Brief stimuli cast a long-term trace in visual cortex

Matthias Fritsche[1,*], Samuel G. Solomon[2] & Floris P. de Lange[1]

[1] Donders Institute for Brain, Cognition and Behaviour, Radboud University, Kapittelweg 29, 6525 EN Nijmegen, the Netherlands

[2] Institute of Behavioural Neuroscience, Department of Experimental Psychology, University College London, London WC1H 0AP, United Kingdom

* Correspondence: m.fritsche@donders.ru.nl

Abstract

Sensory processing and perception are strongly influenced by recent stimulus history – a phenomenon termed adaptation. While perception can be influenced even by brief stimuli presented dozens of seconds ago, neural adaptation to brief stimuli has not been observed beyond time lags of a few hundred milliseconds. Here, using an openly available dataset from the Allen Brain Observatory, we show that neurons in the early visual cortex of the mouse exhibit remarkably long timescales of adaptation in response to brief visual stimuli, persisting over dozens of seconds, despite the presentation of several intervening stimuli. Long-term adaptation was selectively expressed in cortical, but not in thalamic neurons, which only showed short-term adaptation. Visual cortex thus maintains concurrent stimulus-specific memory traces of past input that enable the visual system to build up a statistical representation of the world over multiple timescales, to efficiently represent information in a changing environment.

Introduction

Sensory processing and perception do not only depend on the current sensory input, but are influenced by the recent stimulus history. This is exemplified by perceptual illusions such as the tilt-aftereffect, in which the perceived orientation of a grating is repelled away from a previously seen grating (Clifford et al., 2000; Gibson & Radner, 1937; Magnussen & Johnsen, 1986). Likewise, neurons in visual cortex change their responsivity and stimulus preferences following the presentation of oriented gratings, commonly referred to as neural adaptation (Dragoi et al., 2000, 2001; Kohn & Movshon, 2003, 2004; Müller et al., 1999). Indeed, adaptation-induced changes of neural responses in early visual cortex are consistent in sign and magnitude with the observed changes in perception, thus providing a potential neural basis of the perceptual bias (Clifford et al., 2000; D. Z. Jin et al., 2005; Wainwright, 1999). However, recent studies point towards an apparent disparity between the timescales over which neural and perceptual adaptation effects are observed. Electrophysiological studies in macaque monkeys and mice have found neural adaptation in response to brief stimuli to be short-lived, rarely observable beyond time lags of a few hundred milliseconds (M. Jin et al., 2019; M. Jin & Glickfeld, 2020; Kim et al., 2019; Patterson et al., 2013). In contrast, recent psychophysical studies have revealed long-lived repulsive perceptual biases following briefly presented gratings, biasing subsequent orientation perception over dozens of seconds and despite intervening visual input (Chopin & Mamassian, 2012; Fritsche, Spaak, et al., 2020; Gekas et al., 2019; Suárez-Pinilla et al., 2018). If these long-lived perceptual biases are subserved by neural adaptation in early visual areas, then similarly long timescales of adaptation, robust to intervening visual input, should be measurable in those areas. Importantly, however, neural adaptation of sensory systems in response to brief sensory stimulation has hitherto not been observed over such long timescales.

In order to test whether neurons in early visual areas exhibit long-term adaptation in response to brief stimuli, and despite intervening visual input, we leveraged a large dataset of electrophysiological recordings in the mouse visual system (Allen Brain Observatory – Neuropixels Visual Coding, Siegle et al., 2021). We characterized the recovery time course of adaption in response to brief drifting and static grating stimuli across the visual system of awake mice. Neurons in the mouse primary visual cortex exhibit selectivity for orientation (Liu et al., 2011; Niell & Stryker, 2008; Tan et al., 2011), and undergo orientation-specific adaptation, tuned to the orientation difference between previous and

current stimulus (M. Jin et al., 2019). This makes the mouse visual system, in principle, suitable for probing the timescales of orientation-specific adaptation. The use of high density extracellular electrophysiology probes (Jun et al., 2017) further enabled us to study the temporal dynamics of adaptation across multiple brain areas across the visual hierarchy, most notably the thalamus, primary and extrastriate visual cortex. It has been previously proposed that temporal integration timescales increase along the cortical hierarchy (Hasson et al., 2008; Honey et al., 2012; Lerner et al., 2011; Murray et al., 2014). Beyond testing whether neurons in early visual areas exhibit long-term adaptation, we therefore further investigated whether a similar hierarchy of temporal dynamics may exist for adaptation in the mouse visual system.

To preview, we found remarkably long timescales of adaptation in response to brief visual stimuli in cortical visual areas, persisting over dozens of seconds, despite the presentation of several intervening stimuli. While decay of adaptation was long-lived across primary and extrastriate visual cortex, neurons in the thalamus only showed short-lived adaptation, limited to the processing of adjacent stimuli. Long-term adaptation in visual cortex is thus not inherited from the thalamus, but likely relies on cortical short-term plasticity. This long-term adaptation was also evident after the exposure to more rapidly presented brief static gratings, underlining the robustness of the long-term temporal dependencies. Our findings therefore point towards adaptation in early visual cortex as a potential neural source of long-lived perceptual aftereffects discovered in recent psychophysical experiments, and more broadly highlight the multiple timescales through which stimulus history affects current sensory processing. This dependence on the broader temporal context may enable the visual system to efficiently represent information in a slowly changing environment (Schwartz et al., 2007; Weber et al., 2019).

Results

Orientation-specific adaptation in visual cortex and thalamus

To investigate orientation-specific adaptation in the mouse visual system, we analyzed responses from a total of 2,365 visually responsive neurons in the visual cortex and thalamus of 32 mice, while they were presented with sequences of drifting gratings (**Figure 1A**). We separately analyzed visual responses to gratings that were preceded by a grating of the same orientation (*repeat*) or orthogonal orientation (*orthogonal*). We found that the immediate repetition of stimulus orientation led to a marked reduction in spiking activity in primary visual cortex (V1), predominantly during the early visual response (0 – 100 ms from stimulus onset, n = 562; **Figure 1B**, green shaded area). We quantified this adaptation of the transient visual response by calculating the response to a repeated orientation, relative to that following the orthogonal orientation (*1-back adaptation ratio*, see **Materials & Methods**). Adaptation reduced the response by 17% in V1 (1-back adaptation ratio: 0.83, p = 4e-57, 95% CI [0.81, 0.84]), and had similar impact in higher-level extrastriate visual areas (**Figure 1C**; 1-back adaptation ratios between 0.80 and 0.88, all p < 2e-21, two-sided t-tests, corrected for multiple comparisons). Notably, we also found orientation-specific adaptation in the dorsolateral geniculate nucleus of the thalamus (LGN, n = 140; 1-back adaptation ratio: 0.93, p = 8e-7, 95% CI [0.91, 0.96]), and the lateral posterior nucleus of the thalamus (LP, n = 90; 1-back adaptation ratio: 0.83, p = 3e-10, 95% CI [0.79, 0.88]). The response reductions for immediate stimulus repetition were highly consistent across mice (**Supplementary Figure S1**). While 1-back adaptation was generally strongest when neurons were tested with their preferred orientation, neurons also adapted robustly to other orientations (**Supplementary Figure S2**). In our subsequent analyses of long-term adaptation, we therefore averaged adaptation across all stimulus orientations, regardless of the neurons' orientation preference, but qualitatively similar results were obtained when only considering trials in which stimuli matched a neurons' preferred orientation (see **Supplementary Figure S3** and **Supplementary Tables 1** and **2)**. Overall, these findings indicate robust orientation-specific adaptation of neurons in visual cortex and thalamus to gratings presented in the immediate past.

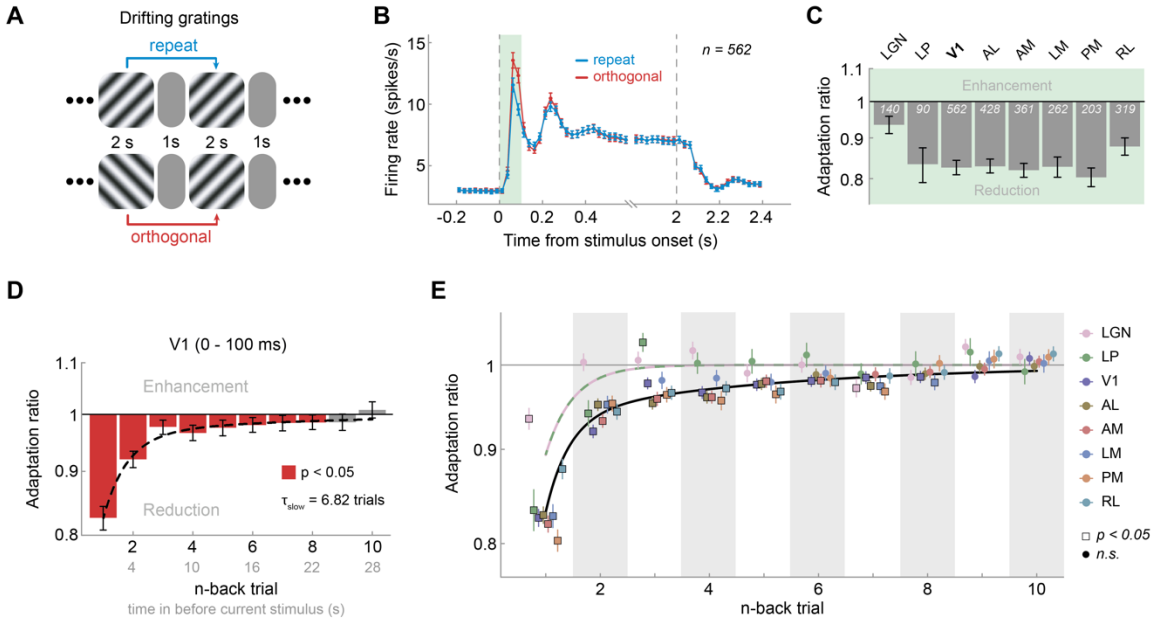


Figure 1. Visual cortex and thalamus exhibit orientation-specific adaptation to the immediately preceding (1-back) grating, but only visual cortex shows long-term adaptation. (A) Presentation sequence of drifting grating stimuli. Mice were shown drifting gratings with a duration of 2 seconds, separated by a 1-second grey screen. Gratings were drifting in one of 8 different directions (0°, 45°, 90°, 135°, 180°, 225°, 270°, 315°) and were presented in random order. For the analysis of orientation-specific adaptation, we contrasted activity to gratings preceded by gratings of the same orientation (*repeat*, blue) with that elicited by gratings preceded by a grating of the orthogonal orientation (*orthogonal*, red). **(B)** Population peristimulus time histograms of neurons in V1 for *repeat* and *orthogonal* conditions. The transient response is reduced when the same orientation is successively repeated, indicating orientation-specific adaptation. Subsequent analyses focused on this transient response (0 – 100 ms, green shaded area). Vertical dashed lines denote stimulus onset and offset, respectively. Binwidth = 25 ms. Error bars show *SEM*s. **(C)** 1-back adaptation ratios of transient responses across visual areas. Adaptation ratios were computed by dividing each neuron's firing rate for *repeat* by that for *orthogonal* stimulus presentations and therefore express the response magnitude to a repeated stimulus orientation relative to that elicited by the same stimulus orientation, but preceded by a grating with the orthogonal orientation. Adaptation ratios smaller than 1 indicate adaptation. All visual areas show significant 1-back adaptation. Error bars denote bootstrapped 95% confidence intervals. White numbers indicate the number of neurons in each area. **(D)** Adaptation ratios of neurons in V1 as a function of the n-back trial. Strongest adaptation occurred in response to the 1-back stimulus, but stimuli encountered up to 8 presentations in the past (seen 22 seconds ago) still exerted significant adaptation effect on the current visual response, despite the presentation of intervening stimuli (red bars, $p < 0.05$, corrected for multiple comparisons). The decay of adaptation over n-back trials was well captured by a double-exponential decay model with a fast- and slow-decaying adaptation component (black dashed line; a_{fast} = 13.99%, τ_{fast} = 0.85 trials, a_{slow} = 3.45%, τ_{slow} = 6.82 trials). Error bars denote bootstrapped 95% confidence intervals. Simulations showed that long-term adaptation was not a spurious effect due to short-term adaptation and the particular stimulus sequences used in the experiment (**Supplementary Figure S4**). **(E)** Adaptation ratio as function of n-back trial for different visual areas (color-coded). While adaptation decays similarly and slowly across cortical visual areas, and is generally significant for up to 6-8 trials back (square-shaped datapoints, $p < 0.05$, corrected for multiple comparisons per area), it decays more rapidly in thalamic areas LGN and LP. Black and lilac-green lines illustrate the best fitting exponential decay models for cortex and thalamus. Error bars denote standard errors of the mean. See **Supplementary Figure S3** for 10-back adaptation ratios computed for the neurons' preferred stimulus orientations only, which yielded qualitatively similar results. List of acronyms: Dorso-lateral geniculate nucleus of the thalamus (*LGN*), latero-posterior nucleus of the thalamus (*LP*), primary visual cortex (*V1*), antero-lateral area (*AL*), antero-medial area (*AM*), latero-medial area (*LM),* postero-medial area (*PM*), rostro-lateral area (*RL*).

Long-term adaptation in visual cortex but not in thalamus

In order to investigate the timescale over which adaptation influences subsequent visual processing, we computed adaptation ratios based on the orientation difference between the current grating, and gratings at different n-back timepoints. Surprisingly, we found that neurons in V1 exhibited significant adaptation effects to stimuli seen up to 8 presentations (or 22 seconds) in the past, despite the presentation of intervening stimuli (**Figure 1D**). In contrast, adaptation in the thalamus appeared to be limited to the 1-back (LGN) or 2-back trial (LP; **Figure 1E**). Indeed, adaptation to temporally remote stimuli was significantly stronger in V1 compared to LGN (2-, 4- and 9-back stimulus, two-sided Welch's unequal variances t-test, all $p = 0.02$) and LP (3-back stimulus, $p = 0.001$, corrected for multiple comparisons), even after accounting for differences in the initial strength of adaptation between thalamus and V1 (i.e. normalizing to the 1-back adaptation ratio). The temporal decay of adaptation in higher-level extrastriate areas was similar to the decay in V1 (**Figure 1E)** and long-term adaptation in cortical areas was very consistent across mice (**Supplementary Figure S1**).

We further characterized the timescale of recovery from adaptation in visual cortex and thalamus by fitting exponential decay models to the n-back adaptation ratios in the respective areas. Recovery in cortical visual areas was better explained by double-exponential decay models, with a fast and a slow-decaying adaptation component, compared to a single-exponential decay (F-tests, all $p < 0.006$, except VISrl: $p = 0.072$). Recovery from adaptation was slowest in V1 with an exponential time constant τ_{slow} of 6.82 trials (bootstrapped 95% CI [3.39, 13.50]; **Figure 1D**, black dashed line), but was relatively similar for extrastriate areas (τ_{slow} ranging from 3.12 to 5.52 trials, all 95% CI intervals overlapping, see **Supplementary Table 1** for all parameter estimates). In contrast, the recovery of adaptation in the thalamus was most parsimoniously captured by a single-exponential decay model (F-tests, $p = 1$ for both LGN and LP), and the time constants of the single-exponential decays were very short (LGN: $\tau_{fast} = 0.02$ trials, 95% CI [0.004, 0.63]; LP: $\tau_{fast} = 0.71$ trials, 95% CI [0.03, 0.93]). Together, these results indicate that adaptation in response to relatively brief 2-second stimuli decays surprisingly slowly in cortical visual areas, over the course of dozens of seconds, and survives the presentation of multiple intervening stimuli. Conversely, while adapting to the immediate stimulus history, neurons in the thalamus exhibit a fast recovery from adaptation, in line with a reduced temporal integration timescale for low- compared to high-level visual areas.

Long-term adaptation following exposure to brief static gratings

So far, we have shown that neurons in mouse visual cortex exhibit long-lived adaptation to 2-seconds long drifting gratings, influencing subsequent visual processing over the time course of dozens of seconds and multiple intervening stimuli. However, it is unclear to which degree the existence of such long-term adaptation effects depends on the particular stimulus type (drifting gratings) and duration (2 seconds). We therefore tested whether similar long-lived adaptation effects can be elicited by the presentation of brief, static gratings. Mice were presented with a rapid stream of static gratings, presented back-to-back for 250 milliseconds each (**Figure 2A**). Similar to our previous analyses, we probed orientation-specific adaptation by contrasting visual responses to gratings that were preceded by a grating of the same or orthogonal orientation. In V1, the repetition of stimulus orientation led to a clear reduction of the visual response to the current grating (**Figure 2B**, green shaded area; $n = 530$; adaptation ratio: 0.90, $p = 2e\text{-}81$, 95% CI [0.89, 0.91]). Very similar adaptation effects were found for extrastriate areas (**Figure 2C**, 1-back adaptation ratios between 0.89 and 0.93, all $p < 8e\text{-}9$, two-sided t-tests, corrected for multiple comparisons). Although there were only relatively few responsive neurons in the thalamus (LGN: $n = 60$; LP: $n = 16$), those neurons also exhibited significant adaptation effects (LGN - adaptation ratio: 0.93, $p = 0.004$, 95% CI [0.93, 0.98]; LP – adaptation ratio: 0.85, $p = 0.01$, 95% CI [0.85, 0.96]). Overall, these findings of orientation-specific adaptation, exerted by the immediately preceding static grating stimulus, parallel those found for adaptation to drifting gratings.

Next, we investigated the timescale over which adaptation to briefly presented static gratings affected subsequent visual processing. Neurons in V1 showed significant adaptation effects to stimuli presented as far as 20 presentations (5 seconds) in the past (**Figure 2D** and **Supplementary Figure S5** showing consistent adaptation across mice). Again, the decay of adaptation was well described by a double exponential decay model with a long time constant $\tau_{slow} = 9.12$ trials (95% CI [6.09, 14.82];

Figure 2D, black dashed line). Higher-level extrastriate cortical areas showed similar decay dynamics (**Figure 2E**), with decay time constants ranging from 6.54 (VISam) to 21.78 trials (VISpm; all 95% CI intervals overlapping; see **Supplementary Table 3** for all parameter estimates). While all cortical areas were significantly better fit by a double exponential decay model (F-tests, all $p < 1e-5$), neurons in LGN were more parsimoniously described ($p = 0.22$) by a single exponential decay with a shorter time constant (τ_{fast} = 2.93 trials, 95% CI [1.36, 7.32]). We did not include thalamic nucleus LP in this analysis due to the low number of visually response neurons in this area (16 neurons across 32 mice). These findings demonstrate that even briefly presented static grating stimuli, which are embedded in a rapid stream of stimulus presentations, still elicit robust long-term cortical adaptation effects. Neurons in the thalamus are less influenced, and thus provide a more context-independent input to visual cortex.

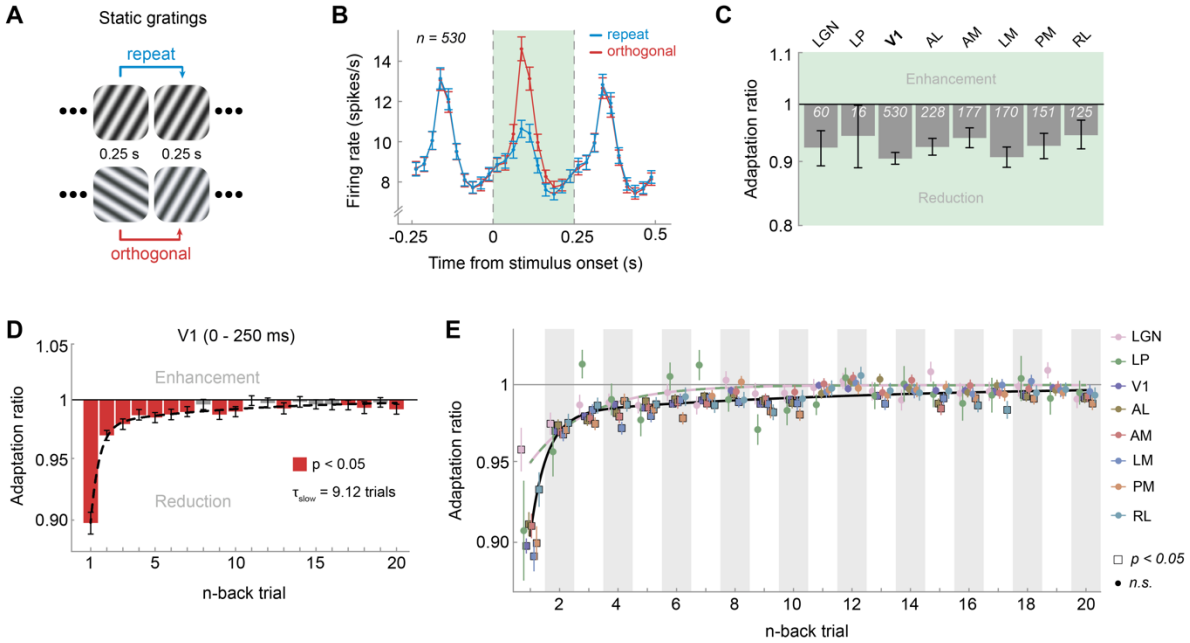


Figure 2. Visual cortex exhibits long-term adaptation following briefly presented static gratings. (A) Presentation sequence of static grating stimuli. Mice were shown static gratings with a duration of 250 ms with no intervening grey period. Gratings had one of six orientations (0°, 30°, 60°, 90°, 120°, 150°), five spatial frequencies (0.02, 0.04, 0.08, 0.16, 0.32 cycles/°), and four phases (0, 0.25, 0.5, 0.75). The order of grating presentations was randomized. Similar to the analysis of drifting gratings, we contrasted activity to gratings preceded by gratings of the same orientation (*repeat*, blue) with that elicited by gratings preceded by a grating of the orthogonal orientation (*orthogonal*, red). **(B)** Population peristimulus time histograms of neurons in V1 for *repeat* and *orthogonal* conditions. The visual response to the current stimulus (green shaded area) was reduced when the previous stimulus had the same orientation as the current stimulus (*repeat*), indicating orientation-specific adaptation. Vertical dashed lines denote onset and offset of the current stimulus, respectively. Binwidth = 25 ms. Error bars show *SEM*s. **(C)** 1-back adaptation ratios across visual areas. All areas show significant 1-back adaptation. Error bars denote bootstrapped 95% confidence intervals. White numbers indicate the number of neurons in each area. **(D)** Adaptation ratios of V1 as a function of the n-back trial. While adaptation was most strongly driven by the previous stimulus (1-back), stimuli encountered up to 20 presentations in the past (5 seconds ago) still exerted significant adaptation effects on the current visual response (red bars, $p < 0.05$, FDR-corrected). Similar to drifting grating adaptation, the decay of adaptation over n-back trials was well captured by a double-exponential decay model with a fast- and slow-decaying adaptation component (black dashed line; a_{fast} = 8.17%, τ_{fast} = 0.54 trials, a_{slow} = 2.04%, τ_{slow} = 9.12 trials). Error bars denote bootstrapped 95% confidence intervals. **(E)** Adaptation ratios as function of n-back trial for different visual areas (color-coded). In cortical areas there is significant adaptation to stimulus orientations presented up to 20 trials back (square-shaped datapoints, $p < 0.05$, FDR-corrected per area). Error bars denote standard errors of the mean. Black and orange-green lines denote the best fitting exponential decay models for cortex and thalamus, respectively. Adaptation was computed over the whole stimulus interval (0 to 250 ms), since due to

the back-to-back presentation of static gratings, visual responses to the previous stimulus overlapped with the initial time window of the current stimulus, thereby increasing response variability in this early time window. However, largely similar results were obtained when performing the analyses on the same time window used in the drifting grating experiment (0 to 100 ms), except for a less clear difference of the decay of adaptation between cortex and thalamus (**Supplementary Figure S6**).

Discussion

We observed that neurons in mouse visual cortex exhibit remarkably long timescales of adaptation effects after brief stimuli, influencing the processing of subsequent input over dozens of seconds and outliving the presentation of several intervening stimuli. The long-term adaptation effect was stimulus-specific - tuned to the orientation differences between past and current stimuli - indicating that the visual cortex maintains a lasting memory trace of individual briefly experienced stimuli. While adaptation decayed at a similar rate in primary and extrastriate visual cortex, and was still observable for stimuli seen 8 trials (or 22 seconds) in the past, adaptation in the thalamus decayed more quickly, limited to the 1- or 2-back stimulus (experienced 1-4 seconds prior). This demonstrates that the long-term component of adaptation observed in the visual cortex is not inherited from the thalamus, but is maintained in cortical circuits. Finally, we replicated our findings of cortical long-term adaptation to drifting gratings with a different stimulus set of rapidly presented static gratings, underlining the robustness of the long-term temporal dependencies. Together, our findings show that visual cortex maintains concurrent stimulus-specific memory traces of briefly presented input, which allow the visual system to build up a statistical representation of the world over multiple timescales. We speculate that this may enable the visual system to exploit temporal input regularities over extended timescales to efficiently encode new visual stimuli under natural conditions (Barlow & Földiák, 1989; Clifford et al., 2000; Müller et al., 1999; Wainwright, 1999; for reviews see Schwartz et al., 2007; Weber et al., 2019).

There is ample evidence that sensory cortex can exhibit long-term adaptation following long exposure to a stimulus. For instance, long stimulus presentations lasting from dozens of seconds to several minutes can alter visual responses of neurons in monkey and cat primary visual cortex over similarly long timescales, persisting for several minutes (Dragoi et al., 2000; Patterson et al., 2013). Furthermore, stimulus-specific adaptation effects can accumulate over many brief intermittent presentations of the same stimulus (Kuravi & Vogels, 2017), and subsequently show a persistence of several seconds (Peter et al., 2020; Ulanovsky, 2004). Crucially, in contrast to these previous studies, here we tested the adaptation effects elicited by individual, briefly presented stimuli. In the stimulus sequences of the current experiments, all stimulus orientations occurred equally often and in random order, precluding systematic accumulation of adaptation to any particular orientation. Despite the absence of such accumulation effects, we find that the presentation of brief individual stimuli alters subsequent visual processing over time spans of at least 22 seconds and affects the processing of many subsequent stimuli. This demonstrates that long-term adaptation effects are not contingent on long adaptor durations or many repeated presentations of the same adaptor stimulus, but can occur in much more naturalistic settings that are also frequently employed in experimental designs, i.e. in response to brief individual visual experiences.

The observation of long-term adaptation effects to brief stimuli is particularly surprising, as previous studies investigating the recovery of adaptation following brief visual stimulation reported only very fleeting adaptation effects. In V1 of anesthetized monkeys, adaptation to 4-seconds long drifting gratings decayed with a half-life of ~1 second, in the absence of any intervening visual input (Patterson et al., 2013). This half-life is much shorter than the ~14 seconds we observed in the current study. We speculate that this difference could be, at least partly, related to the anesthetized versus awake state of the animals in the respective experiments, and that long-term adaptation might be facilitated by deeper, recurrent stimulus processing in awake animals. Nevertheless, recent studies in awake mice point towards similar short-lived adaptation effects in V1. For instance, adaptation to 100 ms gratings has resulted in decay time constants of 0.5 to 1 second (M. Jin et al., 2019; M. Jin & Glickfeld, 2020), and other studies have found no detectible effects of adaptation to 2-seconds drifting gratings after a 6-seconds delay (King & Crowder, 2018), and no history dependencies beyond 1 second in response to 250 milliseconds orientation patterns (Kim et al., 2019). Notably, most of these previous studies have investigated adaptation in the absence of any intervening visual input, making

the current observation of stimulus specific long-term adaptation despite intervening input even more astounding. One major advantage of the current study is the large number of recorded neurons (2,365), which vastly increased our power to reveal subtle but reliable long-term adaptation effects that may have gone unnoticed in previous studies.

It has been previously proposed that temporal integration timescales increase along the cortical hierarchy (Hasson et al., 2008; Honey et al., 2012; Lerner et al., 2011; Murray et al., 2014). Here, we show that the integration window of temporal context, in the form of adaptation, increases from the thalamus to cortex, broadly in line with these proposals. However, we did not find different integration times between lower-level primary and higher-level extrastriate visual cortex, congruent with a recent study in humans (Fritsche, Lawrence, et al., 2020; but see Zhou et al., 2018). Since we measured adaptation in the early feedforward response (0 to 100 milliseconds), it appears unlikely that long-term adaptation in V1 was inherited from higher-level visual areas through feedback connections, but rather suggests that long-term temporal context already influences the earliest stages of cortical processing. The similar decay of adaptation across cortical areas could either be due to the comparatively flat hierarchical structure of mouse visual cortex (for a review see Glickfeld & Olsen, 2017) or may reflect an important difference between the temporal tuning of adaptation and previously reported temporal integration timescales.

Recent psychophysical studies in humans have revealed long-lived repulsive perceptual biases following briefly presented gratings, biasing subsequent orientation perception over dozens of seconds (Chopin & Mamassian, 2012; Fritsche, Spaak, et al., 2020; Gekas et al., 2019; Suárez-Pinilla et al., 2018). Our current findings of long-term orientation-specific adaptation in early visual cortex suggests a potential neural mechanism underlying these perceptual biases. Interestingly, a recent behavioral study in rats revealed similar long-term dependencies in a vibrissal vibration judgment task (Hachen et al., 2020), suggesting potential parallels of long-term perceptual adaptation between rodents and humans. An important future goal will be to quantitatively relate such behavioral adaptation biases to the present long-term history dependencies at the neural level.

To conclude, our findings highlight the ubiquitous influence of the short- and long-term stimulus history on current sensory processing in visual cortex. This dependence on the broader temporal context may enable the visual system to efficiently represent information in a slowly changing environment (Schwartz et al., 2007; Weber et al., 2019).

Materials & Methods

Dataset

All analyses were conducted on the openly available Neuropixels Visual Coding dataset of the Allen Brain Observatory (Siegle et al., 2021). This dataset surveys spiking activity from a large number of neurons across a wide variety of regions in the mouse brain, using high-density extracellular electrophysiology probes (Neuropixels silicon probes, Jun et al., 2017). Experiments were designed to study the activity of the visual cortex and thalamus in the context of passive visual stimulation. Here, we focused on a subset of experiments, termed the Brain Observatory 1.1 dataset. The Brain Observatory 1.1 dataset comprises recordings in 32 mice (16 C57BL/6J wild type mice and three transgenic lines: 6 Sst-IRES-Cre x Ai32, 5 Pvalb-IRES-Cre x Ai32 and 5 Vip-IRES-Cre x Ai32). The three transgenic lines were included to facilitate the identification of inhibitory inter-neuron sub-classes using opto-tagging. For the purpose of the current research question, we analyzed the data of all 32 mice, irrespective of transgenic lines. Mice were maintained in the Allen Institute for Brain Science animal facility and used in accordance with protocols approved by the Allen Institute's Institutional Animal Care and Use Committee. For a detailed description of the entire Neuropixels Visual Coding protocol see Siegle et al. (2021). All data is openly available through the AllenSDK (https://allensdk.readthedocs.io/en/latest/visual_coding_neuropixels.html).

Stimuli

During Brain Observatory 1.1 experiments, mice passively viewed a variety of different stimulus types. Here, we focused on a subset of stimuli: full-field drifting and static grating stimuli (**Figures 1A** and **2A**). Visual stimuli were generated using custom scripts based on PsychoPy (Peirce, 2007) and were displayed using an ASUS PA248Q LCD monitor, with 1920 x 1200 pixels (21.93 in wide, 60 Hz refresh rate). Stimuli were presented monocularly, and the monitor was positioned 15 cm from the mouse's right eye and spanned 120° x 95° of visual space prior to stimulus warping. Each monitor was gamma corrected and had a mean luminance of 50 cd/m^2. To account for the close viewing angle of the mouse, a spherical warping was applied to all stimuli to ensure that the apparent size, speed, and spatial frequency were constant across the monitor as seen from the mouse's perspective. For more details see Siegle et al. (2021).

Full-field drifting gratings were shown with a spatial frequency of 0.04 cycles/deg, 80% contrast, 8 directions (0°, 45°, 90°, 135°, 180°, 225°, 270°, 315°, clockwise from 0° = right-to-left) and 5 temporal frequencies (1, 2, 4, 8, and 15 Hz), with 15 repeats per condition, resulting in a total number of 600 drifting grating presentations, divided across three blocks. Drifting gratings were presented for 2 seconds, followed by a 1 second inter-stimulus interval (grey screen). Gratings of different directions and temporal frequencies were presented in random order and were interleaved by the presentation of 30 blank trials, in which only a grey screen was shown.

Static gratings were shown at 6 different orientations (0°, 30°, 60°, 90°, 120°, 150°, clockwise from 0° = vertical), 5 spatial frequencies (0.02, 0.04, 0.08, 0.16, 0.32 cycles/degree), and 4 phases (0, 0.25, 0.5, 0.75). They were presented for 0.25 seconds, with no intervening gray period. Gratings with each combination of orientation, spatial frequency, and phase were presented ~50 times in a random order, resulting in a total of 6000 grating presentations, divided across three blocks. There were blank sweeps (i.e. mean luminance gray instead of grating) presented roughly once every 25 gratings.

Data analyses

All data analyses were performed using custom code written in Python, Matlab and R. All code will be made openly available on the Donders Institute for Brain, Cognition and Behavior repository at https://data.donders.ru.nl.

Unit exclusion

To filter out units (i.e. putative neurons) that were likely to be highly contaminated or missing lots of spikes, we applied the default quality metrics of the AllenSDK. This entailed excluding units with ISI violations larger than 0.5 (Hill et al., 2011), an amplitude cutoff larger than 0.1 and a presence ratio smaller than 0.9 (for more details see https://allensdk.readthedocs.io/en/latest/_static/examples/nb/ecephys_quality_metrics.html). For the analysis of drifting gratings, we defined visually responsive units as those units whose average firing rate during the first 100 ms of stimulus presentation of the unit's preferred orientation (eliciting the highest firing rate) was larger than 5 Hz and larger than 1 *SD* of the firing rate during the first 100 ms of grey screen presentations. For the analyses of static gratings, we applied the same inclusion criteria, but computed firing rates over the whole stimulus duration (i.e. 250 ms). We chose to use a longer time window for analyzing static grating adaptation, since due to the back-to-back presentation of the static gratings, visual responses to the previous stimulus overlapped with the initial time window of the current stimulus, thereby increasing response variability in this early time window. However, largely similar results were obtained when performing the analyses on the same time window used in the drifting grating experiment (0 to 100 ms; **Supplementary Figure S6**). In order to assess whether the choice of the minimum firing rate threshold of 5 Hz had a substantial impact on our results, we repeated the analyses with a more conservative (10 Hz) and less conservative (2.5 Hz) threshold, but obtained qualitatively similar results (see **Supplementary Tables 4** and **5**). In our further investigation of sensory adaptation, we focused on those regions that contained a minimum of 50 visually responsive units (for an overview of included regions and unit counts per region see **Figures 1C and 2C**). All subsequent analyses were performed on visually responsive units only.

Orientation-specific adaptation to drifting gratings

To investigate orientation-specific adaptation, for each unit we compared firing rates in response to a current grating when this grating was preceded by a grating with the same orientation (*repeat*) or by its orthogonal orientation (*orthogonal*), irrespective of the temporal frequencies of current and previous gratings. Note that repeat trial pairs could consist of gratings with opposite drifting directions, but with the same orientation. Investigating orientation- rather than direction-specific adaptation had the advantage of maximizing the number of repeat and orthogonal trial pairs occurring across the random trialsequence. Adaptation was quantified in the form of an adaptation ratio:

$$adaptation\ ratio = \frac{fr_{repeat}}{fr_{orthogonal}}$$

where fr_{repeat} and $fr_{orthogonal}$, are the firing rates in response to a repeated and orthogonal stimulus, respectively. The adaptation ratio expresses the response magnitude to a repeated stimulus orientation relative to that elicited by the same stimulus orientation, but preceded by a grating with the orthogonal orientation. Adaptation ratios smaller than 1 indicate a relative response reduction for orientation repetitions. Importantly, this analysis quantifies orientation-specific adaptation, as the stimulus features between repeat and orthogonal condition are the same (on average), with the only difference being the relative orientation of the adaptor stimulus. An initial exploratory analysis in one mouse suggested strongest adaptation effects for the early transient response (0-100 ms from stimulus onset). Therefore, we limited our analysis to this time window. This analysis choice was made while blind to adaptation effects beyond the 1-back grating, which were of main interest to the current study. Adaptation induced by n-back gratings was quantified in a similar manner as described above, by conditioning the data on the orientation difference (*repeat* or *orthogonal*) between the current and n-back gratings. For each region, we statistically compared log-transformed adaptation ratios of 1- to 10-back gratings to zero (indicating no adaptation) using two-tailed t-tests, while controlling the false discovery rate at an alpha-level of 0.05 using the Benjamini-Hochberg procedure.

In order to quantify the recovery time course of adaptation, we fitted exponential decay models to the 1- to 50-back adaptation ratios of each region. The recovery of adaptation in cortical areas was significantly better fit by a double exponential, compared to a single-exponential decay model, with a fast and slow decay component, of the form:

$$r(n) = 1 - a_{fast} * e^{-(n-1)/\tau_{fast}} + a_{slow} * e^{-(n-1)/\tau_{slow}}$$

where r(n) denotes the adaptation ratio conditioned on the n-back stimulus orientation, a_{fast}, τ_{fast}, a_{slow} and τ_{slow} determine the magnitude and recovery time of the fast and slow adaptation component, respectively. Adaptation in thalamic regions was more parsimoniously explained by a single-exponential decay model of the form:

$$r(n) = 1 - a * e^{-(n-1)/\tau}$$

For each region, we statistically compared single- and double exponential decay models with an F-test. We used an F-test as the two decay models are nested – the single-exponential decay model is a restricted version of the double-exponential decay model. Since adaptation ratios were not normally distributed, all models were fit to log-transformed adaptation ratios by analogously log-transforming model predictions. We obtained the 95% confidence intervals of the parameter estimates with a bootstrapping procedure. In particular, for each region we resampled units with replacement and refitted the exponential decay model. We repeated this procedure 1,000 times and recorded the resulting parameter estimates of the bootstrapped sample. The 95% confidence interval was taken as the 2.5 and 97.5 percentile of the bootstrapped parameter distribution. We restricted parameter values to a wide range of plausible values ($a_{fast} = a_{slow}$ = [-Inf, 0.5], $\tau_{fast} = \tau_{slow}$= [-50, 50]), and discarded bootstrapped estimates which lay on the boundary of the parameter range, indicating implausible fits (0.3% of bootstrapped fits).

Additionally, we investigated to which degree 1-back adaptation was dependent on the relationship between a unit's orientation preference and the adaptor/test orientation. For instance, one may expect strongest adaptation when the repeated stimuli match the unit's preferred orientation, due to the strong response during the adaptation period. To shed light on this question, we first binned units into three equally sized subgroups per region, based on their orientation selectivity. Orientation selectivity was quantified as

$$OSI = \frac{fr_{preferred} - fr_{non-preferred}}{fr_{preferred} + fr_{non-preferred}}$$

where $fr_{preferred}$ and $fr_{non\text{-}preferred}$ refer to the unit's firing rates to its preferred orientation (eliciting the highest average firing rate) and the orthogonal orientation, respectively. The OSI ranges from 0 to 1, where 0 indicates no selectivity (identical firing rates to preferred and non-preferred orientations) and 1 indicates maximal selectivity (zero firing rate to non-preferred orientation). Subsequently, for each subgroup of units we computed adaptation ratios as a function of the previous (adaptor) and current (test) stimulus orientation relative to the unit's preferred orientation (see **Supplementary Figure S2**). To statistically test the influence of the relative adaptor/test stimulus orientation on adaptation ratios, and to test whether this influence depended on the degree of orientation selectivity of the units, we conducted a 3 x 3 mixed ANOVA, with repeated measures factor "relative adaptor/test orientation" (0, 45 and 90º) and between-unit factor "orientation selectivity" (low, medium and high OSI).

Since we found that adaptation was indeed strongest when the repeated orientations matched the unit's preferred orientation, we repeated our analysis of the recovery time course of adaptation for these trial types. That is, we computed adaptation ratios on a subset of trials, for which the current orientation matched the unit's preferred orientation and the previous orientation either matched (*repeat*) or was orthogonal (*orthogonal*) to the preferred orientation. While this approach had the advantage of quantifying adaptation to the most effective adaptor stimulus, it had the disadvantage of limiting the analysis to a much smaller set of trials compared to computing adaptation for all orientations. We did not observe qualitative differences between the two analysis approaches (see **Supplementary Figure S3** and **Supplementary Table 2**).

Orientation-specific adaptation to static gratings

Analyses of adaptation to static gratings were similar to the analysis of the drifting grating data, with two exceptions. First, we quantified adaptation based on neural responses during the entire stimulus presentation period (0-250 ms). As discussed above, we chose to use a longer time window for analyzing static grating adaptation, since due to the back-to-back presentation of the static gratings, visual responses to the previous stimulus overlapped with the initial time window of the current stimulus, thereby increasing response variability in this early time window. However, largely similar results were obtained when performing the analyses on the same time window used in the drifting grating experiment (0 to 100 ms; **Supplementary Figure S6**). Second, we only analyzed adaptation to all orientations, regardless of the units' orientation preferences. This analysis was similar to the main analysis of drifting grating adaptation described above. Due to the rapid presentation of the static gratings, without intervening grey periods, responses persisted into the presentation period of the next grating. Since sub-selecting data according the units' preferred orientation led to response differences in the adaptation period (i.e. larger response to preferred than orthogonal adaptor), the bleeding of the previous response into the current stimulus time window strongly biased the response to the current grating, thereby confounding genuine adaptation-induced changes in the response to the current grating. Conversely, when analyzing adaptation for all orientations, regardless of the units' orientation tuning, the relationship between the adaptor orientations and the units' preferred orientations were balanced across *repeat* and *non-repeat* trials, and therefore did not bias the analysis of the current response.

References

Barlow, H. B., & Földiák, P. (1989). Adaptation and decorrelation in the cortex. In *In R. Durbin, C. Miall, & G. Mitchison (Ed.) The Computing Neuron* (pp. 54–72). Addison-Wesley.

Chopin, A., & Mamassian, P. (2012). Predictive Properties of Visual Adaptation. *Current Biology*, *22*(7), 622–626. https://doi.org/10.1016/j.cub.2012.02.021

Clifford, C. W. G., Wenderoth, P., & Spehar, B. (2000). A functional angle on some after-effects in cortical vision. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *267*(1454), 1705–1710. https://doi.org/10.1098/rspb.2000.1198

Dragoi, V., Rivadulla, C., & Sur, M. (2001). *Foci of orientation plasticity in visual cortex*. *411*, 7.

Dragoi, V., Sharma, J., & Sur, M. (2000). Adaptation-Induced Plasticity of Orientation Tuning in Adult Visual Cortex. *Neuron*, *28*(1), 287–298. https://doi.org/10.1016/S0896-6273(00)00103-3

Fritsche, M., Lawrence, S. J. D., & de Lange, F. P. (2020). Temporal tuning of repetition suppression across the visual cortex. *Journal of Neurophysiology*, *123*(1), 224–233. https://doi.org/10.1152/jn.00582.2019

Fritsche, M., Spaak, E., & de Lange, F. P. (2020). *A Bayesian and efficient observer model explains concurrent attractive and repulsive history biases in visual perception* [Preprint]. Neuroscience. https://doi.org/10.1101/2020.01.22.915553

Gekas, N., McDermott, K. C., & Mamassian, P. (2019). Disambiguating serial effects of multiple timescales. *Journal of Vision*, *19*(6), 24. https://doi.org/10.1167/19.6.24

Gibson, J. J., & Radner, M. (1937). Adaptation, after-effect and contrast in the perception of tilted lines. I. Quantitative studies. *Journal of Experimental Psychology*, *20*(5), 453–467. https://doi.org/10.1037/h0059826

Glickfeld, L. L., & Olsen, S. R. (2017). Higher-Order Areas of the Mouse Visual Cortex. *Annual Review of Vision Science*, *3*(1), 251–273. https://doi.org/10.1146/annurev-vision-102016-061331

Hachen, I., Reinartz, S., Brasselet, R., Stroligo, A., & Diamond, M. E. (2020). Dynamics of history-dependent perceptual judgment. *BioRxiv*. https://doi.org/10.1101/2020.07.12.199489

Hasson, U., Yang, E., Vallines, I., Heeger, D. J., & Rubin, N. (2008). A Hierarchy of Temporal Receptive Windows in Human Cortex. *Journal of Neuroscience*, *28*(10), 2539–2550. https://doi.org/10.1523/JNEUROSCI.5487-07.2008

Hill, D. N., Mehta, S. B., & Kleinfeld, D. (2011). Quality Metrics to Accompany Spike Sorting of Extracellular Signals. *Journal of Neuroscience*, *31*(24), 8699–8705. https://doi.org/10.1523/JNEUROSCI.0971-11.2011

Honey, C. J., Thesen, T., Donner, T. H., Silbert, L. J., Carlson, C. E., Devinsky, O., Doyle, W. K., Rubin, N., Heeger, D. J., & Hasson, U. (2012). Slow Cortical Dynamics and the Accumulation of Information over Long Timescales. *Neuron*, *76*(2), 423–434. https://doi.org/10.1016/j.neuron.2012.08.011

Jin, D. Z., Dragoi, V., Sur, M., & Seung, H. S. (2005). Tilt Aftereffect and Adaptation-Induced Changes in Orientation Tuning in Visual Cortex. *Journal of Neurophysiology*, *94*(6), 4038–4050. https://doi.org/10.1152/jn.00571.2004

Jin, M., Beck, J. M., & Glickfeld, L. L. (2019). Neuronal Adaptation Reveals a Suboptimal Decoding of Orientation Tuned Populations in the Mouse Visual Cortex. *Journal of Neuroscience*, *39*(20), 3867–3881. https://doi.org/10.1523/JNEUROSCI.3172-18.2019

Jin, M., & Glickfeld, L. L. (2020). Magnitude, time course, and specificity of rapid adaptation across mouse visual areas. *Journal of Neurophysiology*, *124*(1), 245–258. https://doi.org/10.1152/jn.00758.2019

Jun, J. J., Steinmetz, N. A., Siegle, J. H., Denman, D. J., Bauza, M., Barbarits, B., Lee, A. K., Anastassiou, C. A., Andrei, A., Aydın, Ç., Barbic, M., Blanche, T. J., Bonin, V., Couto, J., Dutta, B., Gratiy, S. L., Gutnisky, D. A., Häusser, M., Karsh, B., … Harris, T. D. (2017). Fully integrated silicon probes for high-density recording of neural activity. *Nature*, *551*(7679), 232–236. https://doi.org/10.1038/nature24636

Kim, H., Homann, J., Tank, D. W., & Berry, M. J. (2019). A Long Timescale Stimulus History Effect in the Primary Visual Cortex. *BioRxiv*. https://doi.org/10.1101/585539

King, J. L., & Crowder, N. A. (2018). Adaptation to stimulus orientation in mouse primary visual cortex. *European Journal of Neuroscience*, *47*(4), 346–357. https://doi.org/10.1111/ejn.13830

Kohn, A., & Movshon, J. A. (2003). Neuronal Adaptation to Visual Motion in Area MT of the Macaque. *Neuron*, *39*(4), 681–691. https://doi.org/10.1016/S0896-6273(03)00438-0

Kohn, A., & Movshon, J. A. (2004). Adaptation changes the direction tuning of macaque MT neurons. *Nature Neuroscience*, *7*(7), 764–772. https://doi.org/10.1038/nn1267

Kuravi, P., & Vogels, R. (2017). Effect of adapter duration on repetition suppression in inferior temporal cortex. *Scientific Reports*, *7*(1). https://doi.org/10.1038/s41598-017-03172-3

Lerner, Y., Honey, C. J., Silbert, L. J., & Hasson, U. (2011). Topographic Mapping of a Hierarchy of Temporal Receptive Windows Using a Narrated Story. *Journal of Neuroscience*, *31*(8), 2906–2915. https://doi.org/10.1523/JNEUROSCI.3684-10.2011

Liu, B., Li, Y., Ma, W., Pan, C., Zhang, L. I., & Tao, H. W. (2011). Broad Inhibition Sharpens Orientation Selectivity by Expanding Input Dynamic Range in Mouse Simple Cells. *Neuron*, *71*(3), 542–554. https://doi.org/10.1016/j.neuron.2011.06.017

Magnussen, S., & Johnsen, T. (1986). Temporal aspects of spatial adaptation. A study of the tilt aftereffect. *Vision Research*, *26*(4), 661–672. https://doi.org/10.1016/0042-6989(86)90014-3

Maus, G. W., Chaney, W., Liberman, A., & Whitney, D. (2013). The challenge of measuring long-term positive aftereffects. *Current Biology*, *23*(10), R438–R439. https://doi.org/10.1016/j.cub.2013.03.024

Müller, J. R., Metha, A. B., Krauskopf, J., & Lennie, P. (1999). Rapid Adaptation in Visual Cortex to the Structure of Images. *Science*, *285*(5432), 1405–1408. https://doi.org/10.1126/science.285.5432.1405

Murray, J. D., Bernacchia, A., Freedman, D. J., Romo, R., Wallis, J. D., Cai, X., Padoa-Schioppa, C., Pasternak, T., Seo, H., Lee, D., & Wang, X.-J. (2014). A hierarchy of intrinsic timescales across primate cortex. *Nature Neuroscience*, *17*(12), 1661–1663. https://doi.org/10.1038/nn.3862

Niell, C. M., & Stryker, M. P. (2008). Highly Selective Receptive Fields in Mouse Visual Cortex. *Journal of Neuroscience*, *28*(30), 7520–7536. https://doi.org/10.1523/JNEUROSCI.0623-08.2008

Patterson, C. A., Wissig, S. C., & Kohn, A. (2013). Distinct Effects of Brief and Prolonged Adaptation on Orientation Tuning in Primary Visual Cortex. *Journal of Neuroscience*, *33*(2), 532–543. https://doi.org/10.1523/JNEUROSCI.3345-12.2013

Peter, A., Stauch, B. J., Shapcott, K., Kouroupaki, K., Schmiedt, J. T., Klein, L., Klon-Lipok, J., Dowdall, J. R., Schölvinck, M. L., Vinck, M., Singer, W., Schmid, M. C., & Fries, P. (2020). Stimulus-specific plasticity of macaque V1 spike rates and gamma. *BioRxiv*, 23.

Schwartz, O., Hsu, A., & Dayan, P. (2007). Space and time in visual context. *Nature Reviews Neuroscience*, *8*(7), 522–535. https://doi.org/10.1038/nrn2155

Siegle, J. H., Jia, X., Durand, S., Gale, S., Bennett, C., Graddis, N., Heller, G., Ramirez, T. K., Choi, H., Luviano, J. A., Groblewski, P. A., Ahmed, R., Arkhipov, A., Bernard, A., Billeh, Y. N., Brown, D., Buice, M. A., Cain, N., Caldejon, S., … Koch, C. (2021). Survey of spiking in the mouse visual system reveals functional hierarchy. *Nature*. https://doi.org/10.1038/s41586-020-03171-x

Suárez-Pinilla, M., Seth, A. K., & Roseboom, W. (2018). Serial dependence in the perception of visual variance. *Journal of Vision*, *18*(7), 4. https://doi.org/10.1167/18.7.4

Tan, A. Y. Y., Brown, B. D., Scholl, B., Mohanty, D., & Priebe, N. J. (2011). Orientation Selectivity of Synaptic Input to Neurons in Mouse and Cat Primary Visual Cortex. *Journal of Neuroscience*, *31*(34), 12339–12350. https://doi.org/10.1523/JNEUROSCI.2039-11.2011

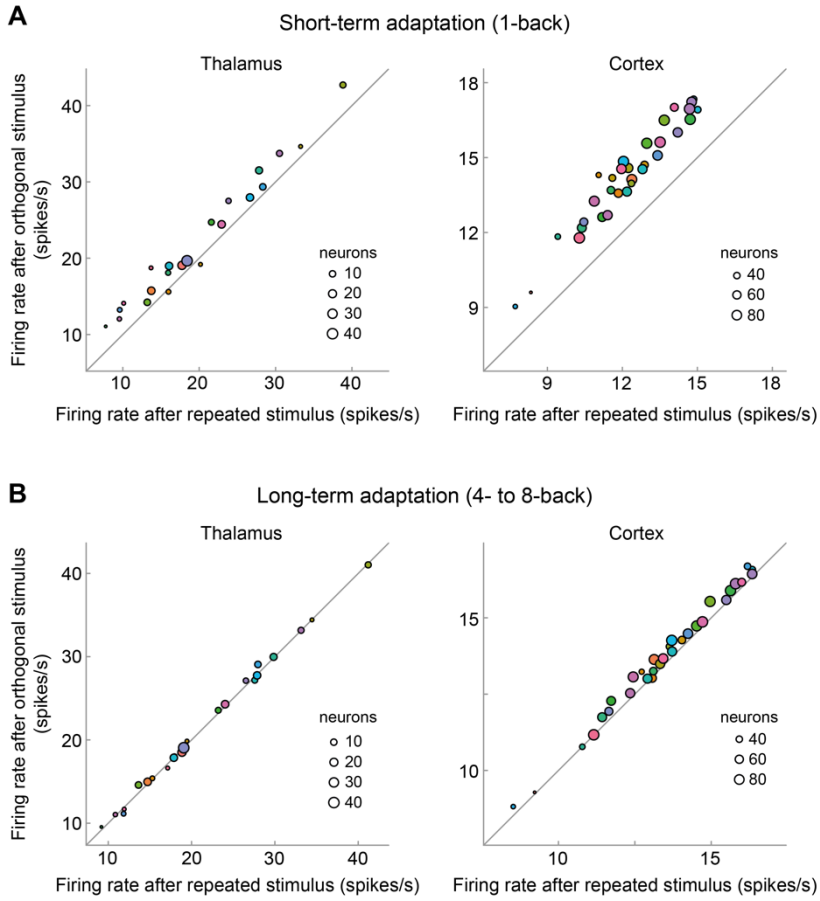
Ulanovsky, N. (2004). Multiple Time Scales of Adaptation in Auditory Cortex Neurons. *Journal of Neuroscience*, *24*(46), 10440–10453. https://doi.org/10.1523/JNEUROSCI.1905-04.2004

Wainwright, M. J. (1999). Visual adaptation as optimal information transmission. *Vision Research*, *39*(23), 3960–3974. https://doi.org/10.1016/S0042-6989(99)00101-7

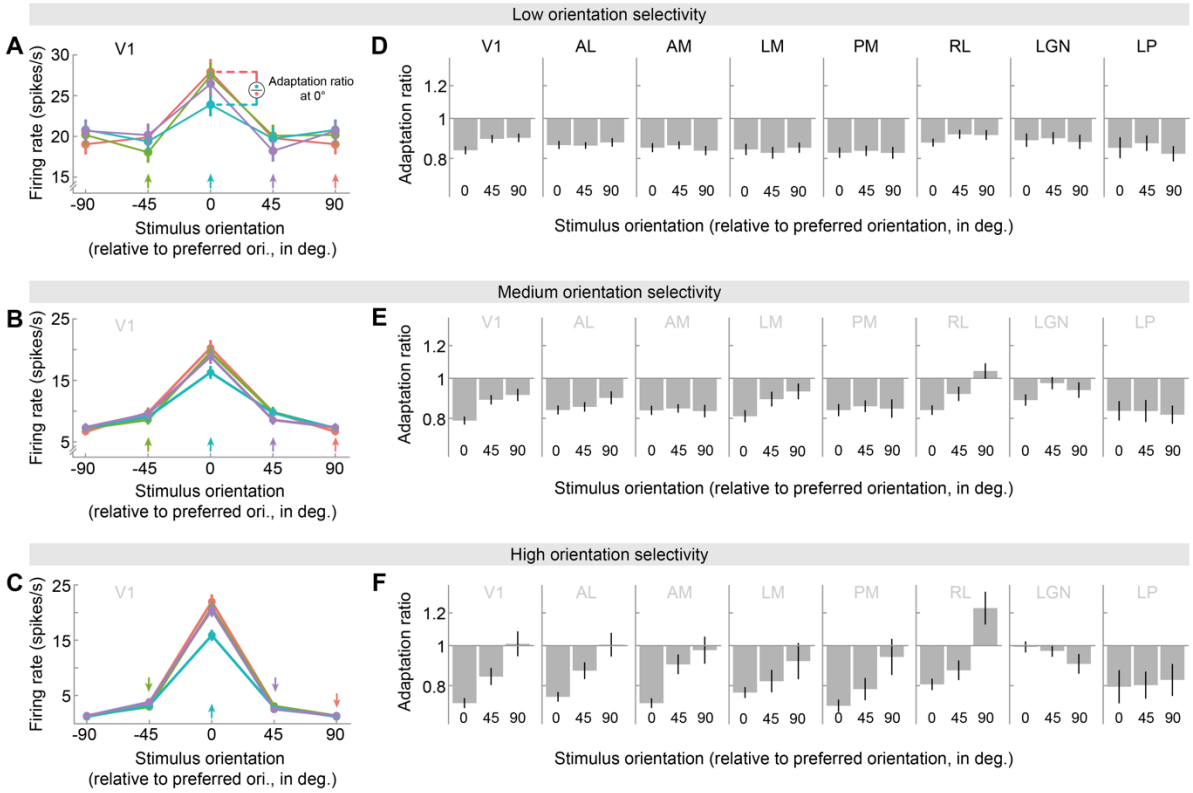
Weber, A. I., Krishnamurthy, K., & Fairhall, A. L. (2019). Coding Principles in Adaptation. *Annual Review of Vision Science*, *5*(1), 427–449. https://doi.org/10.1146/annurev-vision-091718-014818

Zhou, J., Benson, N. C., Kay, K. N., & Winawer, J. (2018). Compressive Temporal Summation in Human Visual Cortex. *Journal of Neuroscience*, *38*(3), 691–709. https://doi.org/10.1523/JNEUROSCI.1724-17.2017

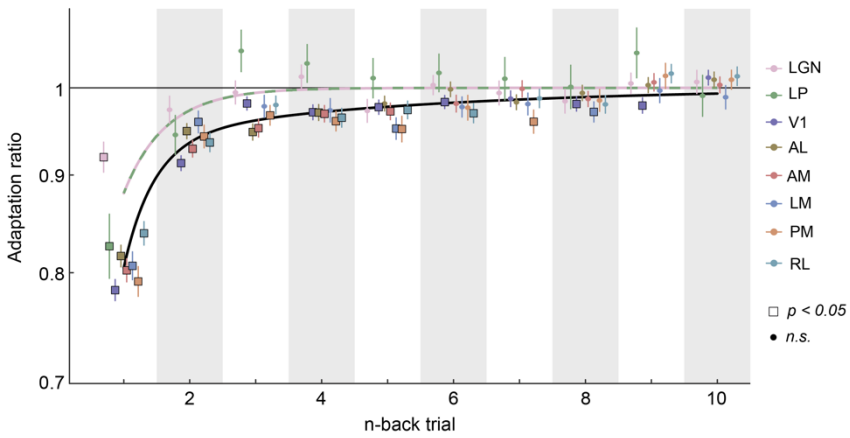
Supplemental Material



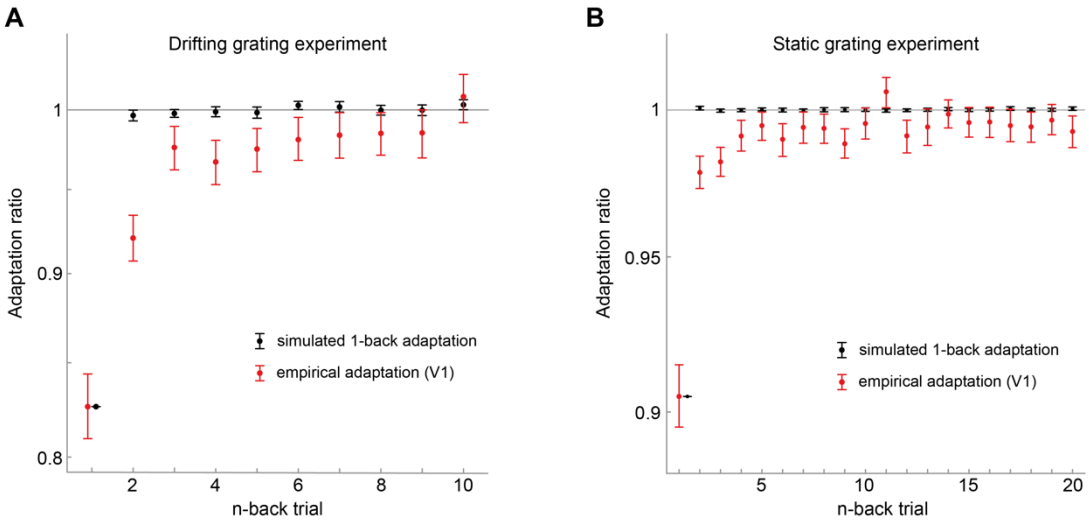
Supplementary Figure S1. Cortical short- and long-term adaptation to drifting gratings is consistent across mice. (A) The average firing rate to a stimulus preceded by a stimulus with the same orientation (*x*-axis) is consistently smaller than the firing rate to a stimulus preceded by a stimulus with the orthogonal orientation (*y*-axis) across mice (colored dots denote different mice; size scaled by the number of neurons of each mouse) in both thalamus (left) and cortex (right), as indicated by datapoints positioned above the diagonal. **(B)** Consistent long-term adaptation in cortex, but not in thalamus. Across the group of mice, average firing rates of cortical neurons remain consistently reduced when the current stimulus is preceded by a stimulus with the same versus orthogonal orientation on the 4- to 8-back trial (10 to 22 seconds ago), as indicated by datapoints positioned above the diagonal (right). No long-term adaptation effects are observed in the thalamus, where datapoints predominantly fall onto the diagonal (left). For all panels adaptation was computed in a time window of 0 to 100 ms relative to the onset of the current stimulus and averaged across thalamic and cortical neurons in each mouse.



Supplementary Figure S2. Adaptation depends on orientation tuning and adaptor/test orientation. (A, B, C) Orientation tuning curves in V1 for units of low **(A)**, medium **(B)** or high **(C)** orientation selectivity (tertile split, see **Materials & Methods**), following adaptation to different 1-back grating orientations (colored arrows). Stimulus and adaptor orientations are expressed relative to each neuron's preferred orientation. Tuning curves show local response reductions to the adapted orientation. **(D, E, F)** Adaptation ratios as a function of the adaptor and test orientation relative to the neuron's preferred orientation. For instance, the adaptation ratio for a relative stimulus orientation of 0° compares the visual response to a test grating with the neuron's preferred orientation, when it is preceded by an adaptor grating with the same (preferred) orientation, versus when it is preceded by the orthogonal (non-preferred) adaptor orientation (see also illustration in panel **A**). In V1 (panels **D**, **E** and **F**, leftmost columns), adaptation was strongest when adaptor and test stimuli corresponded to the preferred orientation of the neuron, and decreased when adapting and testing with less preferred orientations (significant main effect of relative orientation, $p = 4e\text{-}11$). This relationship was particularly strong in neurons exhibiting high orientation selectivity (significant interaction between relative adaptor/test orientation and orientation selectivity, $p = 0.005$; for definition of orientation selectivity see **Materials & Methods**). Nevertheless, there was clear adaptation for all adaptor orientations as indicated by 1-back adaptation ratios consistently smaller than 1 (all $p < 0.004$, corrected for multiple comparisons), except for non-preferred (90°) adaptor and test stimuli of highly selective units (panel **F**, leftmost column, 90°, $p = 0.88$). This overall pattern of adaptation effects was qualitatively similar across cortical visual areas (panels **D**, **E** and **F**, columns 2 to 5). In thalamic areas (panels **D**, **E** and **F**, two rightmost columns), there was no evidence for a dependence of adaptation on orientation preference (no significant main effects of relative adaptor/test orientation: LGN, $p = 0.28$; LP, $p = 0.91$; no significant interactions between relative adaptor/test orientation and orientation selectivity: LGN, $p = 0.24$; LP, $p = 0.92$), likely due to the overall lower degree of orientation selectivity of thalamic neurons.

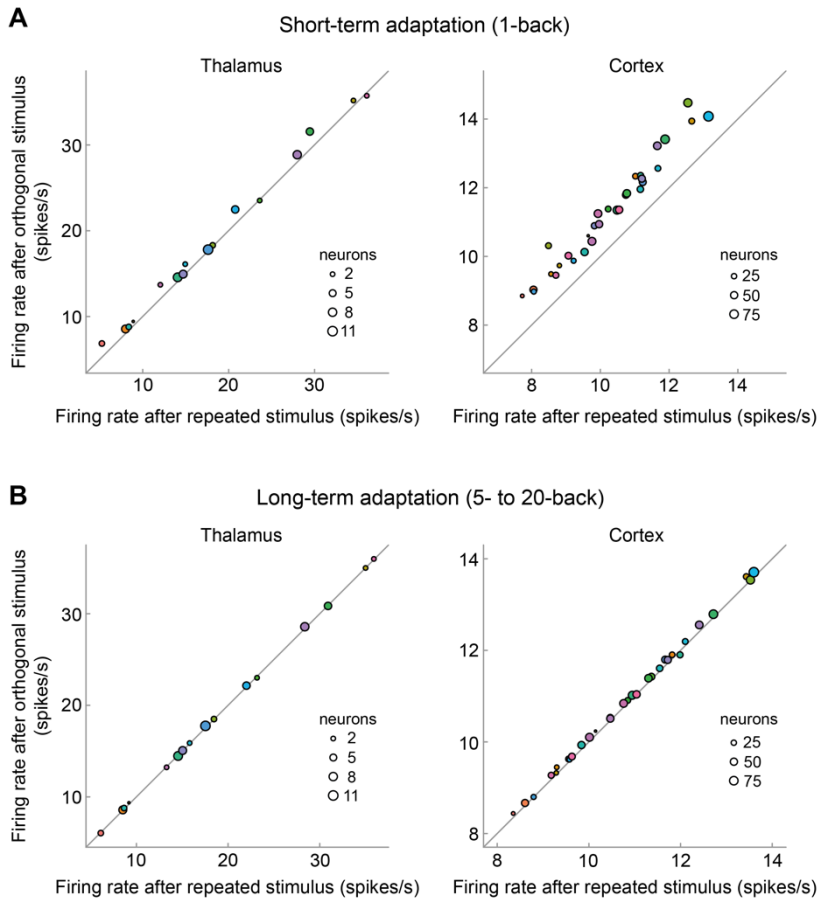


Supplementary Figure S3. 10-back adaptation ratios when adapting and testing with the neurons' preferred orientation (drifting gratings). Adaptation ratio as function of n-back trial for different visual areas (color-coded). The pattern of adaptation is similar to the main results computed for all stimulus orientations regardless of the neurons' tuning preferences (see **Figure 1E**). In particular, adaptation decays similarly and slowly across cortical visual areas and is significant for up to 9 trials back (square-shaped datapoints, $p < 0.05$, FDR-corrected per area). Conversely, it decays rapidly in thalamic areas LGN and LP (no significant adaptation beyond the 1-back trial). Error bars denote standard errors of the mean. Black and lilac-green lines show best-fitting exponential decay models to cortical and thalamic n-back adaptation.

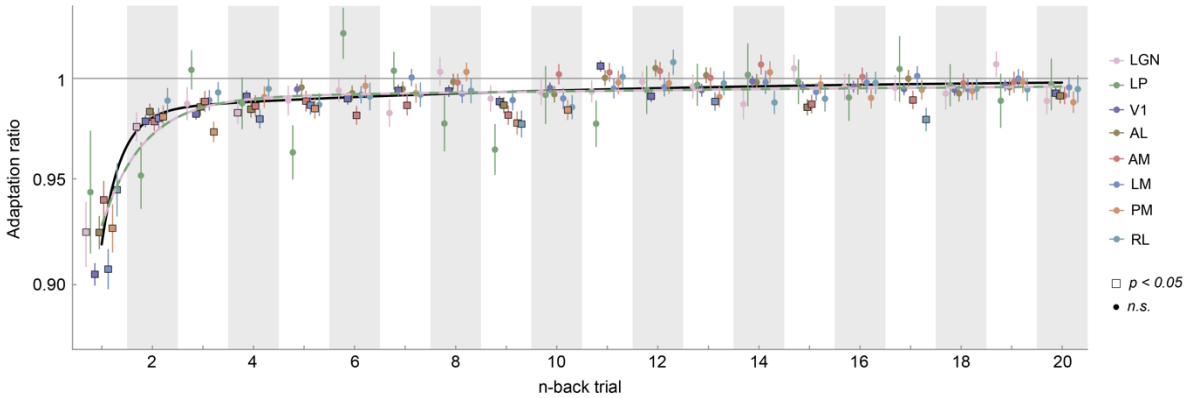


Supplementary Figure S4. Short-term (1-back) adaptation does not introduce spurious long-term adaptation effects for the particular stimulus sequences used in the experiments. Our analysis approach of quantifying adaptation to the n-back stimulus by conditioning the current visual response on the orientation difference between current and previous n-back stimulus (repeat/orthogonal) relies on the assumption that the stimulus sequence is uncorrelated. If the presentations stimulus orientations were correlated across trials, these correlations may introduce spurious adaptation effects, potentially causing short-term adaptation to masquerade as long-term adaptation (Maus et al., 2013). While the presentation order of stimuli of the current experiments was randomized, making such spurious adaptation effects unlikely, we nevertheless assessed this potential confound via the simulation of an artificial neuron that only exhibited short-term (1-back) adaptation. We simulated responses of this neuron to the particular stimulus sequences used in the drifting grating experiment (**panel A**) and static grating experiment (**panel B**). The artificial neuron responded equally to all stimulus orientations, but selectively reduced its responses to a successive

repeated orientation to mimic orientation-specific 1-back adaptation. We chose the strength of this 1-back adaptation effect to match the empirically observed 1-back adaptation of V1. We subsequently analyzed the simulated responses with the same procedure used for the empirical data. The analysis of the simulated responses recovered the ground truth 1-back adaptation effect (black data points). There were no spurious adaptation effects for stimuli further in the past, as indicated by the black data points being centered on an adaptation ratio of 1, markedly different from the empirically observed long-term adaptation effects (red data points - adaptation in V1). Black error bars denote 95% CIs of adaptation across the simulations of the 32 stimulus sequences. Red error bars denote 95% CIs of empirical adaptation across neurons in V1.



Supplementary Figure S5. Cortical short- and long-term adaptation to static gratings is consistent across mice. Same conventions as in **Supplementary Figure S1**, here for adaptation to 250-milliseconds static gratings **(A)** The average firing rate to a stimulus preceded by a stimulus with the same orientation (*x*-axis) is consistently smaller than the firing rate to a stimulus preceded by a stimulus with the orthogonal orientation (*y*-axis) across mice (colored dots denote different mice; size scaled by the number of neurons of each mouse) in both thalamus (left) and cortex (right), as indicated by datapoints positioned above the diagonal. **(B)** Consistent long-term adaptation in cortex, but not in thalamus. Across the group of mice, average firing rates of cortical neurons remain consistently reduced when the current stimulus is preceded by a stimulus with the same versus orthogonal orientation on the 5- to 20-back trial (1 to 5 seconds ago), as indicated by datapoints positioned above the diagonal (right). No long-term adaptation effects are observed in the thalamus, where datapoints are positioned on the diagonal (left). For all panels, adaptation was computed in a time window of 0 to 250 ms relative to the onset of the current stimulus and averaged across thalamic and cortical neurons in each mouse, respectively.



Supplementary Figure S6. Adaptation ratios for static gratings as function of n-back trial, computed for 0 to 100 ms after the onset of the current stimulus. The temporal decay of adaptation is largely similar to the one computed based on visual responses during 0 to 250 ms after stimulus onset (**Figure 2E**). However, a clear difference between the decay of cortical and thalamic adaptation is not borne out in this analysis. Here, estimates of adaptation ratios are more variable compared to adaptation ratios computed over the entire stimulus presentation period, likely because, due to the back-to-back presentation of static gratings, visual responses to the previous stimulus overlapped with the initial time window of the current stimulus, thereby increasing response variability in the early time window.

ROI	a_{fast}	τ_{fast}	a_{slow}	τ_{slow}
V1	**13.99** [10.73, 16.14]	**0.85** [0.59, 1.13]	**3.45** [1.71, 6.64]	**6.82** [3.39, 13.50]
AL	**10.19** [8.14, 12.71]	**0.01** [1e-3, 0.47]	**6.84** [4.90, 8.54]	**4.08** [3.10, 5.94]
AM	**9.96** [6.88, 14.30]	**0.39** [5e-3, 0.87]	**8.03** [3.83, 11.04]	**3.39** [2.37, 6.50]
LM	**13.35** [10.37, 16.07]	**0.46** [0.02, 0.72]	**3.71** [1.89, 5.84]	**5.52** [3.42, 10.60]
PM	**13.20** [10.43, 16.32]	**0.02** [2e-3, 0.45]	**6.50** [3.91, 9.26]	**4.43** [2.75, 9.28]
RL	**4.58** [1.97, 12.00]	**0.26** [5e-3, 1.83]	**7.40** [0.7, 9.82]	**3.12** [2.19, 34.41]
LGN	**6.51** [4.04, 8.85]	**0.02** [4e-3, 0.63]	-	-
LP	**16.76** [12.68, 21.02]	**0.71** [0.03, 0.93]	-	-

Supplementary Table 1. Best fitting parameters of exponential decay models fitted to adaptation ratios (drifting gratings). Amplitude parameters ***a*** are expressed in %-response reduction of *repeat* with respect to *orthogonal* trials. Exponential time constants **τ** are expressed in units of trials. The decay of adaptation in thalamic areas LGN and LP was significantly better fit by single-exponential decay models. Therefore, no parameters for the second exponential component are provided for these areas. Values in parentheses indicate bootstrapped 95% confidence intervals. See **Supplementary Table 2** for parameter estimates of adaptation to the neurons' preferred stimulus orientations only.

ROI	a_{fast}	τ_{fast}	τ_{slow}	t_{slow}
V1	**19.46** [16.02, 21.63]	**0.83** [0.61, 1.03]	**2.29** [1.02 5.27]	**10.81** [3.89, 30.73]
AL	**9.76** [6.45, 12.79]	**0.04** [2e-3, 0.53]	**8.63** [5.96, 11.53]	**2.58** [1.55, 3.75]
AM	**10.84** [5.60, 18.10]	**0.37** [3e-3, 1.05]	**8.95** [1.78, 13.86]	**3.01** [1.82, 15.30]
LM	**14.63** [11.56, 18.07]	**0.03** [3e-3, 0.53]	**4.72** [2.46, 6.91]	**6.93** [4.34, 16.90]
PM	**13.62** [10.27, 18.10]	**0.04** [3e-3, 0.57]	**7.23** [3.52, 10.06]	**4.36** [2.78, 10.47]
RL	**10.51** [5.63, 16.34]	**0.54** [0.01, 1.28]	**5.63** [8e-3, 9.86]	**3.95** [2.21, 43.42]
LGN	**8.07** [4.38, 11.05]	**0.83** [0.03, 3.20]	-	-
LP	**17.58** [11.17, 23.77]	**0.61** [0.03, 0.89]	-	-

Supplementary Table 2. Best fitting parameters of exponential decay models fitted to adaptation ratios computed on preferred test stimuli (drifting gratings). Amplitude parameters ***a*** are expressed in %-response reduction of *repeat* with respect to *orthogonal* trials. Exponential time

constants τ are expressed in units of trials. The decay of adaptation in thalamic areas LGN and LP was significantly better fit by single-exponential decay models. Therefore, no parameters for the second exponential component are provided for these areas. Values in parentheses indicate bootstrapped 95% confidence intervals. The estimated decay parameters are qualitatively similar to those estimated considering all stimulus orientations, regardless of the neurons' orientation preferences (see **Supplementary Table 1**).

ROI	a_{fast}	τ_{fast}	a_{slow}	τ_{slow}
V1	**8.17** [7.11, 9.23]	**0.54** [0.37, 0.71]	**2.04** [1.33, 2.91]	**9.12** [6.09, 14.82]
AL	**7.38** [5.27, 8.64]	**0.68** [0.02, 1.05]	**1.49** [0.77, 3.25]	**14.79** [5.24, 39.75]
AM	**6.24** [4.46, 8.05]	**0.48** [8e-3, 0.96]	**2.77** [1.13, 4.43]	**6.54** [3.82, 19.11]
LM	**7.93** [6.14, 9.80]	**0.40** [0.02, 0.61]	**2.91** [1.89, 4.00]	**7.74** [5.42, 13.36]
PM	**8.46** [6.74, 10.30]	**0.64** [0.36, 0.87]	**1.57** [1.10, 2.25]	**21.78** [13.07, 37.00]
RL	**5.39** [3.50, 7.56]	**0.66** [0.02, 1.09]	**1.39** [0.87, 2.07]	**19.58** [10.99, 37.51]
LGN	**3.91** [1.80, 6.59]	**2.93** [1.36, 7.32]	**-**	**-**

Supplementary Table 3. Best fitting parameters of exponential decay models fitted to adaptation ratios (static gratings). Amplitude parameters ***a*** are expressed in %-response reduction of *repeat* with respect to *orthogonal* trials. Exponential time constants τ are expressed in units of trials (250 ms duration). The decay of adaptation in LGN was significantly better fit by single-exponential decay model. Therefore, no parameters for the second exponential component are provided for LGN. Values in parentheses indicate bootstrapped 95% confidence intervals.

ROI	a_{fast}			τ_{fast}			a_{slow}			τ_{slow}		
V1	13.83	13.99	14.89	0.83	0.85	1.02	3.57	3.45	2.52	5.15	6.82	8.50
AL	9.75	10.19	10.19	4e-3	0.01	0.37	6.58	6.84	7.37	4.00	4.08	3.52
AM	9.75	9.96	9.44	0.42	0.39	0.21	7.12	8.03	8.80	3.58	3.39	3.49
LM	12.78	13.35	13.69	0.29	0.46	0.54	4.10	3.71	4.11	4.48	5.52	5.18
PM	12.25	13.20	14.06	8e-3	0.02	0.03	7.27	6.50	5.86	3.61	4.43	4.30
RL	5.75	4.58	10.18	0.01	0.26	1.44	5.99	7.40	1.35	3.40	3.12	16.80
LGN	5.99	6.51	6.62	0.04	0.02	0.02	-	-	-	-	-	-
LP	15.82	16.76	15.10	0.77	0.71	0.75	-	-	-	-	-	-

Supplementary Table 4. Similar exponential decay parameters across analyses with different unit selection criteria (drifting gratings). Best fitting parameters of exponential decay models fitted to adaptation ratios. The analysis presented in the main manuscript was based on visually responsive units whose average firing rate during to the most effective orientation was larger than 5 Hz (middle white columns). Here we re-ran the analysis using a stricter selection criterion of 10 Hz (red columns) and a less strict criterion of 2.5 Hz (blue column). The estimated exponential decay parameters are qualitatively similar for different unit selection criteria. Amplitude parameters ***a*** are expressed in %-response reduction of *repeat* with respect to *orthogonal* trials. Exponential time constants τ are expressed in units of trials. In all cases, the decay of adaptation in thalamic areas LGN and LP was significantly better fit by single-exponential decay models. Therefore, no parameters for the second exponential component are provided for these regions. Here, exponential decay models were fitted to adaptation ratios computed for all stimulus orientations, regardless of the neurons' orientation preferences.

ROI	a_{fast}			t_{fast}			a_{slow}			t_{slow}		
V1	8.03	8.17	8.94	0.64	0.54	0.52	1.56	2.04	2.01	11.81	9.12	8.92
AL	8.02	7.38	6.32	0.94	0.68	0.33	1.13	1.49	2.81	22.79	14.79	6.51
AM	4.72	6.24	6.84	0.38	0.48	0.51	3.23	2.77	2.81	5.01	6.54	6.70
LM	7.49	7.93	8.49	0.58	0.40	0.44	2.33	2.91	2.37	10.00	7.74	10.28
PM	8.32	8.46	8.49	0.67	0.64	0.66	1.25	1.57	1.50	28.35	21.78	22.61
RL	4.12	5.39	5.38	0.72	0.66	0.71	1.09	1.39	1.47	29.23	19.58	16.73
LGN	2.80	3.91	4.29	3.75	2.93	2.60	-	-	-	-	-	-

Supplementary Table 5. Similar exponential decay parameters across analyses with different unit selection criteria (static gratings). Same conventions as in **Supplementary Table 4**, but for analysis of adaptation to briefly presented static gratings. Different unit selection criteria yield qualitatively similar parameter estimates, demonstrating the robustness of the results against the choice of selection criteria. The decay of adaptation in LGN was significantly better fit by a single-exponential decay model. Therefore, no parameters for the second exponential component are provided for this region.