included animals expressing ChR2 in SST neurons to directly compare optogenetic stimulation of different classes of inhibitory interneurons. PV and SST interneurons each suppress V1 responses (Wilson et al., 2012), but they act through distinct synaptic mechanisms (Kubota et al., 2016).

We collected new data from PV (n=2; 1 female) and SST (n=2, both male) mice that were part of a previously published manuscript (Cone et al., 2019). For the new dataset, mice did a contrast decrement detection task. A 75%-contrast vertically oriented Gabor stimulus (0.1 cycles/deg, centered at 20-25° azimuth, -15-15° elevation, 12° SD, odd-symmetric) was always present on the video display except when its contrast transiently decreased to a lower value (100 ms, Figure 5A). Contrast decrement values spanned psychophysical threshold, and on a randomly selected half of trials, were synchronous with 100 ms illumination of ChR2-expressing PV or SST neurons. We then fit psychometric functions separately to performance values for trials with and without stimulation (see Methods).

Without optogenetic perturbations, there were no detected differences in contrast decrement detection performance (PV median threshold 11%; SST median threshold 11%; p = 0.89; Wilcoxon rank-sum test). PV or SST activation during contrast decrements impaired detection, shifting the psychometric functions to the right (Figure 5B,C). With the illumination powers used, PV or SST stimulation elevated detection thresholds approximately two-fold (PV 2.4-fold, SEM 0.1, range 1.3–4.2; SST 2.0-fold, SEM 0.1, range 1.3–3.0). This effect is the same as was found previously for contrast increment detection (Cone et al., 2019). Across all sessions (4 mice, 38 sessions), contrast decrement detection thresholds were significantly greater when the visual stimulus was paired with either PV or SST activation (Figure 5D, medians = 23% versus 11%, $p < 10^{-7}$; Wilcoxon signed rank test). This effect was significant for both genotypes individually (PV medians = 23% versus 11%, $p < 10^{-4}$; SST medians = 21% versus 11%, $p < 10^{-3}$; Wilcoxon signed rank tests), and in most individual sessions (PV: 20/20; SST 17/18). Combining our previously published data with the current results, PV or SST stimulation never facilitated detection performance across 125 increment or decrement detection sessions (n=8 mice; 5 PV, 3 SST).

Figure 5

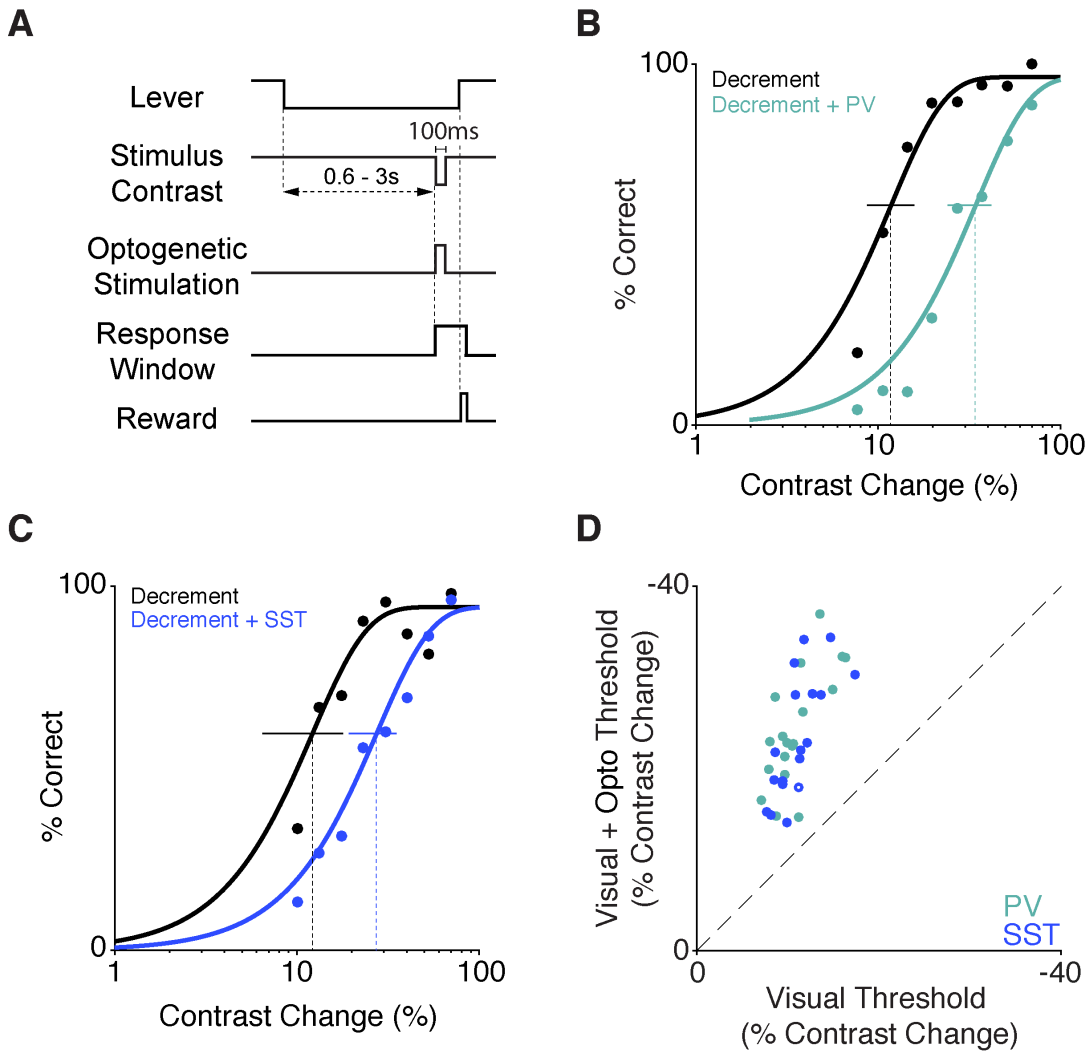

Figure 5. Brief optogenetic stimulation of either PV or SST neurons similarly increases contrast decrement detection thresholds. *A) Trial schematic of the contrast decrement detection task. The visual stimulus contrast is fixed at 75% except during a 100 ms decrement. ChR2-expressing interneurons were illuminated with blue light for 100 ms concurrent with the contrast decrement on a randomly selected half of the trials. B) Representative single session performance in the contrast decrement task for a PV mouse. Dots represent false-alarm corrected performance for trials with (teal) and without (black) activation of PV interneurons. Curves are best fitting Weibull functions that were used to determine detection thresholds (dotted vertical lines) and 95% confidence intervals (solid horizontal lines). C) Same as in B but for an SST mouse. Trials with SST stimulation are depicted in blue. D) Summary of PV and SST stimulation effects in the contrast decrement task. Circles depict the contrast decrement detection threshold from individual sessions with (y-axis) and without (x-axis) PV (teal, 2 mice, 20 sessions) or SST (blue, 2 mice, 18 sessions) stimulation. Filled circles represent sessions with a significant shift in threshold (37/38; bootstrapped). See also Figure S4.*

Different behavioral consequences of optogenetic increments and decrements of excitatory neuron spiking.

The above data suggest that mice cannot detect decrements in spiking in V1. However, this may be because spiking increments and decrements were not tested on equal footing, as the low baseline spike rates typically observed in cortical neurons limits the dynamic range available for decrements to signal information. We wondered if mice might exploit spiking decrements if asked to detect changes in sustained, elevated baseline levels of V1 spiking. To examine this, we trained a new cohort of Emx mice on a modified visual detection task (Figure 6A, 3 mice; 50 sessions) that incorporated optogenetic stimulation that persisted through the random delay period. When the visual stimulus changed contrast, the optogenetic stimulus either increased or decreased to potentiate or reduce its input into V1. In addition to examining changes in optogenetic stimulation from a sustained baseline, this design allowed us to examine sensitivity to V1 spiking increments and decrements in the same animals using perturbation of the same circuit (as opposed to effects in different Cre lines).

In all cases, the sustained optogenetic stimulus was relatively modest (0.10-0.25 mW) and stepped up or down by an amount comparable to the baseline power. The optogenetic stimulus step was brief (75 ms (19 sessions) or 150 ms (31 sessions)), while the contrast change persisted until the end of the trial (900 ms). On a subset of trials, the optogenetic stimulus incremented or decremented after the random delay while the contrast of the visual stimulus did not change. On these trials, the animal was rewarded for responding to increases or decreases in the optogenetic stimulus.

We were primarily interested in the behavioral effects on trials where the visual contrast did not change. Data from trials with concurrent optogenetic and visual stimulation are summarized in Figure S6. Mice did not respond reliably to steps in the optogenetic stimulus in the absence of visual stimulus changes, but were more likely to respond to

optogenetic increments compared to decrements. Figure 6B shows the session-average probability of lever releases for optogenetic stimulus increases and decreases on trials without visual stimulus changes. The task design encouraged animals to set a liberal response criterion, so many releases were false alarms, but there were far more releases on trials when the optogenetic stimulus increased (increment response rate = 38%, 35-40 95% CI; decrement response rate = 29%, 27-31 95% CI; t-statistic = 8.1, $p < 10^{-15}$, binomial logistic regression). Moreover when mice responded to optogenetic stimuli, reaction times were consistently faster for increments than decrements (Figure S6; increments median: 357 ms, 241 – 585 IQR; decrements median: 467 ms, 257-685 IQR; $p < 0.001$; Kruskal-Wallis test).

Because the animals were operating with a high false alarm rate, we could examine how increases or decreases in the optogenetic stimulus affected the probability of lever releases. Figure 6C plots probabilities of lever release following increments (gold) or decrements (aqua) in optogenetic stimulation across all 50 sessions (both 75 ms and 150 ms step durations), together with trial-time-matched false alarm probability (see Methods). The probability of lever responses diverged from the false alarm rate starting 200-300 ms, increasing when the optogenetic stimulus power increased (gold) and decreasing when the stimulus power decreased (aqua). This shows that mice are less likely to report a stimulus change when V1 spiking rate drops.

We performed electrophysiological recordings to measure the effects of this stimulation on V1 spiking (n=3 Emx mice; 5 recording sites). For electrophysiological recordings, we increased the optogenetic step duration to 250 ms to more easily quantify changes in firing. Most V1 units were sensitive to changes in optogenetic input (Figure 6D-F; 62%, 31/50; $p < 0.05$; Wilcoxon signed-rank on firing rates during the optogenetic step relative to baseline). For the subpopulation of optogenetically-modulated units, firing rates closely followed the optogenetic stimulation (Figure 6D,E). We quantified the evoked change in spike rate during the change in optogenetic input relative to baseline. The firing rate change evoked by decrements and increments in optogenetic input followed the step size (Figure 6F; both $p < 10^{-4}$; Kruskal-Wallis test). For smaller step sizes like those used in the behavioral experiments (±0.1 - ±0.2 mW), the spike rate changes for decrements and increments were comparable.

While the effects of optogenetic stimulation through a cranial window are strongest near the cortical surface, we identified ChR2-responsive units throughout cortex (Figure S3I,J) indicating the behavioral consequences of our optogenetic manipulations are unlikely to be restricted exclusively to effects in superficial cortical layers. Together, our observations strongly suggest that mice preferentially rely on increments in V1 output to detect changes in visual stimuli.

Figure 6

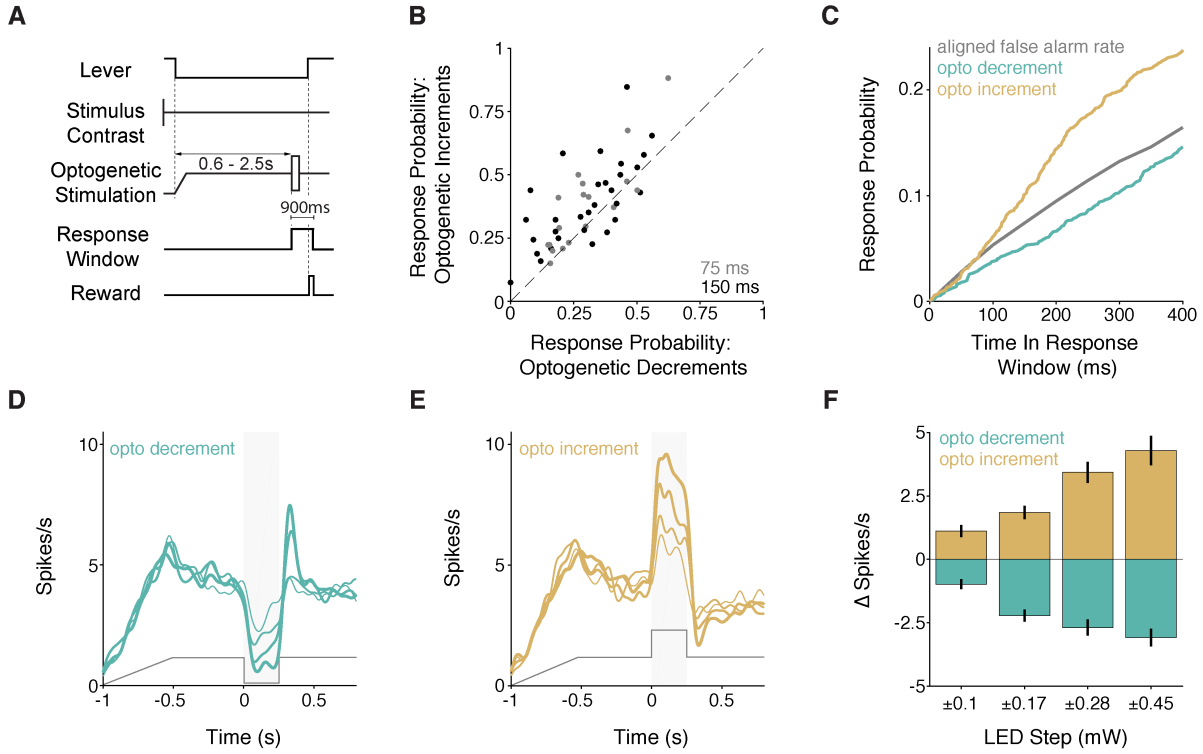

Figure 6. Optogenetically incrementing or decrementing excitatory input into the V1 population asymmetrically effects the probability of behavioral responses. *A) Trial schematic. The visual contrast was fixed at 50%, except during contrast decrements. At trial onset, the LED power ramped up and then held at a fixed value. Concurrently with contrast decrements (or after the random delay on 0% contrast change trials), the LED power briefly (75 or 150 ms) stepped up or down. B) Scatter plot depicts % correct for trials with optogenetic decrements (x-axis) and increments (y-axis) measured in the same session (n=3 mice; 50 sessions). The visual contrast did not change on these trials. Gray points = 75 ms optogenetic step duration; Black = 150 ms optogenetic step duration. C) Time course of changes in the lever release probability following optogenetic decrements (aqua) or increments (gold) compared to a time-in-trial matched false alarm rate (gray). (D-F) Electrophysiological recordings confirm V1 spiking follows complex optogenetic input. D) Average PSTHs for optogenetically responsive units (n = 31 units) in response to different magnitude decrements in optogenetic input. Increasing line thickness corresponds to the magnitude of optogenetic decrement (step sizes listed in F). Gray trace depicts the profile of optogenetic stimulation. E) Same as in D, except for increments in optogenetic input (step sizes listed in F). F) Quantification of V1 spike rate changes in response to decrements (aqua) and increments (gold) in optogenetic input. The magnitude of spike rate change significantly depends on the step size (increments and decrements, both $p < 10^{-4}$; Kruskal-Wallis test). See also Figure S6.*

Spike costs force recurrent neural networks to preferentially rely on increments in spiking to detect changes in contrast.

Preferential readout of spike rate increases compared to decreases might result from constraints that impact the robustness of particular decoding strategies (Gjorgjieva et al., 2014). One potentially strong constraint is the low baseline firing rates of cortical neurons. This limits the dynamic range available for decrements in spike rates to render information to downstream areas and how quickly changes can be detected. As this question of how baseline spiking level influences decoding strategies adopted by neural systems is inaccessible in biological networks, we turned to artificial recurrent neural networks (RNNs).

We trained RNNs (n=50) to perform a contrast change detection task similar to that which we used in mice. RNNs were rewarded for detecting brief decrements and increments in contrast that occurred against an otherwise static baseline contrast (Figure 7A, randomly interleaved). The RNN architecture is described in Methods. Briefly, visual input was passed to a layer of 24 contrast and orientation sensitive units that were either excited or inhibited by visual stimuli (50/50 split). The contrast responsive layer projected to a recurrent layer of 100 units (80 excitatory, 20 inhibitory) and the excitatory units linearly projected onto an output unit that signaled network responses. Because it is likely that the metabolic cost of neuronal activity contributes to the low baseline spike rates in cortical networks (Attwell and Laughlin, 2001; Lennie, 2003), networks were trained using different activity costs. We used rate-based RNNs rather than spiking networks, as recent modeling work suggests that the average firing rate, rather than temporal patterning of spikes, determines the metabolic cost of neuronal activity (Yi and Grill, 2019). This approach allowed us to explore the relationship between activity cost, overall activity, and the decoding strategies used by RNNs to perform the task.

Activity costs profoundly affected the average firing rate observed in RNNs. As activity costs increased, networks adopted significantly lower average rates of firing (Figure 7B; Spearman ρ = -0.96, $p < 10^{-27}$). Consistent with the idea that lower levels of activity impact the dynamic range available for decrements in spiking to signal stimulus changes, the proportion of positive output weights from the recurrent layer depended strongly on activity cost. High activity costs reliably caused RNNs to use exclusively positive weights (Figure 7C; Spearman ρ = 0.79, $p = 10^{-10}$). This demonstrates that the metabolic cost of activity can reduce firing rates and shift networks toward using firing rate increments.

We next used receiver-operating-characteristic (ROC) measures (see Methods) to explore how changes in RNN activity related to behavioral responses. For each network, we calculated a population ROC based on responses to contrast changes. A value of -1 indicates that all units in a network fired less when RNNs correctly detected contrast changes (hit) compared to trials in which the network failed not respond (miss). Conversely, a value of 1 indicates that activity was higher for all units in the network when the network responded. The normalized ROC value converged to 1 as activity

costs increased (Figure 7D; Spearman $\rho = 0.85$, $p < 10^{-14}$). This shows that behavioral responses to increments and decrements in contrast were correlated with firing rate increases for high activity costs in our RNNs.

To examine the causal role for increases and decreases in activity rate in guiding responses, we presented new trials to trained networks but turned off outputs from the recurrent layer that were either positive or negative. To align with our neurophysiological data in V1, this meant that decreases in spiking still exist and can influence activity within the recurrent layer, but only positive or negative outputs from the recurrent layer to the response neuron could impact its responses. Removing negative outputs devastated performance when activity costs were zero, but removing contribution of negative output weights had almost no impact on performance once activity costs became appreciable (Figure 7E; Spearman $\rho = 0.77$, $p < 10^{-10}$). Networks had modest performance when activity costs were low and they were constrained to use positive output weights (Figure 7E). However, networks constrained to negative weights performed poorly, even with low activity costs (Figure S7). These data show that activity costs in artificial neural networks can produce decoding strategies that are consistent with those observed in the mouse: Increments in neuronal activity appear to be preferentially used for detecting changes in visual stimuli.

Figure 7

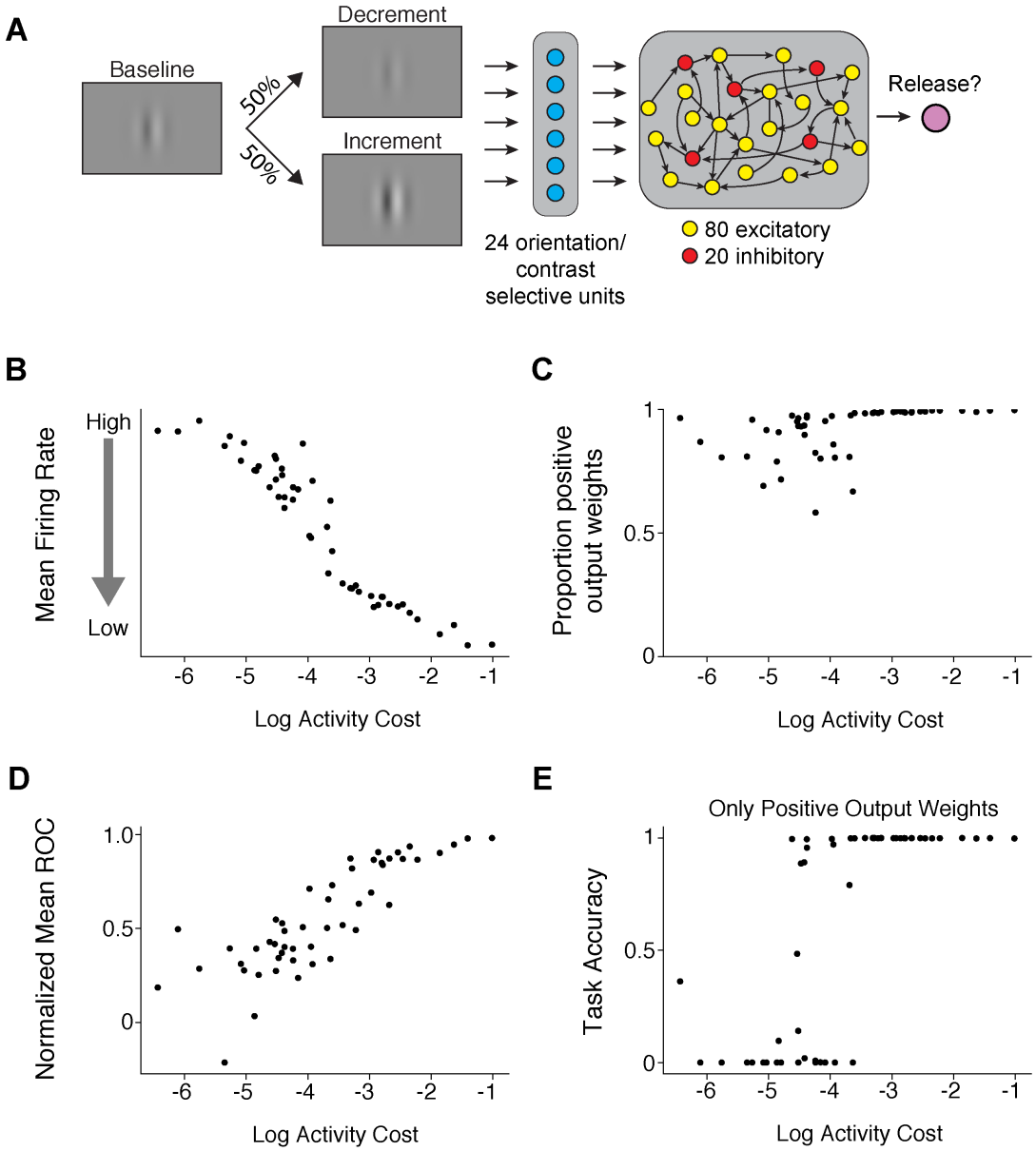

Figure 7. Activity costs force recurrent neural networks to adopt low firing rates and ignore firing rate decrements. *A) Schematic of a network trained to detect interleaved increments and decrements in contrast. The input is a single oriented Gabor that increases or decreases in contrast, which is passed to 24 units that are tuned to orientation and contrast (12 excited or inhibited by contrast changes). The contrast responsive layer passes input to a recurrent layer with 100 interconnected units (80 excitatory/20 inhibitory) that project to a single output node that controls lever releases. B) Increasing activity costs (x-axis) drive recurrent networks to lower basal rates of firing*

(y-axis). Points represent the results obtained from different networks trained with different activity costs. C) The proportion of output weights from the recurrent layer that are positive increases as function of activity cost. D) Normalized ROC (mean of increments and decrements). Range extends from -1 (all units decrease in firing) to 1 (all units increase). E) Task Accuracy as a function of activity cost when all units with negative weights are removed. See also Figure S7.

Discussion

How do downstream brain regions decode changes in V1 spiking? Our data suggest that decreases in the output of V1 neurons are not used in supporting the detection of visual stimuli. Subpopulations of V1 units have firing rates that either increase or decrease in response to contrast changes, likely owing to the ON/OFF juxtaposition of early stage visual receptive fields. This demonstrates that either increases or decreases (or both) in neuronal firing could inform decisions about visual stimuli (Figures 1, S1). Optogenetic excitation of primary neurons facilitated contrast change detection for both increases and decreases in contrast (Figures 3, S2), whereas optogenetic manipulations that reduce the spiking of V1 neurons impaired detection of contrast change of either sign (Figures 4, 5, S2, S4, S5). When mice were trained to respond to optogenetically-induced increments and decrements of V1 spike rates, increments increased their response rates, while decrements lowered their probability of responding (Figures 6, S6). The perceptual effects produced by optogenetic stimulation persisted across testing, suggesting that downstream structures cannot readily learn to decode decreases in V1 spiking.

The brain might not process reductions in cortical spiking because low baseline firing rates in cortex impacts the coding range for decrements. Decrements in low spike rates also require relatively long periods to detect. Low baseline firing rates encourage RNNs to rely exclusively on increases in activity to detect contrast changes (Figures 7, S7). In sum, the current results suggest that increments in the output of V1 neurons are preferentially used by downstream circuits to detect changes in visual contrast.

Implications for signal readout in V1

How do optogenetic perturbations interact with visual signals in a V1 neuronal population? We previously showed that the effects of PV or SST stimulation on contrast detection were better explained by a divisive scaling of the stimulus contrast compared to models in which stimulation acted directly on the probability of lever releases (Cone et al., 2019). This argues strongly against the possibility that interneuron activation impairs performance by disrupting motor planning or produces a signal that distracts the animal. The effects of optogenetic manipulations that increase V1 output might correspondingly increase the apparent contrast of visual stimuli. However, activation of excitatory neurons or vasoactive intestinal peptide expressing (VIP) interneurons can produce visual percepts in the absence of any natural visual stimulus (Cone et al., 2019; Histed and Maunsell, 2014). Thus, optogenetic activation of excitatory neurons when a visual stimulus is present could either enhance stimulus representations, or could

produce a signal that is distinct from the visual signal. Work from others suggests that artificial and natural signals may be merged into a common percept (see Histed et al., 2013). For example, when electrical microstimulation is delivered to direction selective neurons in the middle temporal area (MT) while monkeys view motion stimuli, the animals' reports resemble a vector average of the motion stimulus and the preferred direction of the stimulated column (Nichols and Newsome, 2002). Regardless of the exact mechanism by which potentiating excitation facilitates performance in our experiments, the conclusion is unchanged: adding signal that the brain can decode can increase the probability of detection.

Primates can learn to report electrical microstimulation of all cortical areas tested (Doty, 1965, 1969; Murphey and Maunsell, 2007, 2008; Murphey et al., 2009), suggesting that animals can detect activity arising in any part of cortex. If mice could decode decreases in V1 output, we might expect the effects of PV or SST stimulation to change over time as mice learn to use the resultant signal to guide responses. However, we previously showed that PV or SST activation consistently elevated contrast increment detection thresholds over many weeks of testing (Cone et al., 2019). Here, we retrained a subset of these same animals to detect contrast decrements and found that PV or SST activation still produced a perceptual impairment (Figure 5). Testing spanned weeks (max 21 sessions, ~4500 stimulation trials). For comparison, mice learn to report optogenetic activation of VIP interneurons with far less total exposure (9-15 sessions; ~1800-3000 stimulation trials; (Cone et al., 2019). Other PV mice were tested with different optogenetic stimulus powers or stimulus configurations and testing spanned 20-30 (max = 32) sessions, yet we never observed an increase in detection probability. Despite thousands of exposures to the population level consequences of PV or SST stimulation, mice do not learn to use this signal to guide their responses. Conversely, mice can perceive optogenetic activation of V1 excitatory neurons and the change in spike rate produced near the detection threshold is low (Δ1.1 spikes/unit; Histed and Maunsell, 2014). Our PV manipulations produced spike rate changes that were similar in magnitude but opposite in sign (Figures 4, S3, S4, S5). These data support the idea that only increases in V1 spike rate are decoded by downstream structures.

Why would cortex preferentially decode from increments in neuronal spiking?

Excitation and inhibition are tightly coupled in cortical circuits (Moore et al., 2018; Okun and Lampl, 2008; Wehr and Zador, 2003; Xue et al., 2014). Strong inhibition keeps baseline firing rates low, which limits the dynamic range available for decreases in firing rate to encode information. In the visual system, natural scenes generate sparse firing rates in V1 (Vinje and Gallant, 2000, 2002). Consequently, normal operating regimes provide only a modest pedestal of spiking in which to decrement spike rates, while enhancing the encoding potential for spike rate increments. These features might be adaptations to the metabolic cost of spiking, as energy usage by the brain depends strongly on the rate of action potentials (Attwell and Laughlin, 2001; Lennie, 2003). Such constraints likely favor energy efficient neuronal codes (Gjorgjieva et al., 2014; Levy and Baxter, 1996). Our modeling experiments support this general idea, as increasing the cost of neuronal activity drove RNNs to adopt lower baseline rates of

firing and encouraged preferential reliance on increments in neuronal activity to detect changes in visual stimuli (Figures 7, S7).

In the retina, bipolar cells are functionally specialized to segregate increases and decreases luminance into distinct processing streams (ON and OFF cells; Schiller, 1992). This split allows a given number of retinal ganglion cells (RGCs) to convey information more efficiently than a single channel, especially when the costs of spiking are considered (Gjorgjieva et al., 2014). The fact that pharmacological inhibition of ON cells disrupts behavioral detection of light increments without affecting decrements (Schiller, 1982; Schiller et al., 1986) supports the idea that central structures decode only spike increments. Recently, Smeds and colleagues (2019) used a transgenic mouse with differentially elevated luminance threshold of the ON and OFF RGCs to show that mice doing an absolute luminance detection task depend the ON pathway, even when the OFF pathway could provide greater sensitivity. These results strongly support the idea that decoding by central visual structures depends primarily on increases in spiking.

It remains to be determined how general these results are for cerebral cortex and other brain structures. Increases and decreases in spike output are likely to be equally important for signaling in other brain areas. In the cerebellum and the basal ganglia aspects of eye position (Lisberger and Fuchs, 1978), head position (Barter et al., 2015), speed and rotation (Muzzu et al., 2018), and action initiation (Krause et al., 2010) all appear to critically rely on decrements in spike rate. Most of these systems exhibit relatively higher baseline rates of firing compared to cerebral cortex, which provides a larger pedestal of activity for decrements in firing to signal information. It would be unsurprising if unique constraints have shaped diverse sets of information processing strategies across different brain regions.

Determining how changes in spike rate mediate the functions of neural circuits is critical for understanding the brain. Looking ahead, advances in holographic stimulation techniques are enabling detailed investigations into how individual neurons contribute to behaviors (Carrillo-Reid et al., 2019; Forli et al., 2018; Lerman et al., 2018; Marshel et al., 2019; Russell et al., 2019). Notwithstanding, there is still very little causal evidence describing how the brain uses changes in neuronal spiking to render information to downstream areas. Further work will be required to establish how increments and decrements in spiking contribute to different brain computations. Nevertheless, the data presented here highlight a strong asymmetry in how increases and decreases in neuronal spiking in V1 relate to perceptual reports.

Author Contributions: JJC, NYM, and JHRM designed research. JJC, MLB, NYM, and EAP conducted research. JJC and NYM analyzed data. JJC wrote the manuscript with input from all other authors.

Acknowledgements: The authors thank Dr. Vytas Bindokas and the University of Chicago Integrated Light Microscopy Core Facility for assistance with confocal imaging and Dr. Supriya Ghosh for assistance with electrophysiological recordings. The authors also thank Zaina Zayyad, Dr. Supriya Ghosh, and Dr. Matthew Kaufman for critical feedback on the manuscript. This work was supported by an Albert O. Beckman Postdoctoral Fellowship (JJC), the Vannevar Bush Faculty Fellowship (DJF), and NIH U01-NS090576 and NIH U19-107464.

Declaration of Interests: The authors declare no competing interests.

Methods

Mouse Strains. All animal procedures were in compliance with the guidelines of the NIH and were approved by the Institutional Animal Care and Use Committee at The University of Chicago. Mouse lines were obtained from The Jackson Laboratory. Data come from parvalbumin-Cre mice (PV, 11 mice, 6 female; JAX stock #017320; Hippenmeyer et al., 2005), somatostatin-Cre mice (SST, 2 mice, both male; Jax stock #013044; Taniguchi et al., 2011), and Emx-1 Cre mice (Emx, 7 mice, 2 female, Jax stock #005628; Gorski et al., 2002). Experimental animals were heterozygous for Cre recombinase in the cell type of interest (outbred by crossing homozygous Cre-expressing strains with wild type BALB/c mice, Jax stock #000651). Mice were singly housed on a reverse light/dark cycle with ad libitum access to food. Mice were water scheduled throughout behavioral experiments, except for periods around surgeries. Mice used for electrophysiological recordings had ad libitum access to food and water.

Cranial window implant. Mice (3–5 months old) were implanted with a headpost and cranial window to give stable optical access for photostimulation during behavior (Glickfeld et al., 2013; Goldey et al., 2014; Histed and Maunsell, 2014). Animals were anesthetized with ketamine (40 mg/kg, i.p.), xylazine (2 mg/kg, i.p.) and isoflurane (1.2–2% in 100% O_2). Using aseptic technique, a headpost was secured to the skull using acrylic (C&B Metabond, Parkell) and a 3 mm craniotomy was made over the left cerebral hemisphere (3.0 mm lateral and 0.5 mm anterior to lambda) to implant a glass window (0.8 mm thickness; Tower Optical).

Intrinsic autofluorescence imaging. We located V1 by measuring changes in the intrinsic autofluorescence signal using visual stimuli and epifluorescence imaging (Andermann et al., 2011). Autofluorescence produced by blue excitation (470 ± 40 nm, Chroma) was collected using a green long-pass filter (500 nm cutoff) and a 1.0x air objective (Zeiss; StereoDiscovery V8 microscope; ~0.11 NA). Fluorescence was captured with a CCD camera (AxioCam MRm, Zeiss; 460×344 pixels; 4×3 mm field of view). The visual stimuli were full contrast drifting Gabors (10° SD; 30°/s; 0.1 cycles/deg) presented for 10 s followed by 6 s of mean luminance. The response to the visual stimulus was computed as the fractional change in fluorescence during the first 8 s of the stimulus presentation compared to the average of the last 4 s of the preceding blank.

Viral injections and ChR2 stimulation. Virus injections were targeted to a monocular region of V1 based on each animal's retinotopic map (+25° in azimuth; between -15° to +15° in elevation). Before virus injection, mice were anesthetized (isoflurane, 1–1.5%), and the glass window was removed using aseptic technique. We used a volume injection system (World Precision Instruments) to inject 200-400 nl of AAV9-Flex-ChR2-tdTomato (~10^{11} viral particles; Penn Vector Core) 300 µm below the pial surface. The virus was injected at a rate of 40 nl/min through a glass capillary attached to a 10 µL syringe (Hamilton). Following the injection, a new cranial window was sealed in place. Several weeks after injection, we localized the area of ChR2 expression using tdTomato fluorescence, and attached an optical fiber (400 µm diameter; 0.48 nA; Doric Lenses)

within 500 μm of the cranial window (~1.3 mm above the cortex). We delivered light though the fiber from a 455 nm LED (ThorLabs) and calibrated the total power at the entrance to the cannula. Optogenetic stimulation began no earlier than 4 weeks after injection. We prevented optogenetic stimuli from cueing the animal to respond by wrapping the fiber implant in blackout fabric (Thor Labs) that attached to the headpost using a custom mount.

Behavioral tasks. Mice were trained to respond to changes in a visual display for a water reward using a lever while head fixed (Histed et al., 2012). In the primary experiment, a static 50% contrast Gabor stimulus was continuously on the screen, presented on a uniform background with the same average luminance. Mice initiated trials by depressing a lever. Following a random delay (400-3000 ms), the contrast of the Gabor stimulus either incremented or decremented (interleaved). The Gabor stimulus (SD 5-7°, 0.1 cycles/deg, odd-symmetric) changed contrast for the duration of a brief response window. The size of the contrast change varied randomly from trial to trial across a range that spanned behavioral thresholds for both increments and decrements. The mouse had to release the lever within the response window running from 100 ms to 700 or 900 ms after change onset to receive a reward. Following completion of the trial, the contrast of the Gabor stimulus returned to 50%. Stimuli for each animal were positioned at a location that corresponded to the V1 representation expressing ChR2. Early releases and misses resulted in a brief timeout before the start of the next trial. Behavioral control, data collection and analysis were done using custom software written using Objective-C, MWorks (mworks-project.org), Matlab (MathWorks) and Python.

Optogenetic stimulation did not begin until animals worked reliably for hundreds of trials each day and performance was stable at threshold for both increments and decrements. This typically required ~2.5 months of training. During optogenetic experiments, we activated ChR2 expressing neurons on a randomly selected half of trials for a single contrast change (around ±15% for all mice). These change magnitudes were chosen for stimulation as they approximated the increment and decrement detection thresholds and thus maximized our ability to resolve an impairment or facilitation of detection capability. We aligned the opsin illumination with visually evoked spiking in V1 by delaying the optogenetic stimulus by 35 ms relative to the appearance of the visual stimulus on the monitor. Opsin illumination persisted until the end of the trial to prevent mice from using the offset of optogenetic input as a task-relevant signal. The optogenetic stimulation intensity was fixed within a session and chosen for each mouse basis based on 1 or 2 preliminary testing sessions that were not included in the main analysis. Using these preliminary observations, powers were selected to avoid saturating behavioral performance (ranges for high power sessions: Emx: 0.12-0.25 mW; PV: 0.12-0.25 mW). Following data collection at high powers, we conducted additional sessions in some mice at lower optogenetic stimulus intensities to determine how changes in performance scaled with power (ranges 0.02-0.15 mW).

Follow up experiments included retraining some PV mice (n=2; both female) to detect contrast increments of a counterphasing Gabor (2 or 4 Hz). The average contrast of

Gabor stimulus was held at 20%, except during contrast changes. Changes in contrast were synchronized with zero crossings of the temporal modulation to avoid generating instantaneous luminance steps that could cue the animal to respond. Optogenetic stimulation was delivered on a random subset of trials for a moderate contrast change magnitude (+30%) and the powers used with counterphase-modulated stimuli were identical to those used in the main experiments. As above, optogenetic stimulation was delivered from stimulus onset until the end of the trial. Other task variations included shortening the duration of visual and optogenetic stimulation or ramping and stepping optogenetic stimuli up or down during contrast changes.

Histology. Mice were perfused with 10% pH-neutral buffered formalin (Millipore Sigma Inc.), after which the brain was removed and submerged in fixative for 24 hr. The brain was subsequently rinsed with PBS, placed in a 30% sucrose PBS solution until it sank. Brains were sectioned at 40 μm on a freezing microtome, mounted and cover slipped with DAPI Fluoromount-G (Southern Biotech). tdTomato expression and DAPI labeling were visualized with 561 and 405 nm excitation light respectively, using a Leica SP5 Confocal Microscope.

Behavioral data analysis. The proportion correct for each contrast change level was determined using trials in which the subject either responded correctly (hit) or failed to respond (miss). Trials in which the animal released the lever before stimulus onset (false alarm) were not considered in performance analyses. Sessions in which the false alarm rate or the miss rate was greater than 50% were excluded from analysis. We estimated the likelihood of observing a false hit by calculating the conditional probability that the animal would release the lever in each 100 ms bin given that a stimulus had yet to occur in that or earlier bins. The false hit rate represents the probability that the mouse would get a trial correct due to spontaneous lever releases independent of detecting stimulus changes and is thus a lower bound on performance. The false hit rate was low (Emx median 5.3%; range, 3.8-6.3%; PV median 5.3%; range, 2.8-8.3%), demonstrating that mice were relying on stimulus changes to guide their responses. To correct for false hits, we subtracted a randomly selected fraction of correct trials from each contrast level based on the estimated false hit probability observed in the corresponding session.

When performance data were fit to psychometric functions, we first corrected for the estimated false hit probability as described above. This correction was typically small (median hits removed 6.8%; range, 4.5–12.9% for 38 sessions from 4 mice). Corrected performance data were then fit with a Weibull cumulative distribution function using non-linear least squares and variance weighting of each mean. The two psychometric functions (with and without ChR2 stimulation) were fit simultaneously using four parameters: individual thresholds ($\alpha_{unstimulated}$, $\alpha_{stimulated}$), a common lapse rate (γ), and a common slope (β) such that:

$$\text{Proportion Correct} = (1-\gamma) \times \left(1 - e^{-\left(\frac{contrast}{\alpha}\right)^{\beta}}\right)$$

Threshold confidence intervals were estimated using a bootstrap (1000 repetitions, $p < 0.05$, one-tailed).

To compare how optogenetic increments and decrements affected lever responses relative to the trial-time-matched false rate (Figure 6), we used the stimulus onset times for 0% contrast change trials in which mice correctly responded to optogenetic stimuli. We restricted our analyses to stimulus onset times that occurred before the final 400 ms of all possible onset times as the small number of observations for the longest trial times made this calculation unstable over the response window. Using only trials in which stimuli had yet to occur by the stimulus onset time, we expressed the lever release time for false alarms relative to the stimulus onset time. Thus, the false alarm distribution aligned to the onset of stimuli served as a time-in-trial matched measure for the lever release probability over time for static optogenetic input. This process was repeated within each session and the probability of lever releases were compared for optogenetic increments, decrements, and time-matched false alarms.

Electrophysiological recordings. We recorded extracellularly from V1 in awake, head-fixed mice (n=4 Emx, 1 female; n=4 PV, 3 female) using multisite silicon probes (Neuronexus, Inc.; 32-site model 4 × 8-100–200-177). Some mice were first used for behavioral experiments (2 PV, both female), while the rest were untrained but injected with opsins before recording. Electrodes were electroplated with a gold solution mixed with carbon nanotubes (Ferguson et al., 2009; Keefer et al., 2008) to impedances between 200-500 kΩ.

At the start of recording sessions, mice were anesthetized with isoflurane (1.2–2% in 100% O_2), placed in a sled and head-fixed. While anesthetized, the eyes were kept moist with 0.9% saline. We visualized ChR2-expressing areas of monocular visual cortex by imaging tdTomato fluorescence with a fluorescence microscope and camera (Zeiss, Inc). The cranial window was then removed and the electrodes lowered through a slit in the dura. We then positioned an optic fiber above the cortex at a distance comparable to that used during behavioral experiments (1.0-1.5 mm). The craniotomy was then covered with 3% agarose dissolved in aCSF (Millipore Sigma Inc., TOCRIS respectively) Following the recovery period of 1 hour, anesthetic was removed and we waited an additional hour for recovery from anesthesia.

The electrode was advanced to locate responsive units, and was allowed to settle for 30 minutes before collecting data. Delivery of visual and optogenetic stimuli and data acquisition was computer controlled. Concurrent visual and optogenetic stimuli matched those used during behavioral experiments except that visual and optogenetic changes were presented for 500 ms rather than 700-900 ms and the visual stimuli filled the video display. We recorded at least 25 repetitions of each stimulus condition in a given stimulus set. For optogenetic stimuli presented in isolation (Figure 6), optogenetic input ramped up for 250 ms at the beginning of each stimulus to a moderate baseline power (0.5 mW), where it remained for 750 ms, and stepped up or down in intensity (randomly interleaved) for 250 ms before returning to the baseline for the remainder of the trial.

Electrode signals were amplified, bandpass filtered (750 Hz to 7.5 kHz) sampled around threshold crossings (Blackrock, Inc.) and spikes were sorted offline (OfflineSorter, Plexon, Inc.). Visually responsive units were taken as those with a 10% change in the average firing rate during the 50-250 ms after stimulus onset (stimulus period) relative to the average firing rate during the baseline epoch (baseline period -250 to -50 ms before stimulus onset) for the largest stimulus intensities. Optogenetically responsive units were defined as any unit with a significant difference ($p < 0.05$; signed-rank test) in firing rate during the stimulus period for the same visual stimulus with and without optogenetic stimulation. For optogenetic stimuli presented in isolation (Figure 6), optogenetically responsive units were classified based on significant changes in firing for optogenetic increments or decrements (from 0-250 ms following optogenetic steps) relative to pre-change firing rates (from -250 - 0 ms; $P < 0.05$; signed-rank test). We recorded both single and multi-units but did not differentiate between them because our primary interest was how optogenetic manipulations affect visually evoked responses across the V1 population.

Neural network models. We trained recurrent neural network (RNN) models on a task similar to the one that the mice perform to test whether *in silico* networks adopt the same strategies as *in vivo* networks. RNNs were trained and simulated using the Python machine learning framework TensorFlow (Abadi et al., 2016), and the network architecture was based on our previous study (Masse et al., 2019). Briefly, all networks consisted of orientation and contrast selective input neurons (whose firing rates are represented as $\boldsymbol{u}(t)$) that projected onto 100 recurrently connected neurons (whose firing rates are represented as $\boldsymbol{h}(t)$, which in turn projected onto the output layer (Figure 7). Recurrently connected neurons never sent projections onto themselves.

The activity of the recurrent neurons was modeled to follow the dynamical system (Song et al. 2016):

$$\tau \frac{d\boldsymbol{h}}{dt} = -\boldsymbol{h} + f(W^{rec}\boldsymbol{h} + W^{in}\boldsymbol{u} + \boldsymbol{b}^{rec} + \sqrt{2\tau}\sigma_{rec}\zeta)$$

where τ is the neuron's time constant (set to 50 ms), $f(\cdot)$ is the activation function, W^{rec} and W^{in} are the synaptic weights between recurrent neurons, and between input and recurrent neurons, respectively, $\boldsymbol{b}^{rec}$ is a bias term, ζ is independent Gaussian white noise with zero mean and unit variance applied to all recurrent neurons, and σ_{rec} is the strength of the noise (set to 0.05). To ensure that neuron's firing rates were non-negative and non-saturating, we chose the rectified linear (ReLu) function as our activation function: $f(x) = max(0, x)$.

To simulate the network, we used a first-order Euler approximation with time step Δt:

$$\boldsymbol{h}_t = (1-\alpha)\boldsymbol{h}_{t-1} + \alpha f(W^{rec}\boldsymbol{h}_{t-1} + W^{in}\boldsymbol{u}_t + \boldsymbol{b}^{rec} + \sqrt{\frac{2}{\alpha}}\sigma_{rec}N(0,1))$$

where $\alpha = \frac{\Delta t}{\tau}$ and $N(0,1)$ indicates the standard normal distribution.
The decision to release the lever (when there was a contrast change) or to hold the lever (when there was no contrast change) was mediated by a competition between two output units. The 80 excitatory neurons linearly projected onto the output unit associated with releasing the lever:

$$z_t{}^{release} = W^{out}\boldsymbol{h}_t + b^{out}$$

where W^{out} are the synaptic weights between the excitatory neurons and the output unit, and b^{out} is a bias term. The activity of the output unit associated with holding the lever was simply the negative of the activity of the unit associated with releasing the lever: $z_t{}^{hold} = -z_t{}^{release}$.

We then calculated the network policy, $\boldsymbol{\pi}_t$, which was the probability of holding or releasing the lever, by taking the softmax of these two values:

$$\boldsymbol{\pi}_t = \text{softmax}\left([z_t^{hold}, z_t^{release}]\right)$$

To maintain separate populations of 80 excitatory and 20 inhibitory neurons, we decomposed the recurrent weight matrix, W^{rec} as the product between a matrix for which all entries are non-negative, $W^{rec,+}$ whose values were trained, and a fixed diagonal matrix, D, composed of 1s and -1s, corresponding to excitatory and inhibitory neurons, respectively (Song et al. 2016):

$$W^{rec} = W^{rec,+}D$$

$$D = \begin{bmatrix} 1 & \cdots & \\ \vdots & \ddots & \vdots \\ & \cdots & -1 \end{bmatrix}$$

Initial connection weights between excitatory neurons were randomly sampled from a Gamma distribution with shape parameter of 0.1 and scale parameter of 1.0, and then multiplied by 0.25. Initial connections weights projecting to or from inhibitory neurons were sampled from a Gamma distribution with shape parameter of 0.1 and scale parameter of 1.0 and then multiplied by 0.5. Initial bias values were set to 0.

Networks consisted of 24 orientation and contrast selective input neurons. The tuning of the input neurons followed a Von Mises' distribution, such that the activity of the input neuron i was

$$u^i = g^i(contrast)Aexp(\kappa\cos(\theta - \theta^i_{pref})) + \sqrt{\frac{2}{\alpha}}\sigma_{in}N(0,1)$$

where θ is the orientation of the stimulus (always fixed at 0°), θ^{i}_{pref} is the preferred direction of input neuron i, κ was set to 2, and A was set to $\frac{1}{\exp(\kappa)}$. The strength of the input activity noise, σ_{in}, was set to 0.05. The function $g^{i}(contrast)$ determined how neuron i responded to different contrasts. Half (12) of the input neurons were contrast increasing, defined as $g^{i}(contrast) = contrast$, while the other half were contrast decreasing, defined as as $g^{i}(contrast) = 1/contrast$.

Contrast change detection task for network model. The networks were trained to indicate whether the stimulus contrast changed by responding within a fixed interval. Trials lasted 3000 ms, divided into 10 ms steps. A Gabor patch with an orientation of 0° and a baseline contrast level was presented from the start of the trial, and at a random time, the contrast either doubled or was halved for 100 ms, before returning to baseline. The time of the contrast change was randomly sampled from an exponential distribution with a time constant of 1300 ms, plus 400 ms. If the contrast change did not occur before the end of the trial, the network was rewarded for maintaining hold of the bar throughout the trial. The network received a reward of +1 if it chose to release the lever during the 100 ms duration contrast change, a reward of – 0.1 if it chose to either release the lever before the contrast change (i.e. false alarm), or not release during the contrast change (i.e. miss).

Network training. The RNNs using the actor-critic reinforcement learning method (Barto et al. 1983), in which the networks were trained to maximize the discounted cumulative future reward:

$$R_{\tau} = \sum_{t=\tau}^{T} \gamma^{t-\tau} r_t$$

where $\gamma \epsilon (0,1]$ is the discount factor and r_t is the reward given at time t. The network was trained to estimate this discounted future reward as a linear projection from the recurrent units

$$\boldsymbol{V_t} = W^{value}\boldsymbol{h}_t + b^{value}$$

by minimizing the loss function:

$$\mathcal{L}_V = \frac{1}{2T}\sum_{t=1}^{T}[V_t - r_t - \gamma V_{t+1}]^2$$

Concurrently, the network adjusts the network policy, $\boldsymbol{\pi}_t$ (described above) to select the actions that would lead to the greatest cumulative reward by minimizing the loss function:

$$\mathcal{L}_P = -\frac{1}{T}\sum_{t=1}^{T}[V_t - r_t - \gamma V_{t+1}]\log \boldsymbol{\pi}_t$$

The network was also encouraged to explore different strategies by maximizing the entropy of the policy output:

$$\mathcal{L}_H = -\frac{1}{T}\sum_{t=1}^{T} \boldsymbol{\pi}_t \log \boldsymbol{\pi}_t$$

Finally, the network was encouraged to solve the task using low levels of neural activity by minimizing the L2 norm of the recurrent neuron firing rates:

$$\mathcal{L}_{sc} = \frac{1}{N \times T}\sum_{t=1}^{T}\sum_{i=1}^{N} h_{i,t}^2$$

where $h_{i,t}$ is the neural activity of the ith recurrent neuron at time t.

All together, the overall loss function is the weighted sum of all four terms:

$$\mathcal{L} = \mathcal{L}_P + \beta\mathcal{L}_V - \alpha\mathcal{L}_H + \gamma\mathcal{L}_{sc}$$

where α and β were set to 0.01. To understand different network solutions across various metabolic constraints, γ was randomly sampled for each network from a logarithmically uniform distribution between $10^{-6.5}$ and $10^{-0.5}$.

We trained all network parameters using the Adam version of stochastic gradient descent, with 1st and 2nd moment decay rates set to their default values (0.9 and 0.999, respectively). All networks were trained for 50,000 batches, with a batch size of 1024 trials and a learning rate of 0.001.

Analysis of Neural Network Responses. To link unit responses with network decisions to respond or withhold responses to contrast changes, we calculated the mean normalized ROC value for each network. First, we calculated the ROC for each excitatory unit by comparing its firing rate distributions when the network decided to release versus continue holding the lever following contrast changes. The ROC values for increments and decrements were then averaged for each unit.

Next, we summed a rescaled version of each units ROC value, normalized by the absolute value of this metric, such that:

$$\Sigma(ROC(i) - 0.5) \div \Sigma|(ROC(i) - 0.5)|$$

Where *i* is the index of all units in the network. Here, a value of -1 is the extreme case where all units in the network have lower firing rates when the network releases than

when it withholds, whereas a value of +1 indicates all units had higher firing rates when the network responds compared to trials in which it does not respond.

References:

Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., Isard, M., et al. (2016). TensorFlow: A System for Large-scale Machine Learning. In Proceedings of the 12th USENIX Conference on Operating Systems Design and Implementation, (Berkeley, CA, USA: USENIX Association), pp. 265–283.

Aizenberg, M., Mwilambwe-Tshilobo, L., Briguglio, J.J., Natan, R.G., and Geffen, M.N. (2015). Bidirectional Regulation of Innate and Learned Behaviors That Rely on Frequency Discrimination by Cortical Inhibitory Neurons. PLoS Biol. *13*, 1–32.

Andermann, M.L., Kerlin, A.M., Roumis, D.K., Glickfeld, L.L., and Reid, R.C. (2011). Functional specialization of mouse higher visual cortical areas. Neuron *72*, 1025–1039.

Atallah, B. V., Bruns, W., Carandini, M., and Scanziani, M. (2012). Parvalbumin-Expressing Interneurons Linearly Transform Cortical Responses to Visual Stimuli. Neuron *73*, 159–170.

Attwell, D., and Laughlin, S.B. (2001). An Energy Budget for Signaling in the Grey Matter of the Brain. J. Cereb. Blood Flow Metab. *21*, 1133–1145.

Bak, M., Girvin, J.P., Hambrecht, F.T., Kufta, C. V., Loeb, G.E., and Schmidt, E.M. (1990). Communication: Visual sensations produced by intracortical microstimulation of the human occipital cortex. Med. Biol. Eng. Comput. *28*, 257–259.

Barter, J.W., Li, S., Sukharnikova, T., Rossi, M.A., Bartholomew, R.A., and Yin, H.H. (2015). Basal ganglia outputs map instantaneous position coordinates during behavior. J. Neurosci. *35*, 2703–2716.

Carrillo-Reid, L., Han, S., Yang, W., Akrouh, A., and Yuste, R. (2019). Controlling Visually Guided Behavior by Holographic Recalling of Cortical Ensembles. Cell *178*, 447-457.e5.

Cone, J.J., Scantlen, M.D., Histed, M.H., and Maunsell, J.H.R. (2019). Different inhibitory interneuron cell classes make distinct contributions to visual contrast perception. ENeuro.

DeAngelis, G.C., and Newsome, W.T. (2004). Perceptual “read-out” of conjoined direction and disparity maps in extrastriate area MT. PLoS Biol. *2*.

Doty, R.W. (1965). CONDITIONED REFLEXES ELICITED BY ELECTRICAL STIMULATION OF THE BRAIN IN MACAQUES. J. Neurophysiol. *28*, 623–640.

Doty, R.W. (1969). Electrical stimulation of the brain in behavioral context. Annu. Rev. Psychol. *20*, 289–320.

Ferguson, J.E., Boldt, C., and Redish, A.D. (2009). Creating low-impedance tetrodes by electroplating with additives. Sens. Actuators. A. Phys. *156*, 388–393.

Ferster, D., and Chung, S. (1998). Strength and orientation tuning of the thalamic input to simple cells revealed by electrically evoked cortical suppression. Neuron *20*, 1177–1189.

Forli, A., Vecchia, D., Binini, N., Succol, F., Bovetti, S., Moretti, C., Nespoli, F., Mahn,

M., Baker, C.A., Bolton, M.M., et al. (2018). Two-Photon Bidirectional Control and Imaging of Neuronal Excitability with High Spatial Resolution In Vivo. Cell Rep. *22*, 3087–3098.

Gjorgjieva, J., Sompolinsky, H., and Meister, M. (2014). Benefits of pathway splitting in sensory coding. J. Neurosci. *34*, 12127–12144.

Glickfeld, L.L., Histed, M.H., and Maunsell, J.H.R. (2013). Mouse primary visual cortex is used to detect both orientation and contrast changes. J. Neurosci. *33*, 19416–19422.

Goldey, G.J., Roumis, D.K., Glickfeld, L.L., Kerlin, A.M., Reid, R.C., Bonin, V., and Andermann, M.L. (2014). Versatile cranial window strategies for long-term two-photon imaging in awake mice. Nat. Protoc. *9*, 2515–2538.

Gorski, J.A., Talley, T., Qiu, M., Puelles, L., Rubenstein, J.L.R., and Jones, K.R. (2002). Cortical excitatory neurons and glia, but not GABAergic neurons, are produced in the Emx1-expressing lineage. J. Neurosci. *22*, 6309–6314.

Gu, Y., deAngelis, G.C., and Angelaki, D.E. (2012). Causal links between dorsal medial superior temporal area neurons and multisensory heading perception. J. Neurosci. *32*, 2299–2313.

Guo, Z.V., Li, N., Huber, D., Ophir, E., Gutnisky, D., Ting, J.T., Feng, G., and Svoboda, K. (2014). Flow of Cortical Activity Underlying a Tactile Decision in Mice. Neuron *81*, 179–194.

Hippenmeyer, S., Vrieseling, E., Sigrist, M., Portmann, T., Laengle, C., Ladle, D.R., and Arber, S. (2005). A developmental switch in the response of DRG neurons to ETS transcription factor signaling. PLoS Biol. *3*, 0878–0890.

Histed, M.H., and Maunsell, J.H.R. (2014). Cortical neural populations can guide behavior by integrating inputs linearly, independent of synchrony. Proc. Natl. Acad. Sci. U. S. A. *111*, E178-87.

Histed, M.H., Bonin, V., and Reid, R.C. (2009). Direct Activation of Sparse, Distributed Populations of Cortical Neurons by Electrical Microstimulation. Neuron *63*, 508–522.

Histed, M.H., Carvalho, L.A., and Maunsell, J.H.R. (2012). Psychophysical measurement of contrast sensitivity in the behaving mouse. J. Neurophysiol. *107*, 758–765.

Histed, M.H., Ni, A.M., and Maunsell, J.H.R. (2013). Insights into cortical mechanisms of behavior from microstimulation experiments. Prog. Neurobiol. *103*, 115–130.

Huber, D., Petreanu, L., Ghitani, N., Ranade, S., Hromádka, T., Mainen, Z., and Svoboda, K. (2008). Sparse optical microstimulation in barrel cortex drives learned behaviour in freely moving mice. Nature.

Jazayeri, M., Lindbloom-Brown, Z., and Horwitz, G.D. (2012). Saccadic eye movements evoked by optogenetic activation of primate V1. Nat. Neurosci.

Keefer, E.W., Botterman, B.R., Romero, M.I., Rossi, A.F., and Gross, G.W. (2008). Carbon nanotube coating improves neuronal recordings. Nat. Nanotechnol. *3*, 434–439.

Kim, S., Callier, T., Tabot, G.A., Gaunt, R.A., Tenore, F. V., and Bensmaia, S.J. (2015).

Behavioral assessment of sensitivity to intracortical microstimulation of primate somatosensory cortex. Proc. Natl. Acad. Sci. U. S. A. *112*, 15202–15207.

Krause, M., German, P.W., Taha, S.A., and Fields, H.L. (2010). A pause in nucleus accumbens neuron firing is required to initiate and maintain feeding. J. Neurosci. *30*, 4746–4756.

Kubota, Y., Karube, F., Nomura, M., and Kawaguchi, Y. (2016). The Diversity of Cortical Inhibitory Synapses. Front. Neural Circuits *10*, 1–15.

Lennie, P. (2003). The cost of cortical computation. Curr. Biol. *13*, 493–497.

Lerman, G.M., Gill, J. V, Rinberg, D., and Shoham, S. (2018). Precise optical probing of perceptual detection. BioRxiv 456764.

Levy, W.B., and Baxter, R.A. (1996). Energy Efficient Neural Codes. Neural Comput. *8*, 531–543.

Lisberger, S.G., and Fuchs, A.F. (1978). Role of primate flocculus during rapid behavioral modification of vestibuloocular reflex. I. Purkinje cell activity during visually guided horizontal smooth-pursuit eye movements and passive head rotation. J. Neurophysiol. *41*, 733–763.

Logothetis, N.K., Augath, M., Murayama, Y., Rauch, A., Sultan, F., Goense, J., Oeltermann, A., and Merkle, H. (2010). The effects of electrical microstimulation on cortical signal propagation. Nat. Neurosci. *13*, 1283–1291.

Madisen, L., Mao, T., Koch, H., Zhuo, J., Berenyi, A., Fujisawa, S., Hsu, Y.-W.A., Garcia, A.J., Gu, X., Zanella, S., et al. (2012). A toolbox of Cre-dependent optogenetic transgenic mice for light-induced activation and silencing. Nat. Neurosci. *15*, 793–802.

Marg, E., and Rudiak, D. (1994). Phosphenes induced by magnetic stimulation over the occipital brain: description and probable site of stimulation. Optom. Vis. Sci. *71*, 301–311.

Marshel, J.H., Kim, Y.S., Machado, T.A., Quirin, S., Benson, B., Kadmon, J., Raja, C., Chibukhchyan, A., Ramakrishnan, C., Inoue, M., et al. (2019). Cortical layer–specific critical dynamics triggering perception. Science (80-.). *365*, eaaw5202.

Masse, N.Y., Yang, G.R., Song, H.F., Wang, X.J., and Freedman, D.J. (2019). Circuit mechanisms for the maintenance and manipulation of information in working memory. Nat. Neurosci.

Merabet, L.B., Theoret, H., and Pascual-Leone, A. (2003). Transcranial magnetic stimulation as an investigative tool in the study of visual function. Optom. Vis. Sci. *80*, 356–368.

Moore, A.K., Weible, A.P., Balmer, T.S., Trussell, L.O., and Wehr, M. (2018). Rapid Rebalancing of Excitation and Inhibition by Cortical Circuitry. Neuron *97*, 1341-1355.e6.

Murphey, D.K., and Maunsell, J.H.R. (2007). Behavioral Detection of Electrical Microstimulation in Different Cortical Visual Areas. Curr. Biol. *17*, 862–867.

Murphey, D.K., and Maunsell, J.H.R. (2008). Electrical microstimulation thresholds for behavioral detection and saccades in monkey frontal eye fields. Proc. Natl. Acad. Sci.

105, 7315 LP – 7320.

Murphey, D.K., Maunsell, J.H.R., Beauchamp, M.S., and Yoshor, D. (2009). Perceiving electrical stimulation of identified human visual areas. Proc. Natl. Acad. Sci. U. S. A. *106*, 5389–5393.

Muzzu, T., Mitolo, S., Gava, G.P., and Schultz, S.R. (2018). Encoding of locomotion kinematics in the mouse cerebellum. PLoS One *13*, e0203900–e0203900.

Nagel, G., Szellas, T., Huhn, W., Kateriya, S., Adeishvili, N., Berthold, P., Ollig, D., Hegemann, P., and Bamberg, E. (2003). Channelrhodopsin-2, a directly light-gated cation-selective membrane channel. Proc. Natl. Acad. Sci. *100*, 13940–13945.

Nichols, M.J., and Newsome, W.T. (2002). Middle temporal visual area microstimulation influences veridical judgments of motion direction. J. Neurosci. *22*, 9530–9540.

O'Connor, D.H., Hires, S.A., Guo, Z. V, Li, N., Yu, J., Sun, Q.-Q., Huber, D., and Svoboda, K. (2013). Neural coding during active somatosensation revealed using illusory touch. Nat. Neurosci. *16*, 958.

Okun, M., and Lampl, I. (2008). Instantaneous correlation of excitation and inhibition during ongoing and sensory-evoked activities. Nat. Neurosci. *11*, 535–537.

Penfield and Rasmussen (1950). The Cerebral Cortex of Man: A Clinical Study of Localization of Function. J. Am. Med. Assoc. *144*, 1412.

Pfeffer, C.K., Xue, M., He, M., Huang, Z.J., and Scanziani, M. (2013). Inhibition of inhibition in visual cortex: The logic of connections between molecularly distinct interneurons. Nat. Neurosci. *16*, 1068–1076.

Rattay, F. (1999). The basic mechanism for the electrical stimulation of the nervous system. Neuroscience *89*, 335–346.

Resulaj, A., Ruediger, S., Olsen, S.R., and Scanziani, M. (2018). First spikes in visual cortex enable perceptual discrimination. Elife *7*, 1–22.

Romo, R., Hernandez, A., and Zainos, A. (2000). Sensing without Touching : Somatosensory discrimination based on cortical microstimulation. Neuron *26*, 273–278.

Russell, L.E., Yang, Z., Tan, P.L., Fişek, M., Packer, A.M., Dalgleish, H.W.P., Chettih, S., Harvey, C.D., and Häusser, M. (2019). The influence of visual cortex on perception is modulated by behavioural state. BioRxiv 706010.

Salzman, C.D., Murasugi, C.M., Britten, K.H., and Newsome, W.T. (1992). Microsimulation in visual area MT: Effects on direction discrimination performance. J. Neurosci. *12*, 2331–2355.

Schiller, P.H. (1982). Central connections of the retinal ON and OFF pathways. Nature *297*, 580–583.

Schiller, P.H. (1992). The ON and OFF channels of the visual system. Trends Neurosci. *15*, 86–92.

Schiller, P.H., Sandell, J.H., and Maunsell, J.H. (1986). Functions of the ON and OFF channels of the visual system. Nature *322*, 824–825.

Schmidt, E.M., Bak, M.J., Hambrecht, F.T., Kufta, C. V, O'Rourke, D.K., and Vallabhanath, P. (1996). Feasibility of a visual prosthesis for the blind based on intracortical microstimulation of the visual cortex. Brain *119 (Pt 2*, 507–522.

Smeds, L., Takeshita, D., Turunen, T., Tiihonen, J., Westö, J., Martyniuk, N., Seppänen, A., and Ala-Laurila, P. (2019). Paradoxical Rules of Spike Train Decoding Revealed at the Sensitivity Limit of Vision. Neuron *104*, 576-587.e11.

Taniguchi, H., He, M., Wu, P., Kim, S., Paik, R., Sugino, K., Kvitsani, D., Fu, Y., Lu, J., Lin, Y., et al. (2011). A Resource of Cre Driver Lines for Genetic Targeting of GABAergic Neurons in Cerebral Cortex. Neuron *71*, 995–1013.

Verhoef, B.E., Vogels, R., and Janssen, P. (2012). Inferotemporal Cortex Subserves Three-Dimensional Structure Categorization. Neuron.

Vinje, W.E., and Gallant, J.L. (2000). Sparse coding and decorrelation in primary visual cortex during natural vision. Science (80-.). *287*, 1273–1276.

Vinje, W.E., and Gallant, J.L. (2002). Natural stimulation of the nonclassical receptive field increases information transmission efficiency in V1. J. Neurosci. *22*, 2904–2915.

Wehr, M., and Zador, A.M. (2003). Balanced inhibition underlies tuning and sharpens spike timing in auditory cortex. Nature *426*, 442–446.

Wiegert, J.S., Mahn, M., Prigge, M., Printz, Y., and Yizhar, O. (2017). Silencing Neurons: Tools, Applications, and Experimental Constraints. Neuron *95*, 504–529.

Wilson, N.R., Runyan, C.A., Wang, F.L., and Sur, M. (2012). Division and subtraction by distinct cortical inhibitory networks in vivo. Nature *488*, 343–348.

Xue, M., Atallah, B. V, and Scanziani, M. (2014). Equalizing excitation-inhibition ratios across visual cortical neurons. Nature *511*, 596–600.

Yi, G., and Grill, W.M. (2019). Average firing rate rather than temporal pattern determines metabolic cost of activity in thalamocortical relay neurons. Sci. Rep. *9*, 6940.

Supplemental Information

Emx-Cre	Decrements			Increments		
Mouse	**Hits/Miss unstim**	**Hits/Miss stim**	**p-value**	**Hits/Miss unstim**	**Hits/Miss stim**	**p-value**
1	79/73	153/31	$< 10^{-9}$	69/78	154/32	$< 10^{-11}$
2	82/121	166/123	$< 10^{-3}$	74/110	168/199	$< 10^{-3}$
3	66/86	149/87	$< 10^{-3}$	56/99	123/74	$< 10^{-6}$

Table S1. Summary of changes in detection performance produced by optogenetic stimulation of pyramidal neurons in Emx-Cre mice. We combined data across sessions within individual mice and calculated the relative proportion of hits and misses separately for contrast changes (±15%) with and without optogenetic stimulation of pyramidal neurons. Cumulative data from the three Emx mice is reported above, numbers correspond to the total counts of hits/misses for contrast decrements (left) and increments (right). P-values correspond to the results of Fisher's exact test comparing the proportions of hits/misses with and without stimulation independently for contrast decrements and increments. Optogenetic stimulation of excitatory neurons similarly facilitated contrast change detection, regardless of sign, in all three tested mice.

PV-Cre	Decrements			Increments		
Mouse	**Hits/Miss unstim**	**Hits/Miss stim**	**p-value**	**Hits/Miss unstim**	**Hits/Miss stim**	**p-value**
1	295/113	77/132	$< 10^{-23}$	203/83	92/82	$< 10^{-16}$
2	181/222	14/199	$< 10^{-25}$	149/255	19/195	$< 10^{-15}$
3	76/95	31/95	$< 10^{-7}$	69/47	30/88	$< 10^{-6}$
4	82/81	28/133	$< 10^{-9}$	74/90	34/129	$< 10^{-5}$
5	77/99	41/118	$< 10^{-4}$	74/86	43/116	$< 10^{-3}$
6	178/94	62/184	$< 10^{-19}$	197/79	97/149	$< 10^{-12}$

Table S2. Summary of changes in detection performance produced by optogenetic stimulation of PV interneurons in PV-Cre mice. As was done for Emx mice, we combined data across all sessions at the tested contrast changes (±15%) and calculated the relative proportion of hits and misses separately for trials with and without optogenetic stimulation of PV interneurons. Cumulative data from the six PV mice is shown above, numbers correspond to the total counts of hits/misses for contrast decrements (left) and increments (right). P-values are the results of Fisher's exact test comparing the proportions of hits/misses with and without stimulation independently for contrast decrements and increments. Optogenetic stimulation of PV interneurons similarly impaired detection of decrements and increments in all six tested mice.

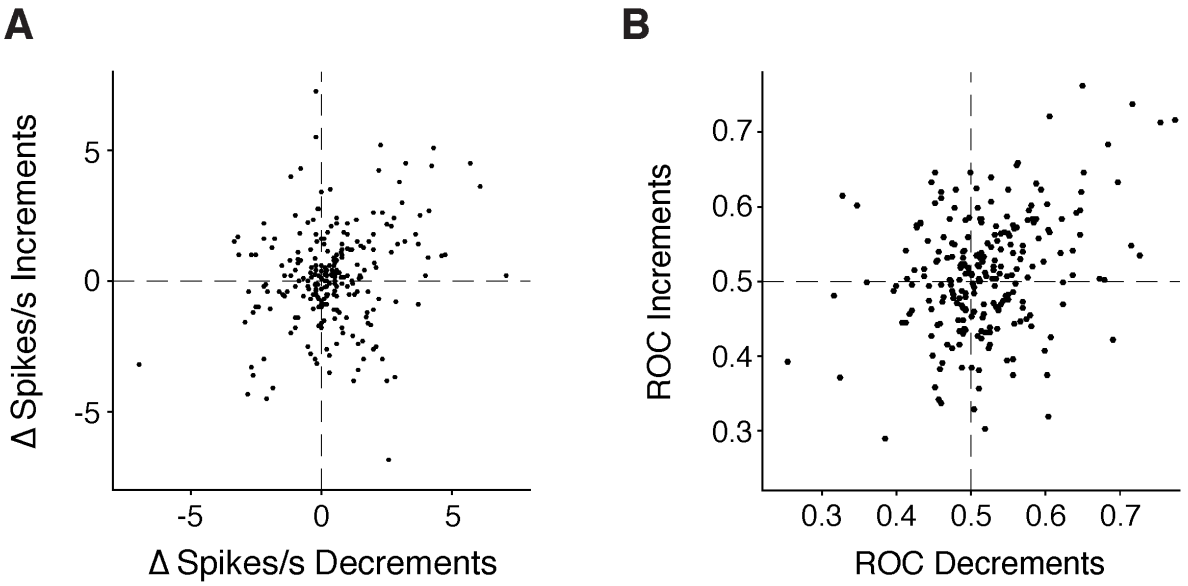

Figure S1. Individual units exhibit diverse responses to increments and decrements in contrast. (A) Points depict the change in firing rate (50-250 ms after stimulus onset – 50-250 ms before stimulus onset) evoked by a halving (decrements, x-axis) or doubling (increments, y-axis) of the baseline contrast. Different units are either excited or inhibited by contrast changes in various combinations. (B) To the measure the ability of individual units to detect visual contrast changes, we performed an ROC analysis for each unit that compared the baseline and stimulus-evoked firing rates. Points depict the ROC value for discriminating between the average baseline and stimulus firing rates for decrements (x-axis) and increments (y-axis) in contrast. While the median ROC values were both near 0.5 (decrements = 0.52; increments = 0.51), the range of ROCs values suggested there was reasonable predictive power for firing rate decreases (ROC<0.5) or increases (ROC>0.5) for either contrast change type (decrements: range 0.29-0.85; increments: range 0.25-0.79).

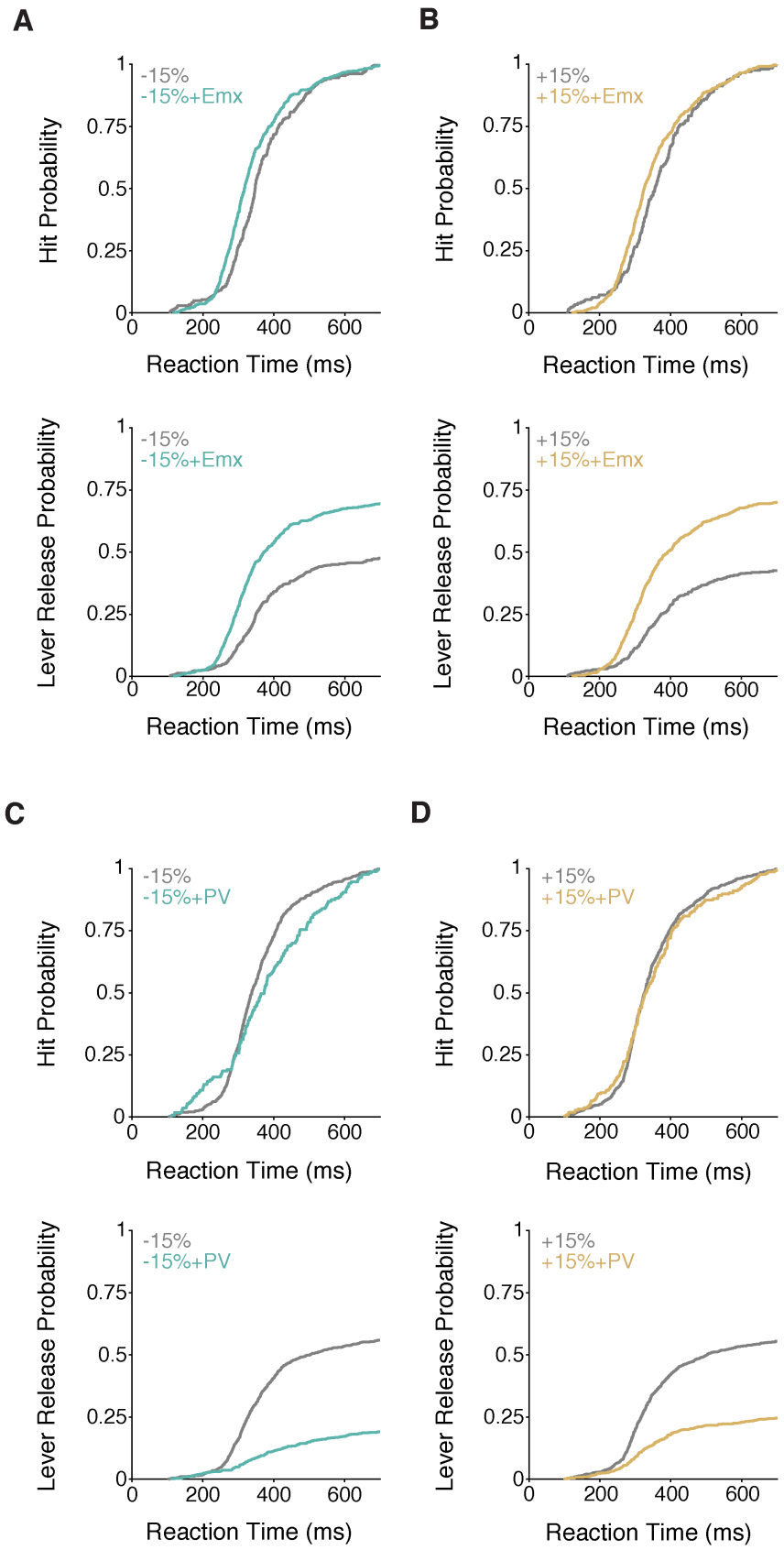
A
B
C
D
-15%
-15%+Emx
+15%
+15%+Emx
-15%+PV
+15%+PV
Hit Probability
Lever Release Probability
Reaction Time (ms)
0
0.25
0.5
0.75
1
200
400
600

Figure S2. Optogenetic stimulation of visual cortex augments reaction times for contrast decrements and increments. The processing time required to detect and respond to a visual stimulus strongly depends on the stimulus intensity (Nissen, 1977; Roitman and Shadlen, 2002). We thus sought to examine effects of optogenetic stimulation on reaction times. We combined reaction time data across all mice (n=3) at the tested contrast changes to determine if the time of lever releases following stimulus onset was affected by optogenetic stimulation. First, we only examined trials in which the mice correctly detected stimuli (A-B, top, hits). We found that optogenetic stimulation of excitatory neurons in Emx-Cre mice significantly enhanced reaction times for both contrast decrements (A, top; unstimulated (gray line) median = 347 ms, 296-414 IQR; stimulated (aqua line) median = 316 ms, 273 – 388 IQR; $p < 10^{-4}$;) and increments (B, top; unstimulated (gray) median = 356 ms, 295 – 424 IQR; stimulated (gold) median = 326 ms, 276 – 409 IQR; $p < 0.01$; Komolgorov-Smirnov test) on hit trials. Note: the absence of responses before 100 ms is due to the fact we only counted responses as hits if the lever release occurred 100 ms after stimulus onset, as otherwise responses would have been too early to be driven by visual stimuli. We also plotted reaction time data for hits and misses combined (A-B, bottom plots). These plots capture the combined effects of optogenetic stimulation on response probability (e.g., the difference in the cumulative probability of responses at the end of the reaction time window) along with the enhancement in reaction time for hit trials (the leftward shift and change in the rising phase of the lever response probability). C-D) Same as in A,B except for PV-Cre mice. PV stimulation augmented the reaction times on hit trials for contrast decrements (C, top: unstimulated (grey line) median = 338 ms, 290 – 406 IQR; stimulated (aqua line) median = 373 ms, 293 – 474 IQR; $p < 0.01$; Komolgorov-Smirnov test). However, we did not find a difference between the stimulated and unstimulated reaction time distributions for successfully detected contrast increments (D, top: unstimulated (gray) median = 326 ms, 284 - 396 IQR; stimulated (gold) median = 330 ms, 278 - 404 IQR; $p = 0.45$; Komolgorov-Smirnov test). We then examined the total response probabilities by including both hits and misses (bottom). PV stimulation strongly suppressed the probability of lever releases throughout the reaction time window.

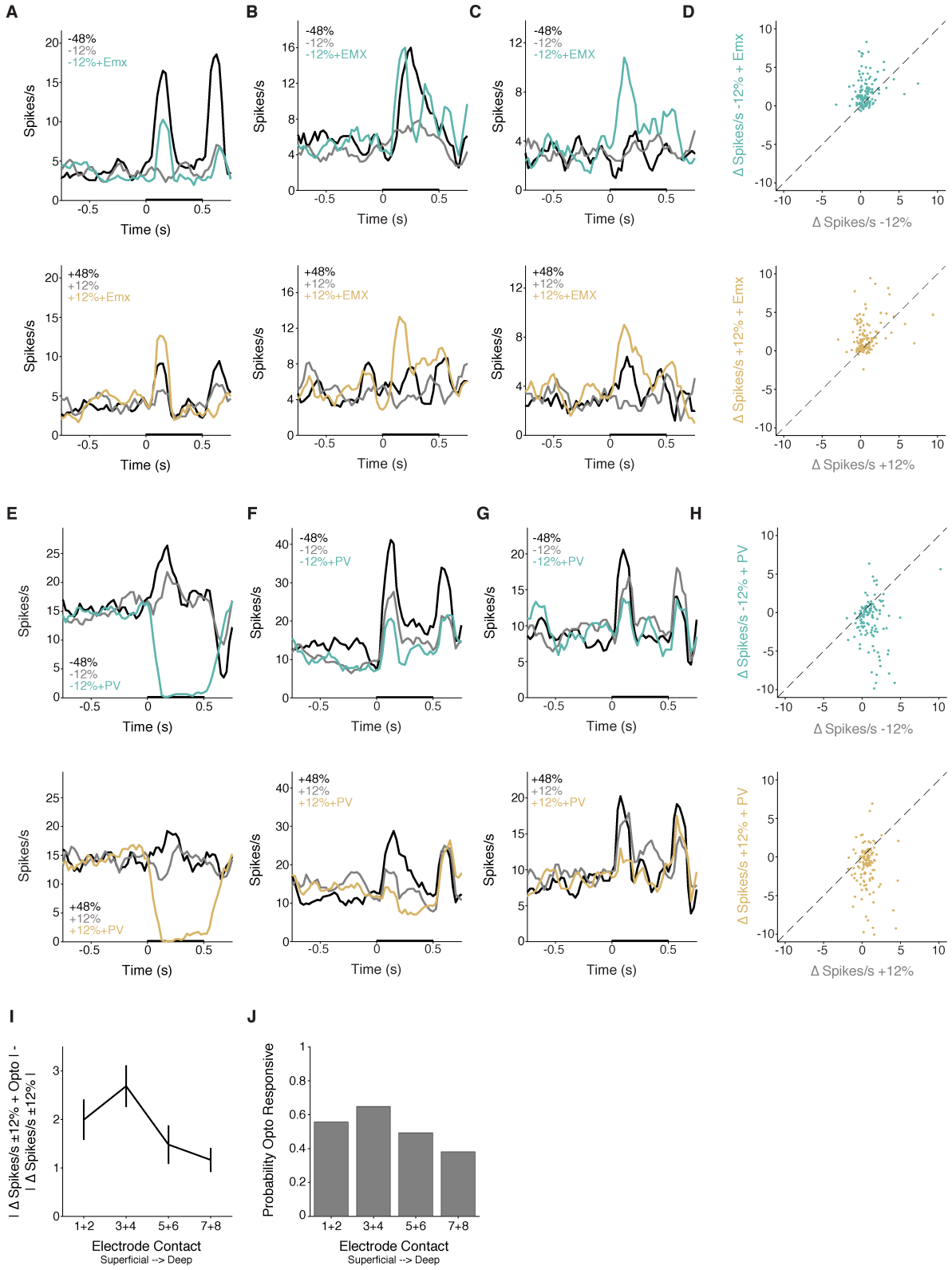
A
B
C
D
-48%
-12%
-12%+Emx
-12%+EMX
+48%
+12%
+12%+Emx
+12%+EMX
Spikes/s
Time (s)
Δ Spikes/s -12% + Emx
Δ Spikes/s -12%
Δ Spikes/s +12% + Emx
Δ Spikes/s +12%
E
F
G
H
-12%+PV
+12%+PV
Δ Spikes/s -12% + PV
Δ Spikes/s +12% + PV
I
J
| Δ Spikes/s ±12% + Opto | - | Δ Spikes/s ±12% |
Probability Opto Responsive
1+2
3+4
5+6
7+8
Electrode Contact
Superficial --> Deep

Figure S3. Pyramidal neuron stimulation potentiates, while PV neuron stimulation suppresses, V1 responses to decrements and increments in contrast and the optogenetic effects on spiking vary by cortical depth. A-C) Average firing rate traces from three different example units for contrast changes with and without optogenetic stimulation of excitatory neurons in Emx-Cre mice. Black line = response to large contrast decrements (+48%). Grey line = response to moderate contrast decrements (+12%). (top panels) Aqua line = response to moderate contrast decrements with concurrent optogenetic stimulation of pyramidal neurons. (lower panels) Gold line = response to moderate contrast increments with concurrent optogenetic stimulation of pyramidal neurons. Optogenetic stimulation (0.3 mW) potentiates contrast change responses. Bin sizes = 25 ms, all traces smoothed. D) Scatter plot shows the evoked change in firing rate (stimulus epoch – baseline) for each unit recorded in Emx-Cre mice (n = 119 units) in response to moderate contrast decrements (top) without (x-axis) and with (y-axis) optogenetic stimulation of pyramidal neurons. Optogenetic stimulation potentiated the response to contrast decrements for most units. 83 of 119 (70%) units had higher evoked responses for trials with optogenetic stimulation. Moreover, the distributions of firing rate changes were significantly different (decrements median = +0.56 ΔSpikes/s, 0.08-1.09 IQR; visual + optogenetic trials median = +1.4 ΔSpikes/s, 0.37-2.76 IQR, $p < 10^{-7}$, Komolgorov-Smirnov test). Similar effects were observed for contrast increments (bottom) as 81/119 (68%) of units had higher evoked responses for contrast increments when they were paired with optogenetic stimulation. The distributions of firing rate changes were significantly different from one another (increments median = +0.39 ΔSpikes/s, -0.07-1.07 IQR; visual + optogenetic trials median = +1.23 ΔSpikes/s, 0.33-2.6 IQR, $p < 10^{-4}$, Komolgorov-Smirnov test). (E-G, top) Same as in A-C, except for optogenetic stimulation of PV interneurons. (E-G, bottom) Responses of the same units as above except for contrast increments. (H, top) Scatter plot shows the evoked change in firing rate for each unit recorded in PV-Cre mice (n = 131 units) in response to moderate contrast decrements without (x-axis) and with (y-axis) optogenetic stimulation of PV interneurons. Optogenetic stimulation reduced the response to contrast decrements for most units. 105 of 131 (80%) units had lower evoked responses for trials with optogenetic stimulation and the distributions of firing rate changes were significantly different (decrements median = +1.04 ΔSpikes/s, 0.25-1.83 IQR; visual + optogenetic trials median = -0.66 ΔSpikes/s, -2.9-0.48 IQR, $p < 10^{-13}$, Komolgorov-Smirnov test). Similar effects were observed for contrast increments (H, bottom). 102 of 131 (78%) of units had lower evoked responses for contrast increments paired with optogenetic activation of PV interneurons compared to the same increment without optogenetic stimulation. (increments median = +0.66 ΔSpikes/s, 0.19-1.13 IQR; visual + optogenetic trials median = -0.94 ΔSpikes/s, -3.45-0.29 IQR, $p < 10^{-16}$, Komolgorov-Smirnov test). I-J) The strength of optogenetic modulation differs by recording depth. As brain tissue is a highly scattering medium, the intensity of opsin illumination decays as a function of depth. We examined how the effects of optogenetic stimulation on population responses varied across recording sites located at different depths. For our recordings, we used 4 shank probes with 8 contacts per shank (A32 4x8, NeuroNexus Inc.) and recording sites were separated by 100 μm. A straightforward proxy for depth is to separate data from

superficial versus deep electrode contacts. To examine how optogenetic modulation varied by depth, we calculated the absolute change in spike rate during the stimulus window (+50 - 500 ms) relative to the pre-stimulus period (-500 to -50 ms) for ±12% contrast changes (increments and decrements). This was done separately for trials with and without optogenetic stimulation. Using the absolute change in spike rate allowed us to combine data from PV-Cre and Emx-Cre mice as the effects of optogenetic stimulation on neuronal activity were opposite in sign. We then subtracted the absolute change in spiking on trials without optogenetic stimulation from the response on trials with stimulation to isolate the contribution of optogenetic input to spiking during the stimulus period. Thus, each unit contributed two data points in this dataset (contrast increments, decrements). This metric captures the absolute difference in spike rate induced by optogenetic stimulation above what would be expected based on the visual stimulus alone for each unit (n = 250 units, 4 Emx-Cre, 4 PV mice). The data were then grouped based on which contact the units were recorded from, starting with the most superficial (1+2) to the deepest (7+8). As would be expected from light scattering through tissue, optogenetic effects on unit responses were strongest for the most superficial contacts and appeared to fall off with increasing depth (mean ΔSpikes/s: sites 1&2: 2.0, 0.42 SEM; sites 3&4: 2.7, 0.43 SEM; sites 5&6: 1.48, 0.28 SEM; sites 7&8: 1.17, 0.25 SEM).

(I) Plot shows the mean±SEM difference in evoked response with versus without optogenetic input for all units recorded at different depths. We found that the strength of optogenetic modulation significantly depended on depth ($p < 0.001$, Kruskal-Wallis test). *Post hoc* comparisons showed the optogenetic modulation was significantly greater for sites 3&4 (versus 5&6 and 7&8, both $p < 0.01$; trend for site 1&2, $p = 0.07$; Dunn–Šidák correction). (J) The depth related differences in the strength of optogenetic modulation was in close agreement with the probability of recording units significantly modulated by the LED. Units were classified as optogenetically responsive if their firing rate during the stimulus period on trials with optogenetic stimulation was significantly different from its' firing rate during the stimulus window on trials without optogenetic input ($p < 0.05$, Wilcoxon signed rank test, see Main Text). The probability of recording from an optogenetically responsive unit differed by depth (sites 1&2: 56%, 29 of 52 units; sites 3&4: 65%, 24 of 37 units; sites 5&6: 50%, 38 of 77 units; sites 7&8: 38%, 32 of 84 units). This pattern of results is consistent with expectations. In addition to the contribution of light scattering with depth, our viral injections were targeted to between 250-400 μm from the cortical surface. Moreover, PV interneurons are most concentrated in layer 4 (Tremblay et al., 2016) while Emx-Cre mice generate widespread transgene expression in all layers below layer 1 (Madisen et al., 2012). Together, these observations suggest our optogenetic manipulations most strongly affected responses in superficial cortical layers (likely II/III) but units were affected by optogenetic stimulation throughout cortex.

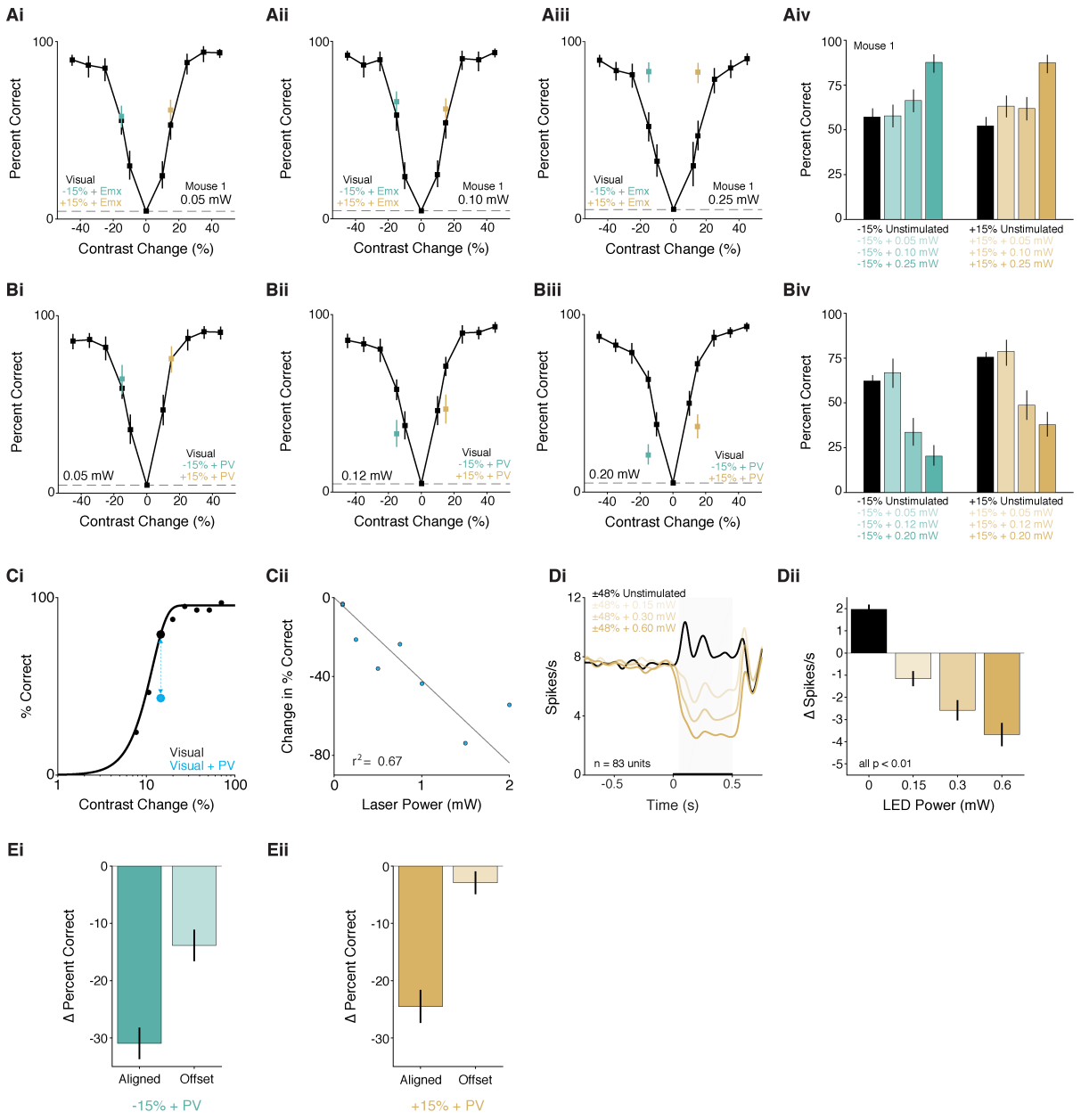

Figure S4. Optogenetic effects on perceptual reports and V1 spiking depend on power. Perceptual impairments produced by PV stimulation depend on retinotopic alignment with the visual stimulus representation. A) Data come from an Emx-Cre mouse trained to detect interleaved increases and decreases in contrast. We tested whether the effects of excitatory neuron activation on contrast perception depends on the stimulation power. Cumulative psychometric functions depict the probability of detection across a range of contrast changes. A subset of ±15% contrast changes were paired with optogenetic stimulation of excitatory neurons and different powers were used on different testing sessions (0.05 mW (Ai, 8 sessions), 0.15 mW (Aii; 8 sessions), 0.25 mW (Aiii, 7 sessions)). We combined the data across all sessions

for the tested contrast changes and performed a logistic regression that compared the probability of hits as a function of stimulation power independently for increments and decrements. We found that the probability of successfully detecting either increments or decrements in contrast was significantly influenced by the power used for optogenetic stimulation (Aiv, Decrements hit rate ± 95% CI: 0 mW = 54%, 48-59%; 0.05 mW = 58%, 51-63%; 0.15 mW = 66%, 60-72%; 0.25 mW = 83%, 77-88%; t-statistic = 5.47, $p < 10^{-7}$; Increments hit rate ± 95%: 0 mW = 49%; 43-55%; 0.05 mW = 61%, 55-67%; 0.15 mW = 62%, 56-68%; 0.25 mW = 82%, 77-88%: t-statistic = 5.48, $p < 10^{-7}$). Thus, the degree to which optogenetic stimulation of excitatory neurons improves detection of both increments and decrements in contrast depends on the power used for stimulation.

B) Same as in A except for a single PV-Cre mouse trained to detect interleaved increases and decreases in contrast. Different powers were used on different testing sessions (0.05 mW (Bi, 7 sessions), 0.12 mW (Bii; 7 sessions), 0.2 mW (Biii, 8 sessions)). We combined the data across all sessions for the tested contrast changes and performed a logistic regression that compared the probability of hits as a function of stimulation power independently for increments and decrements. We found that the probability of successfully detecting either increments or decrements in contrast was significantly influenced by the power used for optogenetic stimulation (Biv, Decrements hit rate ± 95% CI: 0 mW = 60%, 57-64%; 0.05 mW = 64%, 55-72%; 0.12 mW = 33%, 25-41%; 0.2 mW = 21%, 16-27%; t-statistic = -10.5, $p < 10^{-25}$; Increments hit rate ± 95% CI: 0 mW = 73%, 70-76%; 0.05 mW = 76%, 68-83%; 0.12 mW = 47%, 39-55%; 0.2 mW = 37%, 30-44%; t-statistic = -8.34, $p < 10^{-16}$). Thus, the magnitude of the impairment produced by PV-mediated inhibition scales with power.

(Ci) Representative behavioral performance from a single contrast decrement session where one brief contrast decrement (100 ms) was paired with stimulation of PV interneurons. Optogenetic power of 0.5 mW was the stimulation intensity used in this example session. Filled dots represent performance for trials with (blue) and without (black) activation of PV interneurons as the magnitude of contrast decrement is varied. Solid line depicts performance on visual only trials fit with a Weibull function. Blue dashed line connecting dots shows magnitude of perceptual impairment on the selected contrast decrement with (blue) without (black) optogenetic activation of PV neurons. (Cii) The perceptual impairment induced by PV activation scales with optogenetic stimulation intensity. Individual points depict performance impairment from individual sessions (note: there are two sessions at 0.1 mW). Gray line is a maximum likelihood linear regression anchored at the origin ($r^2 = 0.67$).

D) In a subset of electrophysiology sessions in PV mice (n=3; 12 recording sites), we collected additional datasets to assess whether visually evoked changes in firing rate across the population depended on power used for PV interneuron activation. In these datasets, we presented only the largest contrast changes (±48% contrast changes from a 50% contrast static Gabor) so as to elicit the strongest visual responses. In addition to trials without optogenetic stimulation, we also presented visual stimuli concurrently during optogenetic activation of PV neurons. We tested 3 different stimulation powers (0.15, 0.3, 0.6 mW), the highest of which was ~2x greater than the highest power used

in behavioral experiments. Visual and optogenetic stimuli were presented for 500 ms. (Di) Population PSTH (n=83 units) combined for both contrast change types (±48%, each neuron contributes both an increment and decrement response to each trace while the color scheme follows the increment convention used in the manuscript). Spikes were convolved with a Gaussian with σ = 25 ms. Black line depicts the population response without optogenetic activation of PV interneurons. Darkness of the colored lines depicts the population response for different PV stimulation powers. (Dii) PV-mediated suppression of firing scales with power. Bar plot compares the change in firing rate from +50 to +500 ms after stimulus onset relative to the pre-stimulus period (-500 to -50 ms). The strength of PV-mediated suppression significantly depends on LED power ($p < 10^{-34}$, Friedman's test; post hoc all comparisons $p < 0.01$, Dunn–Šidák correction). These observations held when increments or decrements were tested in isolation (Decrements: $p < 10^{-15}$, post hoc: 0.15 mW versus 0.3 mW $p > 0.05$, otherwise all other comparisons $p < 0.01$; Increments: $p < 10^{-16}$, *post hoc*: 0.3 mW versus 0.6 mW $p > 0.05$, otherwise all other comparisons $p < 0.05$; both Friedman's test with Dunn–Šidák correction).

E) The effects on PV stimulation on task performance depend on the alignment between the visual stimulus and the inhibited patch of visual cortex. In a subset of PV mice (n=3), we collected additional behavioral sessions (n=22) in which we moved the visual stimulus away from the retinotopic location of optogenetic stimulation. In these additional sessions (n=22 sessions), we offset the stimulus by 15-20° from the location used for primary data collection (n=25 sessions) while the LED power remained the same. Thus, despite comparable levels of PV activation, offsetting the visual stimulus should attenuate suppression of stimulus-evoked activity in V1. Indeed, optogenetic perturbations produced larger changes in performance when the visual stimulus was aligned with the stimulated patch of V1 (ΔPercent Correct = stimulated – unstimulated). This was true for both decrements (Ei, median ΔPercent Correct, Aligned: -31.2%, -0.41 - -0.17 IQR; Offset: -11%, -0.21 - -0.03 IQR, $p < 10^{-4}$; left) as well as increments (Eii, median ΔPercent Correct, Aligned: -22%, -0.35 - -0.13 IQR; Offset: -0.02%, -0.09 - 0.05 IQR, $p < 10^{-5}$; right). Thus, the ability of PV stimulation to reduce behavioral responses depends on alignment between the optogenetic stimulus and the visual representation in V1. This argues against the possibility that PV stimulation is somehow affecting performance by distracting the animal or impairing motor planning or execution.

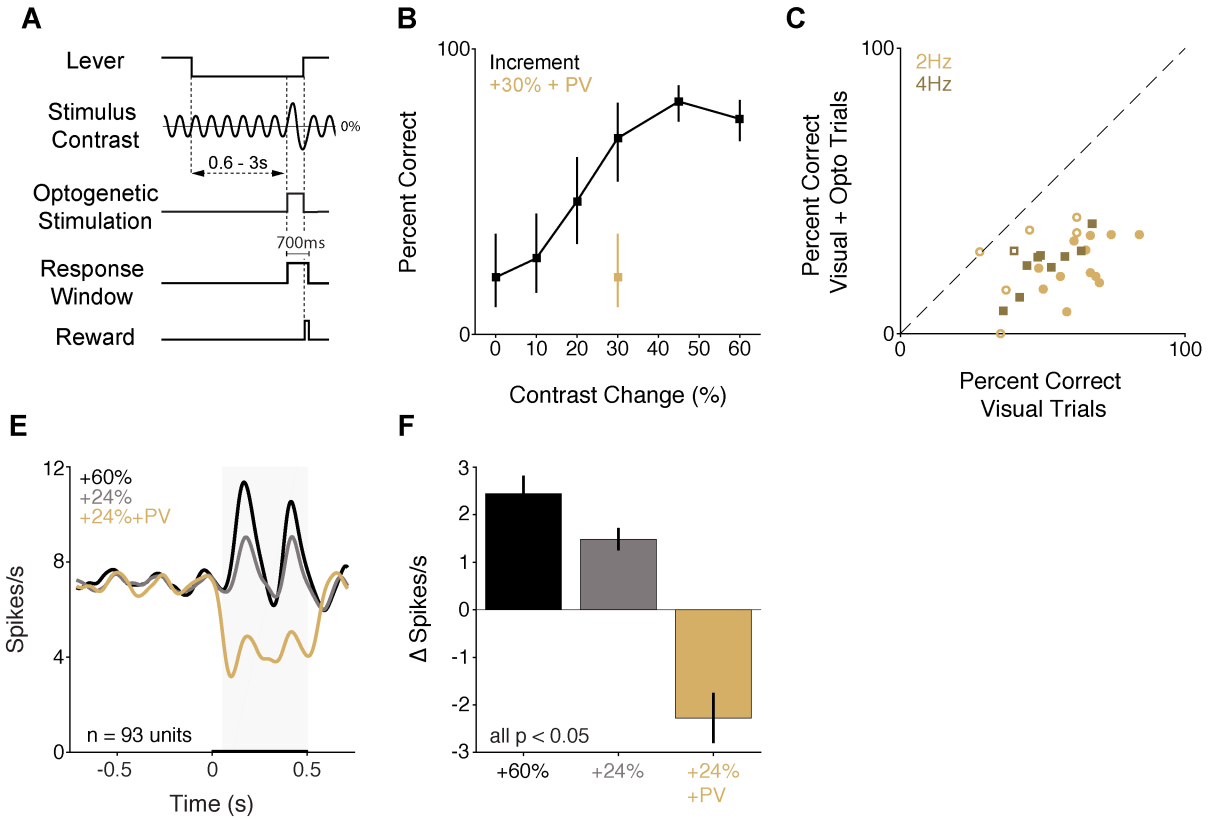

Figure S5. Optogenetic stimulation of PV interneurons impairs detection of contrast increments presented on a counterphase modulated Gabor. We sought to test if the effects of PV stimulation on performance depended on visual adaptation. In our main experiment, a static 50% Gabor stimulus was always present on the video display, which can attenuate visual responses due to adaptation (Blakemore and Campbell, 1969; Maffei et al., 1973; Movshon and Lennie, 1979). We trained PV mice (n=2; both female; 1 was retrained after completing the main experiment) to detect contrast increments of a Gabor stimulus that was counterphase modulated (2 or 4 Hz).

(A) Trial schematic. The average contrast of the Gabor stimulus was held at 20%. Contrast changes were synchronized with the zero crossings of the phase modulation and were randomly selected from values spanning threshold. We reasoned that the counterphase modulation would drive the V1 population into higher rates of firing, therefore providing a larger pedestal on which to inhibit neuronal activity. We optogenetically activated PV interneurons on a random half of trials for a moderate contrast increment (+30%). As before, optogenetic stimulation was delivered starting at stimulus onset through the end of the trial. (B) Representative single session performance (hit rate ± 67% CI) from a mouse trained to detect contrast increments of counterphase modulated stimuli. Across all sessions (2 mice, 28 sessions; for 2Hz (n=18 sessions) and 4Hz (n=10 sessions), the proportion of trials in which mice successfully detected a 30% contrast increment was significantly lower on trials with optogenetic stimulation of PV neurons (median = 57.1% versus 26.9; $p < 10^{-5}$, Wilcoxon signed rank test). Performance was impaired for either counterphase modulation rate

when tested in isolation (2 Hz: median = 61.5% versus 25.7, $p < 10^{-3}$; 4 Hz: median = 47.8% versus 26.7 $p < 0.01$; both Wilcoxon signed rank test). For both mice, we found a significant decrease in the cumulative proportion of hits relative to misses on trials with PV stimulation (Table S3; 2 Hz: both mice at least $p < 10^{-11}$; 4 Hz: both mice at least $p < 10^{-10}$; all Fisher's exact test) (C) Symbols depict the percent correct in individual behavioral sessions (2 mice, 28 total sessions) with (y-axis) and without (x-axis) optogenetic stimulation, separately for sessions in which the counterphase modulation frequency was 2 Hz (gold circles; 18 sessions) or 4 Hz (brown squares; 10 sessions). Filled circles indicate a significant change in detection performance (21/28 sessions; 2 Hz: 12/18 sessions $p < 0.05$; 4 Hz: 9/10 sessions $p < 0.05$; Fisher's exact test).

(D) We recorded neuronal activity in passively viewing, awake mice while presenting contrast increments on a background of a 20% contrast counterphase modulated (2 Hz) Gabor stimulus. We recorded a total of 93 units across 9 sites in 3 PV mice. Optogenetic stimulation was delivered on a near-threshold contrast change (+24%) concurrently with the onset of the visual stimulus. Modulating the stimulus generated a higher pedestal of spontaneous and stimulus-evoked activity in V1. We calculated the average firing rate during the 500 ms preceding the onset of the visual stimulus across all recorded units when mice viewed either counterphasing or static stimuli. Baseline firing rates were significantly elevated for a 20% contrast counterphase modulated stimulus compared to the 50% contrast static stimulus used in the main experiments (median firing rate 5.69 spikes/S, IQR 2.4-9.0 versus 4.22, IQR 2.1-7.9; $p < 0.05$, Kruskal-Wallis test). Very few units were significantly inhibited by counterphase modulated contrast increments (6.5%, 6/93 units, $p<0.05$ relative to baseline, Wilcoxon's signed rank test), whereas the vast majority of responsive units were excited (43%, 40/93 units).

The population response to contrast increments was strongly positive. However, when contrast changes were paired with optogenetic stimulation of PV neurons, the average population signal changed sign (gold). Traces depict average Gaussian filtered (σ = 25 ms) PSTH in response to contrast increments from all units recorded with counterphase modulated in PV mice. Optogenetic stimulation produces a robust decrease in firing that is comparable in magnitude, but opposite in sign, compared to responses evoked by large contrast increments. 75 of 93 (80%) units had lower evoked responses for trials with PV stimulation and distributions of firing rate changes are significantly different (visual trials median = +1.03 ΔSpikes/s, 0.2-2.0 IQR; visual + optogenetic trials median = -0.87 ΔSpikes/s, -4.0-0.78 IQR, $p < 10^{-9}$, Komolgorov-Smirnov test). Gray square depicts the analysis window (+50 - +500 ms) used for spike rate quantification. (E) Average change in spike rate compared to the time matched baseline period for all recorded units (stimulus – baseline). The population responses were significantly different across conditions ($p < 10^{-53}$; Friedman's test). Post hoc tests indicated that the population response in each condition was significantly different from all other stimulus configurations (Dunn–Šidák correction, all comparisons $p < 0.05$). Thus, decreasing V1 spike output on the background of a dynamic visual stimulus and higher background levels of activity impairs contrast change detection. These results are consistent with our primary findings.

	2 Hz			4 Hz		
Mouse	**Hits/Miss unstim**	**Hits/Miss stim**	**p-value**	**Hits/Miss unstim**	**Hits/Miss stim**	**p-value**
1	146/103	55/170	$< 10^{-15}$	273/221	129/371	$< 10^{-20}$
2	150/120	70/187	$< 10^{-11}$	167/236	83/313	$< 10^{-10}$

Table S3. Summary of changes in detection performance produced by optogenetic stimulation of PV interneurons for counterphase modulated visual stimuli. We combined data across all sessions at the tested contrast change (+24%) for sessions in which the counterphase modulation rate was 2 Hz (left) or 4 Hz (right) and calculated the relative proportion of hits and misses separately for trials with and without optogenetic stimulation of PV interneurons. Cumulative data from the two PV mice is shown above, numbers correspond to the total counts of hits/misses for contrast increments of a 2 Hz or 4 Hz counterphase modulated stimulus. P-values correspond to the results of Fisher's exact test comparing the proportions of hits/misses with and without stimulation. Optogenetic stimulation of PV interneurons similarly impaired detection of contrast increments for both modulation rates.

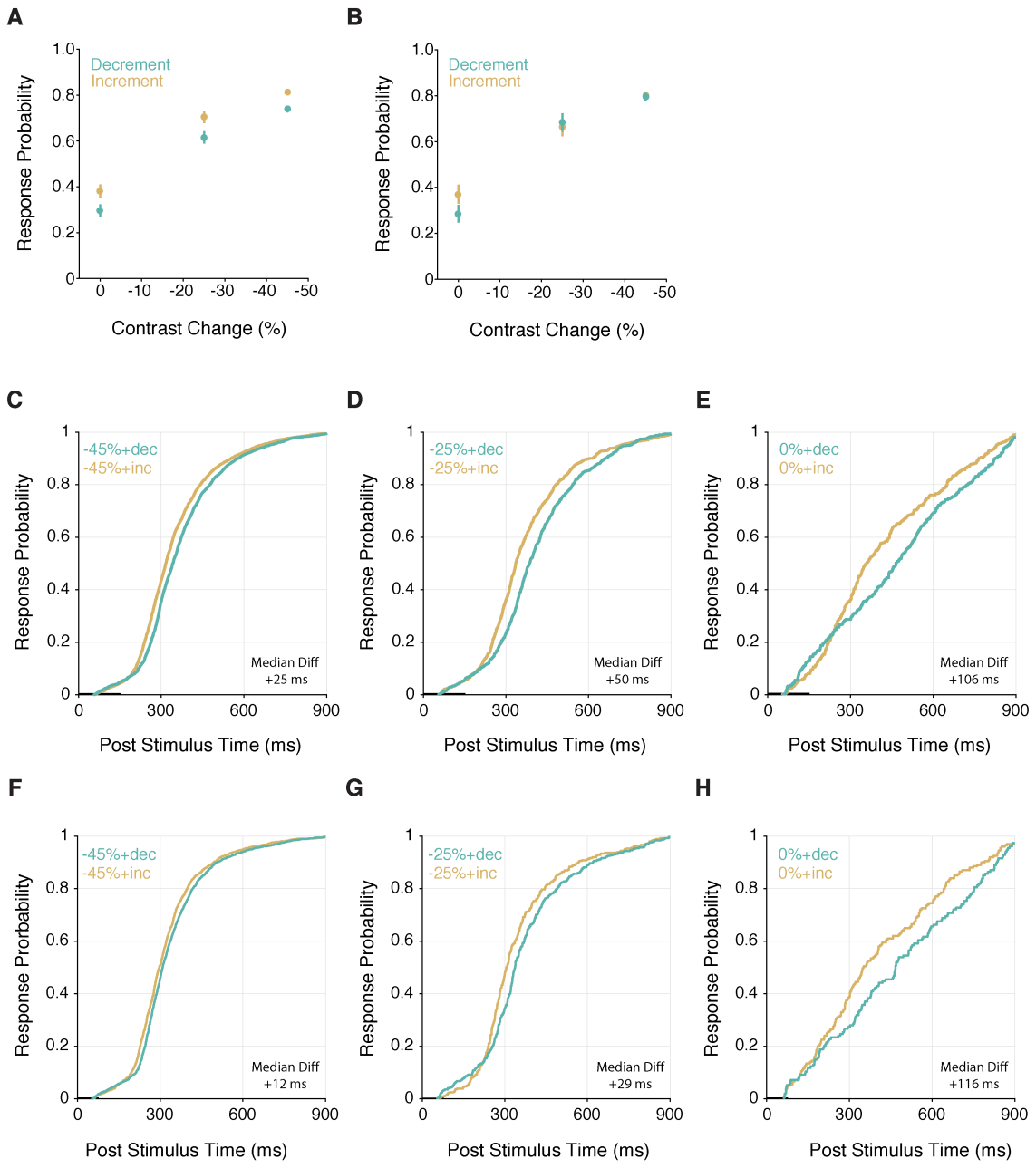

Figure S6. Increments versus decrements in excitatory input produce dissociable effects on the probability of behavioral responses. In addition to trials with only changes in optogenetic stimulation, many trials featured increments and decrements in optogenetic input that were concurrently presented with large and moderate contrast decrements. We found differential effects on behavioral responses to contrast changes when those changes were paired with increases versus decreases in optogenetic input.

A) Points represent probability of a response within the 900 ms reaction time window for different visual contrast decrement magnitudes depending on whether optogenetic input decremented (aqua) or incremented (gold) for 150 ms before returning to the baseline power (n = 31 sessions in 3 mice). Mice consistently responded more when the optogenetic input incremented than when it decremented. (Response probability, -45%: increment = 0.81, 0.80-0.83 95% CI, decrement = 0.74, 0.72-0.75 95% CI; -25%: increment = 0.70, 0.68-0.73 95% CI, decrement = 0.62, 0.59-0.64 95% CI; 0%: increment = 0.38, 0.35-0.41 95% CI, decrement = 0.30, 0.27-0.32 95% CI; Effect of Contrast, $p < 10^{-20}$; Effect of Optogenetic step direction, $p < 10^{-5}$; Interaction, p = 0.39; logistic regression). B) Same as in A, except for optogenetic pulses of 75 ms (n = 19 sessions). Comparing across stimuli, shorter pulses had no effect on the proportion correct (Response probability, -45%: increment = 0.80, 0.78-0.82 95% CI, decrement = 0.79, 0.78-0.81 95% CI; -25%: increment = 0.66, 0.64-0.72 95% CI, decrement = 0.68, 0.64-0.72 95% CI; 0%: increment = 0.37, 0.32-0.41 95% CI, decrement = 0.28, 0.25-0.32 95% CI; Effect of Contrast, $p < 10^{-20}$; Effect of Optogenetic step direction, p = 0.28; Interaction, p = 0.06; logistic regression). However, when tested in isolation, the response probability for optogenetic increments was significantly greater than decrements when the visual stimulus did not change ($p < 0.01$; logistic regression). This shows that 75 ms optogenetic pulses were still sufficient to augment the response probability when presented in isolation. C-H) Cumulative reaction time distributions on trials in which mice released the lever during the reaction time window (900 ms) following contrast changes presented concurrently with optogenetic increments and decrements. C-E) Cumulative distributions for sessions in which the optogenetic pulse duration was 150 ms when contrast changes were large (C), moderate (D), or zero (E). Trials with optogenetic increments speed responses compared to trials with optogenetic decrements (median reaction time, -45%: increment = 317 ms, 253-413 IQR, decrement = 342, 278-439 IQR, $p < 10^{-12}$; -25%: increment = 335 ms, 266-455 IQR, decrement = 342, 308-503 IQR, $p < 10^{-9}$; 0%: increment = 358 ms, 243-583 IQR, decrement = 464, 249-665 IQR, $p< 0.01$; all Kruskal-Wallis test). Optogenetic increments presented without contrast changes (panel E) produce an increase in response probability compared to decrements. Thick portion of the x-axis indicates the LED pulse duration, while the visual stimulus change remained on the screen for the duration of the reaction time window (900 ms). Median Diff = difference in medians between the two reaction time distributions (increment – decrement). (F-H) Same as C-D except for sessions in which the LED pulse duration was reduced to 75 ms (19 sessions in the same 3 mice). While the distributions when visual stimuli are present are significantly different, the differences between the distributions are smaller in magnitude compared to the same stimuli presented with longer optogenetic pulse durations (median reaction times on hits, -45%: increment = 297 ms, 240-371 IQR, decrement = 309, 256-395 IQR, $p < 10^{-4}$; -25%: increment = 306 ms, 252-403 IQR, decrement = 337, 308-503 IQR, $p < 0.01$; 0%: increment = 351 ms, 220-603 IQR, decrement = 467, 276-716 IQR, $p< 0.05$; all Kruskal-Wallis test). The smaller difference between reaction time distributions is consistent with the smaller integrated difference in LED input with shorter pulses. (H) When contrast changes are absent, 75 ms optogenetic increments produce a visible increase in response probability compared to decrements.

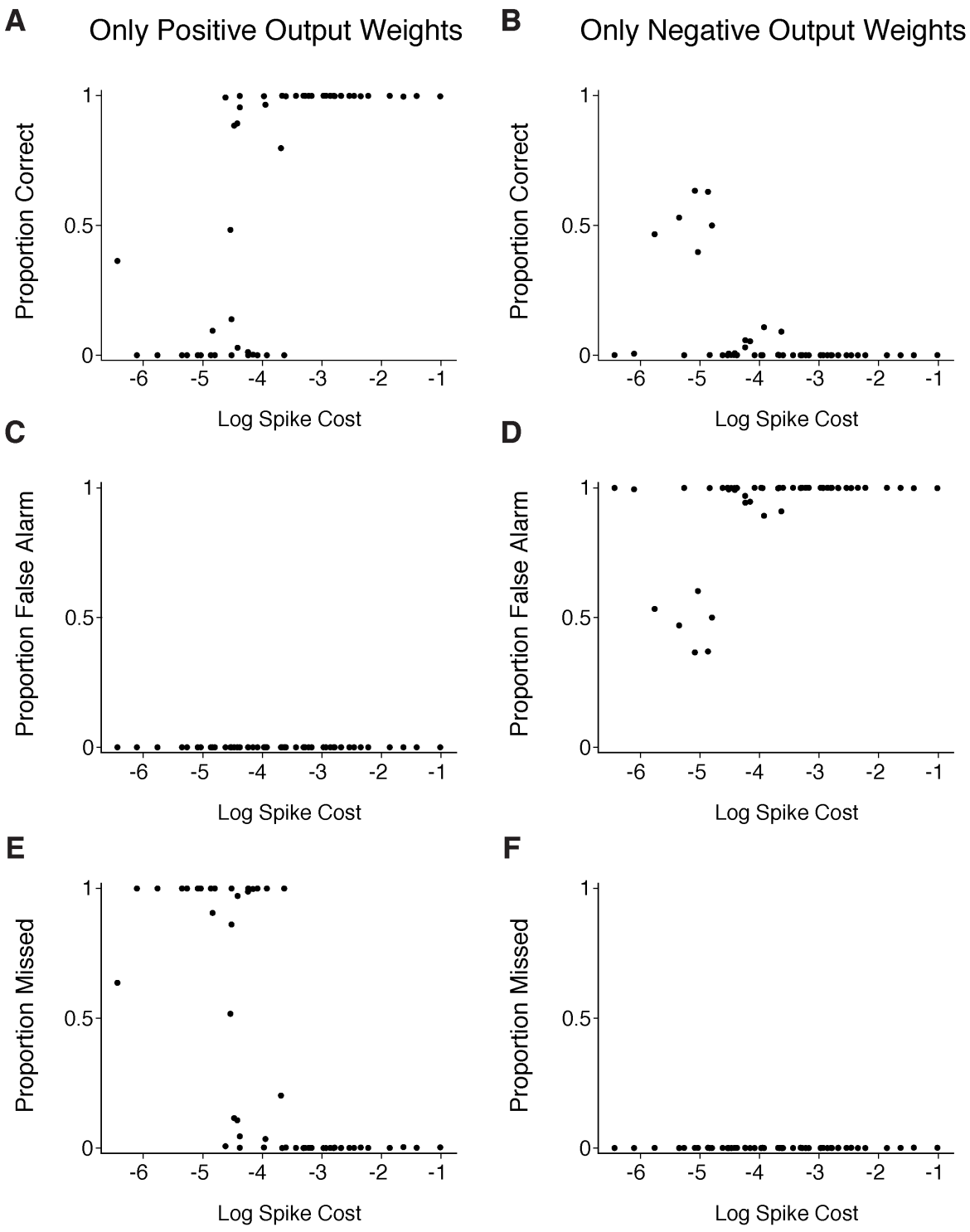
A
Only Positive Output Weights
B
Only Negative Output Weights
Proportion Correct
Proportion False Alarm
Proportion Missed
Log Spike Cost
C
D
E
F
1
0.5
0
-6
-5
-4
-3
-2
-1

Figure S7. Contributions of positive and negative output weights to task performance in RNNs. Following training, networks were presented with new trials and either positive or negative connections from the recurrent layer to the release neuron were shut off. (A-B) Proportion of successfully detected contrast changes for as a function of spike cost when trained networks must perform the task using exclusively positive (A) or negative (B) output weights from the recurrent layer. Each point represents the performance of a single trained network. RNNs perform well using positive but not negative output weights across a range of spike costs. (C-D) Proportion of false alarms as a function of spike cost when RNNs use only positive (C) or negative (D) output weights. Networks restricted to exclusively negative weights exhibit high false alarm rates. (E-F) Proportion of contrast changes that networks failed to detect (misses) as a function of spike cost for RNNs restricted to exclusively positive (E) or negative (F) output weights.

Supplemental References:

Blakemore, C., and Campbell, F.W. (1969). On the existence of neurones in the human visual system selectively sensitive to the orientation and size of retinal images. J. Physiol. *203*, 237–260.

Madisen, L., Mao, T., Koch, H., Zhuo, J., Berenyi, A., Fujisawa, S., Hsu, Y.-W.A., Garcia, A.J., Gu, X., Zanella, S., et al. (2012). A toolbox of Cre-dependent optogenetic transgenic mice for light-induced activation and silencing. Nat. Neurosci. *15*, 793–802.

Maffei, L., Fiorentini, A., and Bisti, S. (1973). Neural correlate of perceptual adaptation to gratings. Science *182*, 1036–1038.

Movshon, J.A., and Lennie, P. (1979). Pattern-selective adaptation in visual cortical neurones. Nature *278*, 850–852.

Nissen, M.J. (1977). Stimulus intensity and information processing. Percept. Psychophys. *22*, 338–352.

Roitman, J.D., and Shadlen, M.N. (2002). Response of Neurons in the Lateral Intraparietal Area during a Combined Visual Discrimination Reaction Time Task. J. Neurosci. *22*, 9475–9489.

Tremblay, R., Lee, S., and Rudy, B. (2016). GABAergic Interneurons in the Neocortex: From Cellular Properties to Circuits. Neuron *91*, 260–292.