

Age-related changes in behavioural and neural variability in a decision-making task

Received: 15 September 2025

Accepted: 28 May 2026

Published online: 30 June 2026

Fenyng Zang¹✉, Anup Khanal^{2,3}, Sonja Förster¹, International Brain Laboratory*, Anne K. Churchland^{2,3} & Anne E. Urai^{1,3}✉

Age-related cognitive decline in learning and decision-making may arise from increased variability of neural responses. Here, we investigated how ageing affects behavioural and neural variability by recording >18,000 neurons across 16 brain regions (including cortex, hippocampus, thalamus, midbrain, and basal ganglia) in younger and older mice performing a visual decision-making task. Older mice showed more variable response times, reproducing a common finding in human ageing studies. Ageing globally increased firing rates and post-stimulus neural variability (quantified using the Fano factor), and decreased variability quenching—the reduction in neural variability upon stimulus presentation. Older animals showed higher overall firing rates across areas of visual and motor cortex, striatum, midbrain, and hippocampus, but lower firing rates in thalamic areas. Age-related attenuation in stimulus-induced variability quenching was most prominent in visual and motor cortex, striatum, and thalamic area. These findings show how large-scale neural recordings can help uncover regional specificity of ageing effects in single neurons, ultimately improving our understanding of the neural basis of age-related cognitive decline.

Age-related cognitive impairments have long been thought to arise from higher levels of neural noise or variability. In the second half of the 20th century, theoretical accounts proposed that increased age-related neural variability reduces the effective signal-to-noise ratio within the central nervous system, leading to cognitive impairments^{1–4}. Behavioural studies at the time supported this idea, showing that the absolute threshold for detecting stimuli increases with age⁴ and that adding random stimulus noise (simulating internal neural noise) could mimic age-related differences in task performance, particularly in visual tasks¹. More recent theoretical accounts propose that declining neuromodulation in healthy ageing may impair neural networks' gain control, the efficient modulation of a neuron's input-output function⁵. Such changes reduce the precision of neural representations, implying that neural dedifferentiation lies at the basis of cognitive changes in healthy ageing⁶.

Human studies of healthy ageing, typically comparing young adults (20–30 years) with older adults (often up to >85 years), present a complex and sometimes contradictory picture of age-related changes in neural variability. While some studies found widespread age-related increases in blood oxygen level-dependent (BOLD) variability across both cortical and subcortical regions⁷, others reported that many brain regions become less variable in older people^{8–10}. The behavioural implications of age-related changes in BOLD variability remain debated. Some studies find lower brain variability in older people, linked to slower and more variable response times (RTs)⁹, while others report increased BOLD variability in older adults, negatively affecting decision-making¹¹. Older adults' EEG activity consistently shows an increase in neural variability, as measured by metrics such as P300 latency variability¹², weighted permutation entropy¹³, and particularly the slope of 1/f power spectra^{14–16}. The latter measure, characterized by

¹Cognitive Psychology Unit, Institute of Psychology, Leiden University, Leiden, The Netherlands. ²Department of Neurobiology, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, CA, USA. ³Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, USA. *A list of authors and their affiliations appears at the end of the paper. ✉e-mail: f.zang@fsw.leidenuniv.nl; a.e.urai@fsw.leidenuniv.nl

the aperiodic component of 1/f power spectral density¹⁶, has been interpreted as reflecting an increase in baseline neural noise, and correlates with age-related impairments in cognitive functions such as visual working memory¹⁶ and visual processing¹⁵. However, flatter slopes of EEG power spectra were recently shown to arise from a cardiac, rather than neural, source that changes with age¹⁷. This highlights the importance of using direct neural recordings to study brain function without age-related vascular and cardiac confounds that may affect the interpretation of BOLD or EEG data¹⁷.

Observations from single-neuron measurements in cortical areas show a complex picture of age-related change. Some studies report increased firing rates^{18–21}, while others have found decreased rates with ageing²². Only a few studies have looked directly into age-related changes in the variability of single-neuron activity. In anaesthetized rhesus monkeys, comparing young adults (5–9 years; ~18–30 human-equivalent years) with aged animals (23–31 years; ~70–90 human-equivalent years), cortical areas V1 and MT show increased age-related trial-by-trial spiking variability (measured using Fano factor, defined as the ratio of spike-count variance to mean spike count across trials), reduced signal-to-noise ratio²¹, and increased noise correlations²³. Extending this evidence to rodents, a recent *in vivo* study in mouse V1 reported higher single-neuron Fano factor and increased noise correlations in 12-month-old mice compared with 8-week-old mice in response to repeated grating stimuli²⁴. Although informative, these findings are limited by relatively small datasets: for example, Yang et al. recorded 172 and 173 neurons from 3 young and 4 old monkeys, respectively²¹, while the recent mouse V1 study recorded 136 neurons from 6 young mice and 140 neurons from 5 old mice²⁴. Moreover, their generalization to task-engaged animals and other brain regions remains unclear. In addition, these studies did not examine the time course of neural variability. Neural variability, as measured by the Fano factor, typically decreases following stimulus onset—a phenomenon known as variability quenching—which has been consistently observed across species and cortical areas^{25–29}. Yet, how this reduction of variability in response to sensory stimuli changes with ageing remains unknown, particularly in behaving animals.

More broadly, rodent studies have also reported age-related changes in neural activity across multiple levels of organization, from cellular and synaptic changes in excitability, synaptic transmission, and excitation–inhibition balance to circuit- and population-level alterations in firing rates and population activity patterns across regions^{30–37}. Related changes have also been reported in rodent models of neurodegenerative disease, including Alzheimer’s disease, which are often characterized by circuit hyperexcitability^{38–40}. These studies span diverse experimental contexts, but often focus on one or a small number of regions and include many experiments conducted in passive paradigms or in anaesthetized/*ex vivo* preparations. As a result, they provide limited insight on how ageing impacts trial-to-trial neural variability during perceptual decision-making across brain regions.

In recent years, large-scale neural recording techniques have offered exciting opportunities to record from many neurons as animals perform complex decision-making tasks⁴¹. We here build on the work of the International Brain Laboratory (IBL), which has standardized training and recording pipelines for mouse decision-making behaviour^{42–44}. Recording brain-wide neural activity during a visual decision-making task revealed that decision-related signals are distributed across much of the mouse brain⁴⁵. Visual stimuli evoked transient responses in early visual areas, which were followed by choice-related ramping in midbrain and hindbrain regions. Neural signals related to movement, reward, and feedback were widespread across the brain⁴⁵. Here, we build on this well-curated dataset and contribute additional recordings from older mice, allowing us to investigate age-related changes in behaviour and neural activity. We openly share all data and code to promote reproducibility and further analyses of this rich dataset.

In this study, we use large-scale neural recordings from behaving mice to investigate age-related changes in single-neuron variability across the brain. We analysed extracellular Neuropixels recordings from 149 mice (aged 3–20 months; roughly corresponding to 23–70-year-old humans⁴⁶) across 16 brain regions (in cortex, striatum, midbrain, hippocampus, and thalamus; Table 1) while animals performed a standardized perceptual decision-making task⁴³.

We found that older animals showed higher trial-to-trial variability in RTs. Neural recordings show that ageing is accompanied by higher firing rates across the brain, increased post-stimulus neural variability (as measured with Fano factors), and attenuated stimulus-induced variability quenching. Different metrics to quantify age-related neural changes show complex regional patterns. These results suggest that behavioural differences in older animals may be accompanied by specific changes in neural variability after stimulus onset.

Results

We combined a previously released dataset of extracellular Neuropixels recordings⁴⁷ from young mice ($N = 130$, 89 male, mean age = 6.64 months, range 3.10–15.13 months)⁴⁵ with additional neural recordings in older mice ($N = 19$, 11 male, mean age = 16.58 months, range 10.58–19.90 months), all acquired using standardized protocols for behaviour and neural recordings (Fig. 1a; see “Methods”) ^{43,45}. Briefly, mice were trained to discriminate the location of a visual stimulus on the left or right side of the screen by turning a small steering wheel in front of them. The stimuli varied in contrast, modulating task difficulty. After applying a rigorous set of standardized quality control measures (see “Methods”), we included a total of 149 mice that performed 367 sessions and underwent 503 Neuropixels insertions. Animals’ age at the day of recording ranged from 3–20 months (Fig. 1b), roughly corresponding to a human age range of 23–70 years⁴⁶—approximately spanning the typical working lifespan from early adulthood to retirement. For visualization, mice were categorized based on their age at recording into a young group ($N = 97$, mean age = 5.60 months) and an old group ($N = 52$, mean age = 12.07 months), using 7.6 months (mean age in the dataset) as the age cutoff (Fig. 1b). Age was treated as a continuous variable in all statistical analyses.

Older mice show more variable behaviour on a standardized decision-making task

Older animals did not take longer to learn the standardized task, and animals of different ages eventually met the standardized criteria to be classified as trained (Supplementary Fig. 1). Although animals that took longer to learn were older by the time of neural recording, the main source of age variability at recording time arises from starting the training program later (Supplementary Fig. 2). Only behavioural data from the recording sessions were included in our main analyses.

During recording sessions, older mice performed slightly more trials until automated behaviour-based stopping criteria ended the session (range: 401–1279; Supplementary Fig. 3a)⁴³. This may reflect larger weight-dependent water requirements in older, heavier animals. To control for this potential confound in our statistical analyses (especially when computing neural metrics sensitive to trial counts), we restricted our analyses to the first 400 trials of each session. Additionally, only trials with RTs between 80 ms and 2 s were included, to exclude anticipatory (<80 ms) and very slow (>2 s) responses that are unlikely to reflect stimulus-driven decisions (Fig. 1g), consistent with prior IBL analyses⁴⁵. Unless otherwise specified, regression lines are shown only when the Bayes Factor indicates strong or moderate evidence in favour of H1. The number of in- or excluded trials did not show a significant correlation with age (included trial counts: Supplementary Fig. 3c; number of trials excluded by RT filtering: Supplementary Fig. 3b).

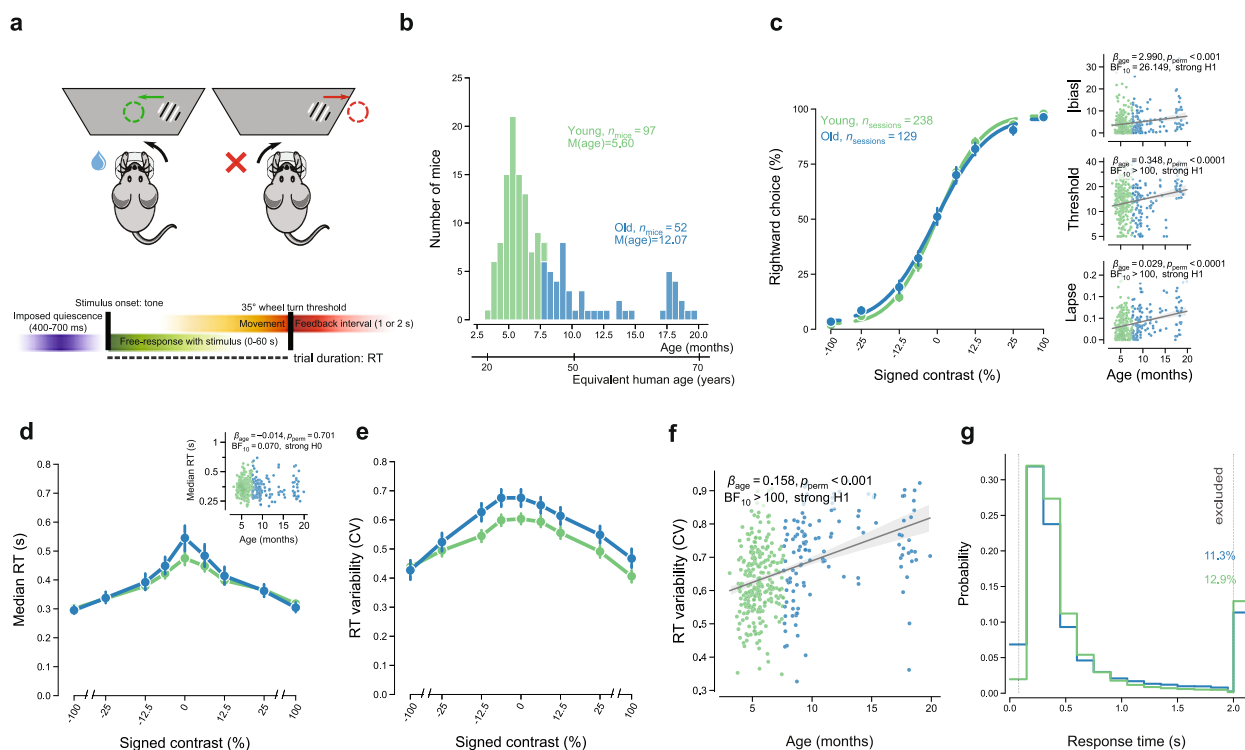


Fig. 1 | Older mice show larger RT variability and slightly worse performance in a standardized decision-making task. a Schematic of the task, showing a correct response (the visual stimulus was brought to the centre of the screen) vs. a wrong response (the visual stimulus was moved off the screen). Response times (RT) were defined as the interval from stimulus onset to response completion. Adapted from ref. 45. (© The Author(s) 2025; licensed under CC BY 4.0). **b** Distribution of mouse age on the day of neural recording ($n = 149$ mice; young $n = 97$; old $n = 52$; age cutoff: 7.6 months). Group splits were only used for visualization; all statistics used age as a continuous variable. The age range roughly corresponds to a human age range of 23–70 years⁴⁶. **c–f** Behavioural statistics were computed at the session level ($n = 367$ sessions; young $n = 238$; old $n = 129$). Age effects were estimated using general linear models with age as a continuous predictor. All statistical conclusions were based on Bayes Factors (BF_{10}); two-sided permutation p -values are unadjusted for multiple comparisons and are reported as complementary statistics (see

“Methods”). Error bars and shaded bands indicate 95% confidence intervals. Unless otherwise noted, regression lines are plotted only when the Bayes Factor indicates strong or moderate H1. **c** Left: average psychometric curve across sessions for each age group. Right: relationship between psychometric parameters and mouse age. Each dot represents one session. Absolute bias: $p_{\text{perm}} = 0.00069$; threshold: $p_{\text{perm}} = 0.00009$; lapse: $p_{\text{perm}} = 0.00009$. Note the log-scaling on the y-axis of the threshold panel. **d** Average chronometric curve across sessions for each age group. Inset: Relationship between median RT and mouse age. Median RT was calculated across all included trials, irrespective of stimulus contrast. $p_{\text{perm}} = 0.70093$. Note the log-scaling on the y-axis. **e** RT variability, quantified by coefficient of variation (CV), across stimulus conditions for each age group. **f** Relationship between RT variability and mouse age. $p_{\text{perm}} = 0.00009$. **g** RT distribution histogram. Bars at the far left and far right show the percentage of trials excluded due to RT shorter than 0.08 s or longer than 2 s (see also Supplementary Fig. 3).

Older mice performed slightly worse on the task. Psychometric curves fit to each session (Fig. 1c) showed that three psychometric function parameters—absolute bias, threshold, and mean lapse rate—showed an increase with age (Fig. 1c, right panels). We did not observe age-related changes in choice bias conditioned on blocks or previous choices (Supplementary Fig. 4), suggesting that the age-related differences in overall psychometric parameters are not driven by changes in the integration of prior information with sensory evidence. While history-dependent biases are well established in this task⁴⁸, they do not seem related to the age effects observed here.

Older mice did not respond more slowly, but were more variable in their RTs. We defined RT as the time animals took to complete their choice (distinct from the time at which they initiated their first movement). Although older mice did not exhibit slower RTs on average (Fig. 1d, right), their RTs were notably more variable (Fig. 1e, f). The variability in RTs, as measured by the coefficient of variation (CV; standard deviation/mean), increased with age. This finding aligns with many human studies that demonstrate increased trial-to-trial RT variability in older adults across a wide range of cognitive tasks^{5,49–51}. These results suggest that age-related increases in trial-to-trial RT variability, widely reported in humans, can also be observed in behaving rodents. This result was robust to various control analyses:

other measures of variability (Supplementary Fig. 5a) and for a definition of RT that captures the first movement initiation, rather than movement completion (Supplementary Fig. 5b), and differences between dataset sources or lab-specific environmental factors (Supplementary Figs. 6 and 7).

Large-scale Neuropixels recordings across the mouse lifespan

To investigate age-related changes in single-neuron activity, we analysed data from 149 mice, comprising 367 sessions and 503 insertions (up to two per recording session) of Neuropixels recordings across 16 brain regions of interest (Fig. 2a, b and Table 1). These included 6 cortical regions, along with structures from the hippocampus, thalamus, midbrain, basal ganglia, and olfactory areas (Fig. 2c). These regions were selected based on a combination of scientific, statistical, and practical considerations: scientifically, frontal cortical areas are known to show early age-related decline; statistically, the set includes the repeated site used to assess cross-laboratory reproducibility⁴⁴; and practically, these areas offered reliable surgical accessibility. We applied a series of standardized quality control metrics to the neural data before proceeding with subsequent analyses (Fig. 2d; see “Methods” for details)⁴⁵. From 242,671 recorded units (including multi-neuron activity) across these 16 regions, the quality control process

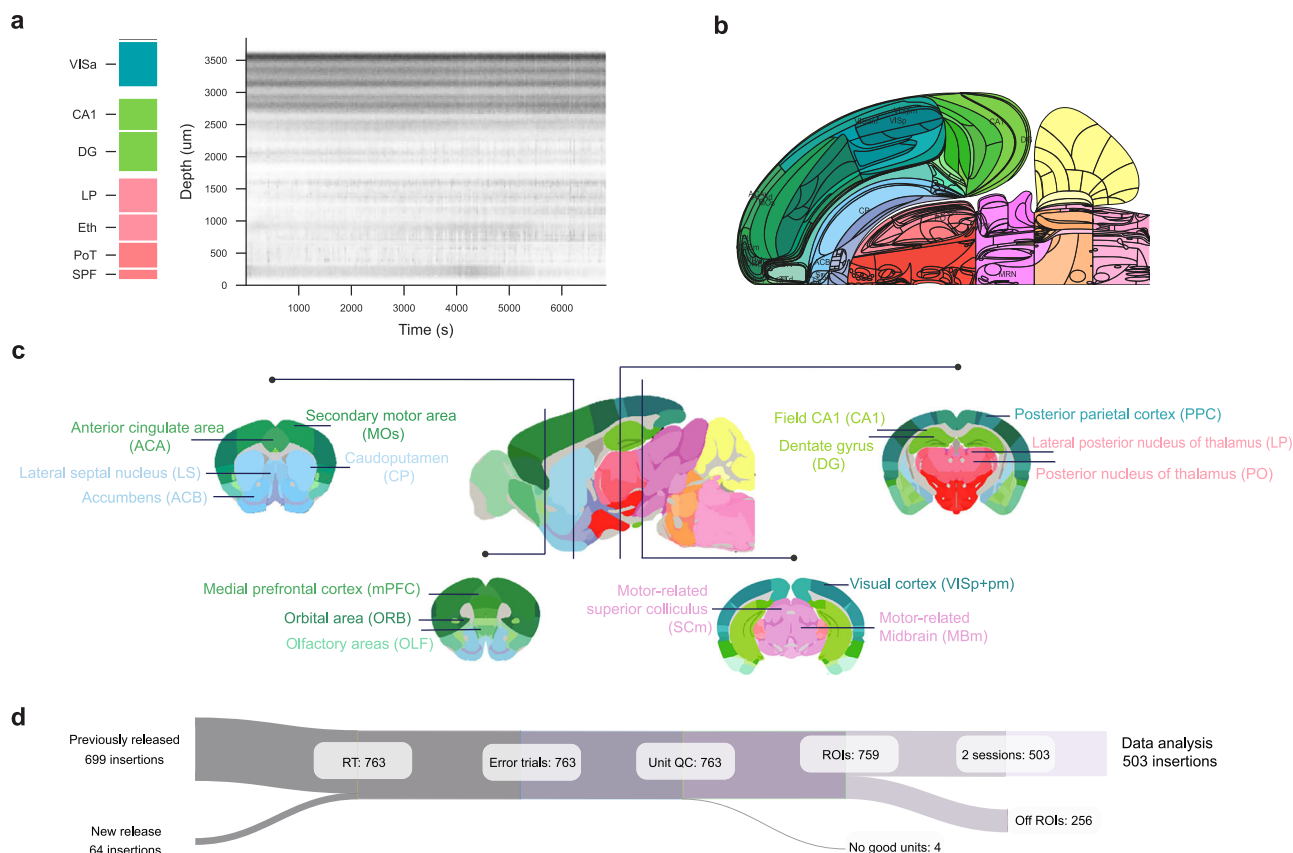


Fig. 2 | Large-scale neural recording across the mouse brain. **a** Raster plot from one example recording probe, with aligned regions indicated on the left. Brain regions: VISa (anterior visual area), CA1 (field CA1), DG (dentate gyrus), LP (lateral posterior nucleus of the thalamus), Eth (ethmoid nucleus of the thalamus), PoT (posterior triangular thalamic nucleus), SPF (subparafascicular nucleus). The y-axis shows brain regions this probe traverses, and the x-axis shows the time elapsed from session start. Each dot indicates one spike, with dark bands showing brain regions with many spiking neurons. **b** The 2D flatmap representation of the mouse brain⁴⁵. Labelled regions indicate regions of interest (ROIs), as detailed in (c) and

Table 1. **c** A 2D-sagittal mouse brain slice and four corresponding coronal slices, adapted from the Allen Mouse Brain Atlas and Allen Reference Atlas—Mouse Brain^{102,103} (mouse.brain-map.org and atlas.brain-map.org), with ROI acronyms and annotations added. **d** The number of insertions remaining after each quality control (QC) step. Previously released: the public dataset from IBL; new release: dataset recorded during the present project. RT: RT and missing events; Error trials: minimum 3 error trials; Unit QC: single unit QC; ROIs: ROIs only; 2 sessions: minimum 2 sessions per region. See Table 2 and “Methods” for more details.

identified 18,755 good neurons. Overall neural yield was slightly reduced in older mice ($\beta_{\text{age}} = -0.106$, $p_{\text{perm}} = 0.001$; $\text{BF}_{10} > 100$, strong HI), specifically in PPC, LS, and ACA (Supplementary Fig. 8).

After confirming high-quality neural data across age groups, we first assessed age-related changes in two general measures of neural functioning: overall firing rates and contrast modulation of neural responses. Throughout, we present different neural metrics first globally across recorded brain regions and then show their regional distribution.

Increased firing rate in older animals

Ageing increased firing rates globally. Previous studies have reported increased firing rates with ageing across species. In rodents, age-related increases in firing rates have been reported in multiple brain areas, including sensory cortex and hippocampus^{33,52,53}. Increased firing rates have also been reported in the sensory cortex of other mammals, including cat V1¹⁹. In non-human primates, ageing has been associated with increased firing rates in cortical areas such as macaque V1 and MT^{20,21}, as well as in specific cell types in the monkey prefrontal cortex¹⁸. In contrast, other studies have reported decreased firing rates with ageing, particularly in the monkey prefrontal cortex²². Here, when pooling neurons across all brain regions, we found global age-related changes in firing rates in both pre- and post-stimulus time windows (Fig. 3a–c, left panels). These effects were present in areas of visual and

motor cortex (VISp+pm, MOs), striatum (LS, ACB, CP), midbrain (SCm, MBm), and hippocampus (CA1, DG) (right panels of Fig. 3b, c and Supplementary Figs. 9 and 10). Notably, the two thalamic areas LP and PO showed the reverse pattern of reduced firing rates in older animals (right panels of Fig. 3b, c and Supplementary Figs. 9 and 10). In PPC, an area causally implicated in age-related effects on learning and decision making⁵⁴, we found evidence for no effect of ageing on firing rates. Frontal cortical, orbital, and cingulate areas showed weak or inconclusive ageing patterns.

Neural responses throughout the brain were modulated by stimulus contrast⁴⁵ (Fig. 3d and Supplementary Fig. 11), but we found no significant effects of ageing on this global contrast modulation (Fig. 3e, left panel). Previous studies show that aged neurons in the primary visual cortex of both cats and macaques exhibit weaker tuning to specific features like orientation and direction, and lower signal-to-noise ratios^{19,20}, and aged mice show declines in frequency selectivity within the auditory midbrain⁵⁵. This neural dedifferentiation, characterized by a loss of selectivity, means individual neurons respond to a broader range of stimuli, potentially leading to a less precise and noisier representation of sensory information⁶. We computed a contrast modulation slope for each neuron by fitting a linear regression between stimulus contrast levels [0, 6.25, 12.5, 25, 50, 100%] and the change in baseline-corrected firing rate (Δ firing rate, defined as post-stimulus minus pre-stimulus firing rate). An omnibus test revealed no

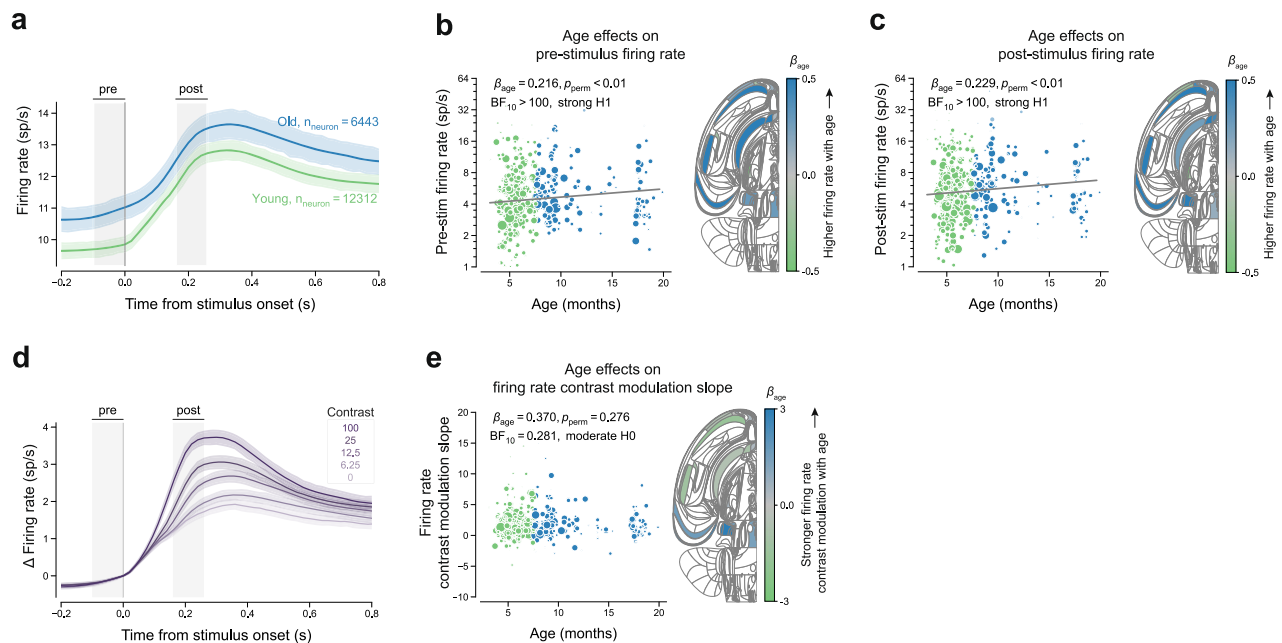


Fig. 3 | Age effects on global firing rates and neural contrast modulation. **a** Time courses of overall firing rates, aligned to stimulus onset. Thick lines represent the mean firing rate for each age group; shaded areas show 95% confidence intervals obtained by pooling neurons within each age group. The analysis included $n = 18,755$ neurons from $n = 503$ insertions and $n = 149$ mice. Sliding window width, 100 ms; step size, 20 ms. Grey areas indicate the pre-stimulus (–100 to 0 ms) and post-stimulus (160–260 ms) windows. For left panels in (**b**, **c**, and **e**), dots show insertion-level summaries for visualization, with dot size indicating the number of neurons in each insertion. Statistical analyses and fitted regression lines were based on single-neuron metrics using age as a continuous predictor. Age effects were estimated as β_{age} using general linear models with age as a continuous predictor. All statistical conclusions were based on Bayes Factors (BF_{10}); two-sided permutation p -values are unadjusted for multiple comparisons and are reported as

complementary statistics (see “Methods”). **b** Left: relationship between pre-stimulus firing rate and mouse age. Note the log-scaled y-axis. $p_{\text{perm}} = 0.00899$. Right: a flatmap representation showing region-specific age effects on pre-stimulus firing rate. The colour bar indicates the slope of age. White areas were not covered in our recordings. **c** Same as (**b**), but for post-stimulus firing rate. $p_{\text{perm}} = 0.00200$. **d** Average firing rate time courses across contrast levels. Saturation represents different contrast levels; shaded areas indicate 95% confidence intervals. **e** Same as (**b**), but for contrast modulation slope. Contrast modulation slopes were computed for each neuron by estimating the slope between different stimulus contrast levels and the change in baseline-corrected firing rates. $p_{\text{perm}} = 0.27572$. Because the omnibus analysis does not support a global age effect, this flatmap representation is shown for completeness and is not interpreted in detail in the main text.

significant age effects on contrast modulation overall (Fig. 3e, left panel). Region-wise estimates are shown on a flatmap representation of the brain⁴⁵ (Fig. 3e, right panel) for completeness, but we do not interpret these regional patterns further given the lack of a robust omnibus effect. Notably, while β_{age} estimates showed heterogeneous and bidirectional regional trends (Fig. 3e, right panel and Supplementary Fig. 12), we describe regional patterns in detail only when a corresponding global age effect is supported, in order to avoid over-interpretation of region-specific effects.

Reduced stimulus-induced variability quenching in older animals

To quantify single-neuron response variability, we next computed Fano factors. Neural responses vary substantially across trials with identical external stimuli. The Fano factor, defined as the ratio of spike-count variance to mean spike count, serves as a measure of this single-neuron spiking variability, with its conceptual roots extending back to early analyses of neural firing patterns^{56–59}. While in vivo spike trains often approximate Poisson statistics (characterized by a Fano factor of 1, indicating that the variance of spike counts equals the mean, Supplementary Fig. 13), studies using Fano factor have revealed significant regional differences in spiking variability across brain regions^{29,60–62}, as well as phenomena like quenching, where variability decreases at stimulus onset²⁵. Reporting neural variability as Fano factor also facilitates direct comparison with prior work in this line of research^{21,23,24}. To control for biases in Fano factor estimates due to condition-specific differences in mean firing

rates, we applied a mean-subtraction correction method (see “Methods”) ⁶³.

Time courses showed different effects of ageing on the Fano factor between pre- and post-stimulus periods (Fig. 4a). Ageing did not affect pre-stimulus Fano factors (Fig. 4b, left panel), suggesting that baseline neural variability remains stable with age. Following stimulus presentation, however, older animals exhibited increased Fano factors (Fig. 4c, left panel). This effect was most robustly observed in the midbrain (SCm, MBm) and thalamic area PO (Fig. 4c, right panel and Supplementary Fig. 14c). In contrast, cortical, striatal, and hippocampal areas showed only weak or inconclusive evidence of a post-stimulus ageing effect on Fano factor, showing a regionally specific pattern of age-related neural variability. These findings are consistent with previous work in anaesthetized rhesus monkeys that reported increased Fano factors in V1 and MT neurons of older animals²¹. Our findings extend these observations by demonstrating that this age-related increase in neural variability is not a property of baseline activity but follows stimulus onset. We also note a short-lived transient Fano factor increase immediately after stimulus onset (more pronounced in older mice), the origin of which remains to be determined.

Older animals showed attenuated variability quenching following stimulus onset. Neural variability, as measured by the Fano factor, typically decreases upon stimulus onset across various cortical regions and species^{25–29}, also called variability quenching. We quantified a quench index as the difference between post- and pre-stimulus Fano factors. We found quenching across the brain attenuated with age, becoming less strongly negative (Fig. 4d, left panel). Our region-

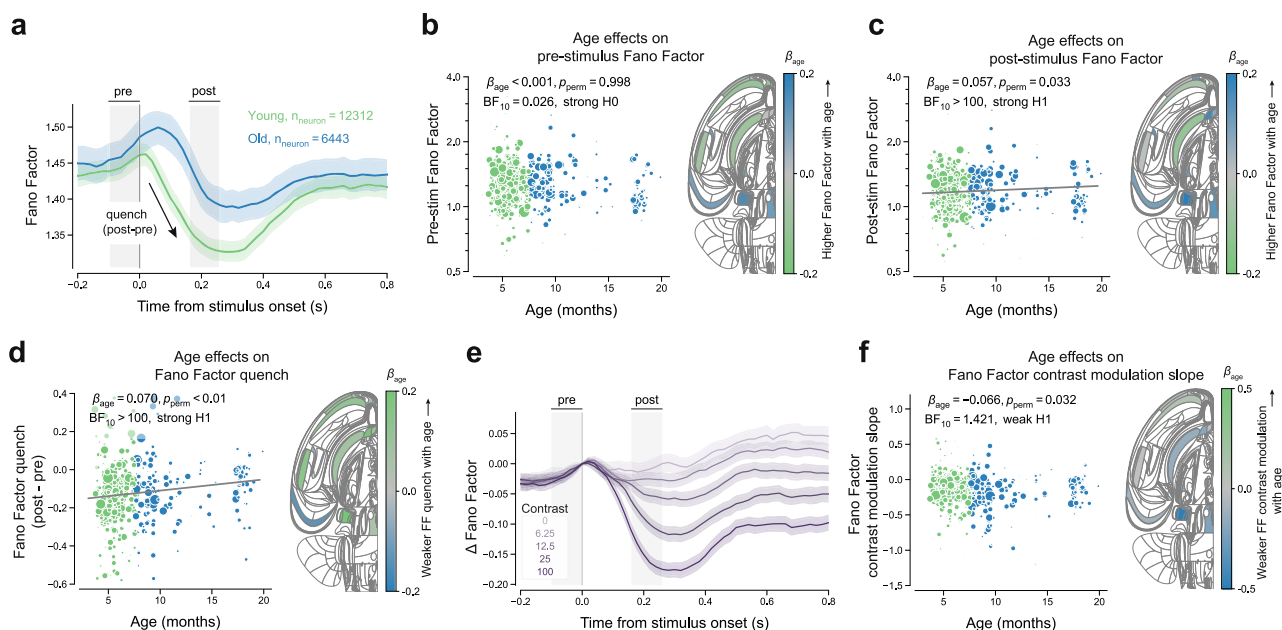


Fig. 4 | Age effects on global mean-subtracted Fano factor and contrast modulation. **a** Time courses of the overall mean-subtracted Fano factor, aligned to stimulus onset. Thick lines show the mean within each age group; shaded areas show 95% confidence intervals obtained by pooling neurons within each age group. The analysis included $n = 18,755$ neurons from $n = 503$ insertions and $n = 149$ mice. Grey areas indicate the pre-stimulus (-100 to 0 ms) and post-stimulus (160 – 260 ms) windows. For left panels in (b–d and f), dots show insertion-level summaries for visualization, with dot size indicating the number of neurons recorded in each insertion; statistical analyses and fitted regression lines were based on single-neuron metrics. All statistical conclusions were based on Bayes Factors (BF_{10}); two-sided permutation p -values are unadjusted for multiple comparisons and are reported as complementary statistics (see “Methods”). **b** Left:

relationship between pre-stimulus Fano factor and mouse age. $p_{\text{perm}} = 0.99800$. Right: a flatmap representation showing region-specific age effects on pre-stimulus Fano factor. The colour indicates the slope of age. White areas were not covered in our recordings. **c** Same as (b) but showing post-stimulus Fano factor. $p_{\text{perm}} = 0.03297$. **d** Same as (b) but showing Fano factor quenching (post-stimulus Fano factor minus pre-stimulus Fano factor). $p_{\text{perm}} = 0.00799$. **e** Contrast modulation of Fano factor quenching. Saturation indicates stimulus contrast level; thick lines show the mean within each stimulus contrast level; shaded areas indicate 95% confidence intervals. **f** Same as (b), but for contrast modulation slope. Slopes were computed for each neuron by fitting a linear relationship between stimulus contrast and the change in baseline-corrected mean-subtracted Fano factor. $p_{\text{perm}} = 0.03197$.

specific analyses revealed age-related attenuation in variability quenching in visual and motor cortex (VISp + pm, MOs), striatum (LS, ACB), and thalamic area LP (Fig. 4d, right panel and Supplementary Fig. 15).

Contrast-modulated Fano factor quenching changed weakly with age. Similar to the contrast modulation of firing rate, we examined how the Fano factor varied with stimulus contrast (Fig. 4e and Supplementary Fig. 16). We again computed a contrast-modulation slope for each neuron by fitting a linear relationship between stimulus contrast levels and the change in baseline-corrected Fano factor (or quench index), revealing a weak, negative effect of age on contrast modulation overall (Fig. 4f, left panel). Region-specific analyses identified age-related decreases in the contrast-modulation slope in midbrain regions (SCm, MBm), and strong evidence for no age-related changes in the striatum (LS) and thalamus (LP) (Fig. 4f, right panel and Supplementary Fig. 17b).

Lastly, although we here focus on animals’ chronological age, neural differences may also be driven by other factors correlated with age. Here, we specifically investigate two sources of interest: training duration and movement.

Age effects on neural variability persist after accounting for training duration

Age-related effects measured at recording could, in principle, be entangled with differences in training history, as mice vary substantially in learning speed. Although individual training duration explains only a small fraction of variance in age at recording within our dataset (Supplementary Fig. 2), and we did not observe slower learning in older mice (Supplementary Fig. 1), training duration in the IBL task

varies systematically across animals and laboratories⁶⁴. Thus, the neural differences we observe could reflect older animals having followed a longer training trajectory rather than chronological age per se. To address this alternative explanation, we quantified each animal’s training duration and tested its association with our neural metrics.

To directly compare the two predictors, we fitted extended models including both age and training duration for each metric. Across metrics, the evidence for age effects was essentially unchanged after adding training duration (Supplementary Tables 1 and 2): pre- and post-stimulus firing rates, post-stimulus Fano factor, and FF quenching all retained strong evidence for an age effect in both the age-only and extended models.

Training duration showed limited, metric-specific contributions and did not change conclusions about effects of age (Supplementary Table 2). Training duration provided strong additional evidence for post-stimulus firing rate and moderate additional evidence for pre-, post-stimulus Fano factor, but did not reduce the evidence for age. For several metrics (e.g., pre-stimulus firing rate, firing rate contrast-modulation slope, FF contrast-modulation slope), the extended models instead provided strong evidence against a training-duration effect. Critically, for Fano factor quenching, the joint models continued to favour an age effect while providing moderate evidence for the absence of a training-duration effect, indicating that reduced neural variability quenching in older animals is not explained by differences in training duration.

Age effects are robust to video-based movement covariates

Neural activity is modulated by movement^{65–67}, including in well-trained animals⁶⁸, motivating a control analysis to test whether the

reported age-related neural differences could be explained by movement. Using IBL right-camera videos, we extracted time-resolved paw speed (task-relevant) and nose-tip speed (less task-specific) (Supplementary Fig. 18a, d) from DeepLabCut tracking. Across QC-filtered sessions with valid right-camera tracking ($n = 343$), neither movement measure showed a systematic association with age in the pre- and post-stimulus windows used to define our neural metrics (Supplementary Fig. 18b, c, e, f). We next incorporated both movement measures (paw speed and nose-tip speed) as covariates in extended models for each neural metric (Supplementary Table 3). Movement covariates provided additional explanatory power in several models, but the qualitative evidence for age effects was broadly preserved after including them. Only two metrics with initially weak-to-moderate evidence were sensitive to the inclusion of movement covariates (firing rate contrast modulation: moderate H0 to weak H1; Fano factor contrast modulation: weak H1 to moderate H1). All other metrics retained the same qualitative conclusion.

To further test whether the age-related neural effects could be accounted for by movement variability beyond average movement speed, we performed an additional control analysis using the coefficient of variation (CV) in trial-to-trial right-paw and nose-tip speed within the same pre- and post-stimulus windows. Session-level movement variability did not show consistent evidence for a systematic association with age, across measures and time windows (Supplementary Fig. 19). Although these covariates explained additional variance in several cases, adding these movement-variability measures as covariates to the extended models left the main age-related pattern broadly unchanged (Supplementary Table 4).

Together, these analyses suggest that while video-derived movement explains additional variance in firing-rate and variability metrics, the reported age-related neural effects persist after accounting for both movement magnitude and movement variability.

Discussion

In this study, we investigated age-related changes in behaviour and neural activity using large-scale extracellular recordings in mice. Older mice performed slightly worse (showing increases in absolute bias, perceptual thresholds, and lapse rate) and showed increased variability in their RTs. We observed global changes in single-neuron baseline firing rates, post-stimulus Fano factors, and an age-related reduction in the amplitude of stimulus-induced variability quenching. Older animals showed higher overall firing rates across areas of visual and motor cortex, striatum, midbrain, and hippocampus, but lower firing rates in thalamic areas. Age-related attenuation in stimulus-induced variability quenching was most prominent in visual and motor cortex, striatum, and thalamic area LP.

Older animals' slightly worse task performance could be due to a combination of peripheral (sensory, motor) and cognitive sources. Age-related decline in sensory processing, particularly in vision, is well-documented across species^{69–73}. However, visual deficits in C57BL/6 mouse strains start manifesting most dramatically after 20 months, beyond the age range of this study⁷⁴. Crucially, a sole failure to process the sensory stimulus (due to poor eyesight or the inability to hear the stimulus onset tone) would lead to an overall slowing of responses in older animals, which we did not observe. Our findings thus suggest that cognitive and decision factors also contribute to the behavioural and neural effects we observed.

The more variable RTs in older mice align with a well-established pattern in human ageing research: in human cohorts of a roughly equivalent age range, response time variability increases in a large range of cognitive tasks^{5,49,51}. The same age-related increase in RT variability in mice provides evidence for cross-species consistency in behavioural variability, highlighting a potentially fundamental aspect of behavioural ageing conserved across mammalian species.

The slight reduction in overall neural yield observed in older mice may reflect a combination of technical (recording-related) and biological factors. From a technical perspective, increased tissue and dural rigidity during probe insertion may increase superficial damage and reduce the number of well-isolated units; however, the regional specificity of the effect suggests that insertion-related factors alone are unlikely to account for the observed pattern. Biologically, ageing has been associated with modest and region-specific neuronal loss in some cortical and subcortical regions^{75–77}. More broadly, ageing is accompanied by cytoarchitectural and structural remodeling, including changes in cellular organization, dendritic morphology, and synaptic integrity^{78,79}. Such age-related changes may reduce the number of stable, well-isolated units obtained in large-scale extracellular recordings, even in the absence of pronounced neuron loss.

Our investigation across 16 brain regions showed age-related increases in single-neuron firing rates. This finding is in line with several studies reporting increased firing rates in aged primary sensory cortices, including cat V1¹⁹ and macaque V1²⁰, and in specific cell types in monkey prefrontal cortex, including pyramidal cells¹⁸. We also found that post-stimulus Fano factors increased with age, consistent with previous findings from anaesthetized rhesus monkeys²¹. Future work could also examine complementary measures of neural variability beyond Fano factor, such as spike train irregularity^{80,81}, to test the robustness and specificity of these effects. With our broad sampling of neurons across cortical and subcortical brain regions, surpassing that of prior studies, our findings indicate that age-related changes in firing rates and neural variability may be present across species, brain regions, and behavioural states.

We also observed significant age-related differences in stimulus-induced variability quenching. Previous studies found that neural variability decreases (quenches) upon stimulus onset in various cortical regions and species, including a wide range of cortical regions of monkeys^{25,29}, the PPC of rats²⁷, and the olfactory cortex of both mice²⁶ and rats²⁸, although thalamic nuclei did not show stimulus-induced quenching of neural variability²⁹. Here, we found an age-related reduction in variability quenching, most prominently in the visual and motor cortex (VISp + pm, MOs), striatum (LS, ACB), and thalamic area LP. In older animals, attenuated (i.e., less strongly negative) variability quenching may reflect a lower temporal precision of task-relevant neural processes underlying decision-making that in turn drives greater trial-to-trial variability in RTs. Future work could further explore and quantify how these single-neuron changes in variability propagate to neural circuits and networks, and lead to the computations that drive age-related changes in behaviour.

Changes in single-neuron variability may arise from various biological mechanisms. Age-related degradation of inhibitory and neuromodulatory systems^{5,51,82,83} may cause reduced gain modulation, shorter intrinsic neural timescales^{84,85}, and lower temporal precision of single-neuron responses. These could in turn be reflected in pairwise noise correlations^{23,86}, neural population dynamics^{16,87,88}, and cortical state (e.g., as reflected in multi-unit activity power spectra)^{62,89}. Future work on such population-level measures may further reveal how age-related neural variability across scales impacts behaviour.

While we here used animals' chronological age at the time of their neural recording, this does not imply that age itself is the best predictor of the neural effects we observed. More broadly, the correlational nature of these analyses highlights a general challenge in interpreting ageing effects: chronological age covaries with multiple behavioural and biological factors that may jointly shape neural activity and variability, and these influences cannot be fully disentangled in the current dataset. Specifically, other factors such as training duration, movement patterns during task performance, or different decision strategies may give rise to different task strategies or different computational mechanisms that produce similar overt task behaviour. While we controlled for video-derived movement

covariates (paw, nose-tip), future work could further refine such integrated predictors using richer behavioural characterizations⁹⁰, for example, motion–energy–based video features or higher-dimensional, data-driven representations of behavioural modules^{91,92}. Accordingly, a range of biological, behavioural, and life history factors could be further integrated to compute an individual animal's brain age⁹³ as opposed to their chronological age, which may reveal higher predictive power in explaining between-animal differences in behavioural and neural variability.

In conclusion, we present a survey of age-related neural variability in >18,000 neurons across 16 brain regions in behaving mice, complementing the previously released brain-wide map recording in younger animals⁴⁵. Our findings replicate earlier work on neural correlates of ageing in other species and brain areas, and show how multi-area neural recordings in behaving animals can help uncover regional specificity of age-related changes in neural variability. Lastly, we provide a well-curated, open dataset in a standardized data format⁶⁴ that further allows investigating the computational mechanisms through which neural variability gives rise to age-related cognitive decline.

Methods

Data

This study included a previously released dataset (the brain-wide map dataset)⁴⁵ from the IBL and our new release (the lifespan dataset). Both datasets contain extracellular recordings of mice performing a visual decision-making task, collected using the same standardized protocols^{94,95}. All experimental procedures for the newly collected lifespan dataset were conducted in accordance with local laws and approved by the institutional IACUC at Cold Spring Harbor Laboratory (licences 141117 and 19.5).

Animals

We included 130 mice from a previously released dataset (C57BL/6, 89 male, mean age = 6.64 months, range 3.10–15.13 months) and 19 mice from a newly recorded dataset (C57BL/6, 11 male, mean age = 16.58 months, range 10.58–19.90 months). All mice were housed and cared for following standardized protocols⁴³. Age was calculated on the day of electrophysiological recording. Across mice, age at recording varied due to differences in (1) age at the start of behavioural training, (2) training duration, and (3) non-training intervals between training start and the first recording (e.g., experimental or logistical constraints). A decomposition of these sources indicates that variability is dominated by differences in age at training start (Supplementary Fig. 2). After quality control, 149 mice were included in the analysis. For visualization, mice were categorized based on their age at recording into a young group ($N = 97$, mean age 5.60 months) and an old group ($N = 52$, mean age 12.07 months), using 7.6 months (mean age in the dataset) as the age cutoff (Fig. 1b). Age was treated as a continuous variable in all statistical analyses.

Behavioural task

Mice were trained to perform a standardized visual decision-making task⁴³. In this task, mice had to decide the location of a visual stimulus presented on the screen in front of them. Specifically, each trial began when the mouse held the wheel still for 0.4–0.7 s. Then, an auditory cue (a 100-ms tone, 5 kHz sine wave) and a visual stimulus (Gabor patch) were presented on either the left or right side of the screen. Mice indicated the stimulus location by turning the response wheel to bring the stimulus to the centre of the screen. They had up to 60 s to make a response. The contrast level of the visual stimulus varied across trials. There were five different contrast levels (%): 100, 25, 12.5, 6.25, 0. Combining the five contrast levels with the two stimulus sides (left, right) resulted in nine conditions: –100, –25, –12.5, –6.25, 0, 6.25, 12.5, 25, 100 (negative for left stimuli, positive for right stimuli). Correct responses were rewarded with sugar water, while incorrect responses

were followed by a noise burst and a longer inter-trial interval. Stimulus probabilities were fixed within each block: sessions began with an unbiased block (50:50 left vs. right) and then alternated between blocks biased toward the left (80:20) and toward the right (20:80). The task was based on the standardized IBL visual decision-making task introduced in the original task paper⁴³. Animals were trained according to the IBL training protocol⁹⁶. We here provide details relevant to the present study.

Behavioural measures

Psychometric curve. Psychometric curves illustrate how the probability of a rightward choice changes with stimulus position and stimulus contrast. For each session, responses were fitted with a parametric error function using a maximum likelihood procedure:

$$P = \gamma + \frac{(1 - \gamma - \lambda) \operatorname{erf}\left(\frac{c - \mu}{\sigma} + 1\right)}{2} \quad (1)$$

P is the probability of rightward choice;

c is the contrast level;

$\gamma, \lambda, \sigma, \mu$ are fitted parameters:

γ : lapse low (the lapse rate for left stimuli);

λ : lapse high (the lapse rate for right stimuli);

σ : the decision threshold;

μ : the response bias (horizontal shift of the curve);

For details on the fitting process and parameter computation, see the IBL protocol⁹⁶. The fitted psychometric parameters were obtained for each session. In the statistical analyses, we examined ageing effects by correlating the absolute bias $|\mu|$, threshold σ , and mean lapse $\frac{\gamma + \lambda}{2}$ to the age of mice.

Chronometric curve and RT variability. Chronometric curves show how response time varies with stimulus position and contrast. RT was defined as the time between stimulus onset (stimOn_times, when the visual stimulus appeared on the screen) and response (response_times, recorded either after 60 s, i.e., timeout, or when the rotary encoder indicated the stimulus reached $\pm 35^\circ$ azimuth). A chronometric curve was computed for each session.

To assess RT variability, we computed the coefficient of variation (CV) for each session.

$$CV = \frac{\sigma}{\mu} \quad (2)$$

Where:

σ = standard deviation of response times

μ = mean response time

Neuropixels recordings

Extracellular recordings were acquired using Neuropixels probes⁴⁷. For a detailed description of the animal surgery, apparatus, and recording procedure, see Appendix 2 and 3⁴⁴. Up to two probes were inserted during each recording session.

Spike sorting

All sessions were spike sorted using ibl-sorter (version 2.35.0)⁹⁷, matching the version used for the original IBL dataset processing⁴⁵. The ibl-sorter code is publicly available⁹⁸.

Quality control

Before analyzing the data, we applied a set of inclusion criteria (at the session, insertion, and neuron level) to ensure data quality. These criteria were primarily based on previously established IBL criteria⁴⁵, with additional requirements for single-neuron quality. These additional requirements were included to ensure that Fano factor estimates were computed from neurons with stable activity, as this metric

is particularly sensitive to unstable firing. The inclusion and exclusion criteria are described below.

Sessions and insertions. (1) Sessions with more than 400 trials and at least 90% accuracy on 100% contrast trials were included. (2) There must be at least 3 trials with incorrect responses (after applying the trials filter below). (3) Sessions had to pass the hardware tests, as defined at https://int-brain-lab.github.io/iblenv/_autosummary/ibllib.qc.task_metrics.html. (4) Sessions were visually inspected to exclude those which have pronounced instability of recordings (drift), epileptiform activity, noisy channels, and artifacts (following the Recording Inclusion metrics and Guidelines for Optimal Reproducibility (RIGOR) criteria⁴⁴). (5) Insertions must have resolved alignments (see Appendix 6⁴⁴ for definition). (6) Insertions were also required to have spike-sorted data processed using the same version of the ibl-sorter algorithm (version 2.35.0).

Trials. To control for differences in trial counts across sessions, we included only the first 400 trials of each session. Additional exclusion

criteria were applied: (1) trials missing any of the following events—choice (response type), probabilityLeft (stimulus side probability), feedbackType (positive for correct, rewarded responses; negative for incorrect or timed-out trials), feedbackTimes (time of feedback), stimOnTimes (stimulus onset), or firstMovementTimes (the time of the first detected movement with sufficient amplitude); and (2) trials with RTs outside the range of 0.08–2 s⁴⁵.

Neurons. Neurons included in the analyses met the following criteria: (1) passed three single-unit quality control (QC) metrics—a refractory period violation metric, a noise cutoff metric, and a median amplitude threshold—based on the computed single-unit metrics of RIGOR⁴⁴; (2) had an average firing rate greater than 1 spike/s; and (3) presence ratio exceeded 0.95. Final analyses were further restricted to 16 ROIs (see Table 1).

ROIs include a combination of areas from the Beryl parcellation⁴⁵. We chose this set of insertion targets for a combination of scientific (earlier age-related degradation in more frontal regions), statistical (inclusion of the repeated site that was used to confirm reproducibility across labs⁴⁴), and practical reasons (surgical accessibility).

Table 2 shows the number of sessions and insertions that survived after each quality control criterion was applied. After quality control, we included 367 recording sessions and 503 insertions in the final analyses.

Fano factor

Trial-by-trial neural variability was quantified using the Fano factor, defined as the variance of spike counts across trials divided by the mean spike count:

$$FF = \frac{\sigma^2}{\mu} \quad (3)$$

Where:

σ^2 = variance of spike count across trials

μ = mean spike count across trials

Because neural responses can differ across stimulus conditions, we first divided trials into nine stimulus conditions (side × contrast: −100, −25, −12.5, −6.25, 0, 6.25, 12.5, 25, 100%) to prevent differences in mean responses from artificially inflating variance estimates.

Mean-subtracted Fano factor. To ensure that our Fano factor analyses were not biased by differences in mean firing rates across conditions, we adapted the mean-subtraction method⁶³ to pool trials while controlling for condition-specific differences in mean firing rates. Specifically, for each trial and time window, we subtracted the condition-specific mean spike count to obtain a residual. Mean-subtracted Fano factors were then computed from the variance of these residuals using a sliding-window approach (width = 100 ms, step

Table 1 | ROI definitions

ROI	ROI full name	Beryl areas	Cosmos
MOs	Secondary motor area	MOs	Isocortex
ACA	Anterior cingulate area	ACAv, ACAd	Isocortex
CP	Caudoputamen	CP	Cerebral nuclei
LS	Lateral septal nucleus	LSr, LSc, LSV	Cerebral nuclei
ACB	Nucleus accumbens	ACB	Cerebral nuclei
mPFC	Medial prefrontal cortex	PL, ILA	Isocortex
ORB	Orbital area	ORBd, ORBm, ORBvl	Isocortex
OLF	Olfactory areas	TTd, DP, AON	Olfactory areas
VISp + pm	Visual cortex	VISp, VISpm	Isocortex
SCm	Superior colliculus motor related	SCm	Midbrain
MBm	Motor related midbrain	APN, MRN	Midbrain
PPC	Posterior parietal cortex	VISa, VISam	Isocortex
CA1	Field CA1	CA1	Hippocampal formation
DG	Dentate gyrus	DG	Hippocampal formation
LP	Lateral posterior nucleus of thalamus	LP	Thalamus
PO	Posterior complex of thalamus	PO	Thalamus

This table details the composition of each defined ROI, including its abbreviation and the constituent brain regions.

Table 2 | Quality control process

QC	Level	Previously released dataset ⁴⁵		Newly recorded dataset	
		n_sessions	n_insertions	n_sessions	n_insertions
Session and insertion QC	Insertion, session	459	699	38	64
Response time and missing events	Trials	459	699	38	64
Minimum 3 error trials	Session	459	699	38	64
Single unit QC	Neuron	458	695	38	64
ROIs only	Neuron	329	439	38	64
Minimum 2 sessions per region	Neuron	329	439	38	64

The table shows the step-by-step filtering of sessions and insertions, using the inclusion criteria described in the main text. The numbers in the first row are based on the data release.

size = 20 ms) for each neuron.

$$FF_{\langle N_i \rangle} = \frac{\text{Var}[Z_U]}{\sum_{i=1}^M \frac{n_i}{N_i} N_i} \quad (4)$$

It is defined as the variance of the union of residuals from all conditions Z_U , divided by the weighted average of the mean spike count for the conditions, where n_U is the total number of trials across all M conditions, n_i and N_i are the number of trials and the mean count for the i -th condition, respectively. Statistical tests were performed on these mean-subtracted Fano factors.

Statistical tests

We assessed the effect of age on each behavioural and neural metric. Behavioural metrics (absolute bias, threshold, mean lapse rate, median RT, and RT variability) were computed at the session level. Neural metrics included neural yield, firing rate, contrast modulation of firing rate, Fano factor, Fano factor quench, and contrast modulation of Fano factor. For each neuron, contrast-modulation slopes were obtained by fitting a linear regression to the change in firing rate (Δ firing rate, post-minus pre-stimulus window) or Fano factor (Δ Fano factor, post-minus pre-stimulus window) across all contrast levels. Neural yield was defined and analysed at the probe-insertion level, whereas all other neural metrics were computed at the single-neuron level, and all statistical analyses were performed on these neuron-level metrics.

To test for age-related effects, we used the age-related slope, β_{age} , from a general linear model (Gaussian family, identity link, using a generalized linear modelling framework) as the test statistic. Strictly positive and right-skewed measures (median RT, psychometric threshold, firing rates, Fano factors) were analysed on a log scale to reduce skewness and stabilise variance. In contrast, RT variability, signed change indices (e.g., FF quench and contrast modulation metrics), nonnegative metrics with substantial mass near zero (absolute bias), and bounded proportions (mean lapse) were analysed on their natural scale, since log transforms are undefined at zero and can introduce artificial heavy lower tails when values are concentrated near zero. We inspected residual Q-Q plots for representative models; residuals for log-transformed metrics showed approximate normality with mild tail deviations, whereas bounded and difference metrics exhibited somewhat heavier but unimodal tails. Importantly, we verified that our conclusions did not change depending on the log-scaling of these variables.

The following metric-specific transformations and covariates were used:

Behavioural metrics. Median RT and psychometric threshold were analysed as $\log(y)$, with age as the only predictor. RT variability, absolute bias, and mean lapse were analysed on their natural scale, with age as the only predictor.

Neural metrics. Neural yield was analysed on its natural scale with age and brain region as predictors. Pre- and post-stimulus firing rates were analysed as $\log(y)$, with age, brain region, absolute signed contrast, and trial count as predictors. Firing rate contrast modulation slope was analysed on its natural scale with age and brain region as predictors. Pre- and post-stimulus Fano factors were analysed as $\log(y)$, with age, brain region, and trial count as predictors. Fano factor quench was analysed on its natural scale with age, brain region, and trial count as predictors. Fano factor contrast modulation slope was analysed on its natural scale with age and brain region as predictors.

Bayes Factors. We computed Bayes Factors (BF_{10}), using the BayesFactor⁹⁹ and Pingouin¹⁰⁰ packages, to quantify evidence for including mouse age as a predictor in each model. BF_{10} compares models with and without age, with values > 1 indicating support for an

age effect (i.e., age modulates the behavioural or neural metric) and < 1 favouring the null. Statistical conclusions throughout the manuscript were based on BF_{10} .

For visualization purposes, regression lines are plotted only when the BF_{10} provides strong ($BF_{10} > 10$) or moderate ($10 \geq BF_{10} > 3$) evidence in favour of the alternative hypothesis (H1). While the choice of these thresholds is inherently arbitrary, we adopted them to ensure that plotted regression lines correspond to effects with at least moderate statistical evidence.

Permutation tests. We additionally computed permutation-based p -values, which are shown alongside the Bayes Factors. Note that all conclusions are based on the Bayes Factors, which provide a primary graded measure of evidence for or against an effect. Permutation tests additionally account for the hierarchical data structure by shuffling age labels at the session level. In omnibus analyses, both methods led to the same conclusions. For region-specific results, we highlighted regions with strong or moderate evidence for the alternative (H1), as well as those with strong evidence for the null (H0).

To compute permutation-based p -values, the null hypothesis was that there is no association between the metric and mouse age. To generate a null distribution for each analysis, we performed a permutation procedure ($n = 10,000$ iterations for behavioural metrics; $n = 2000$ iterations for neural metrics). Importantly, for the neural metrics (except for neural yield), we randomly shuffled the mouse age labels across recording sessions in each iteration while maintaining the grouping of neurons recorded within each session. This session-based shuffling preserved the inherent dependencies among neurons recorded simultaneously. After each shuffle, we recalculated the test statistics for the specific metric being examined. The p -value for each metric was then calculated as the proportion of the permuted age slopes whose absolute values were equal to or greater than the absolute value of the observed slope from the original data. This two-tailed p -value represents the probability of observing a relationship as strong as, or stronger than, the one found, under the null hypothesis of no relationship between the metric and age.

Extended models including age and training duration. To assess whether our age effects could be explained by training duration, we fitted extended models for key metrics that included both age and training duration as predictors simultaneously. Training duration was defined as the number of training days until the mouse first passed the standardized behavioural criterion and was expressed in years to match the age scale used in the model. All other modelling choices (predictors, covariates, and log-transform conventions) were identical to those described above.

In each extended model, we quantified evidence for including each predictor (age, training duration) using Bayes Factors that compared otherwise identical models with and without the respective term. This allowed us to test whether age effects persisted after accounting for training duration, and whether training duration explained additional variance beyond age.

Extended models including age and movement. We used IBL right-camera videos to derive movement covariates from DeepLabCut tracking of two representative keypoints: right paw (task-relevant) and nose tip (task-irrelevant) (Supplementary Fig. 18a, d). From the tracked x/y coordinates, we computed frame-to-frame speed (coordinates normalized by camera resolution and scaled by the camera sampling rate), and aligned the resulting speed time courses to stimulus onset. We summarized movement by computing median speed within the same pre- and post-stimulus windows used for neural metrics (pre: -100 to 0 ms; post: 160 to 260 ms), yielding pre- and post-stimulus movement measures for both paw and nose tip. To quantify movement variability, we additionally computed the coefficient of variation

(CV; standard deviation/mean) in trial-to-trial paw and nose-tip speed within the same pre- and post-stimulus windows.

Sessions were required to pass IBL extended quality control for the right camera (videoRight and dlcRight). This yielded $n = 343$ sessions (out of total 367 sessions) with valid right-camera movement measures. We then fitted extended statistical models for each neural metric by adding either movement-speed covariates or movement-variability covariates to the age-only models. Movement covariates were matched to each neural metric by time window and contrast structure. Metrics defined in a single window used the corresponding pre- or post-stimulus movement measure, whereas metrics spanning both windows included both pre- and post-stimulus covariates. For neural metrics retaining contrast-specific values (pre- and post-stimulus firing rate), movement measures were computed separately for each session and contrast condition; for neural metrics not retaining contrast (contrast modulation of firing rate, pre- and post-stimulus Fano factor, Fano factor quench, and contrast modulation of Fano factor), movement measures were then aggregated across contrasts to yield one value per session, window, and feature. We applied the same matching scheme in the movement-variability control models. All other modelling choices (predictors, covariates, log-transform conventions, and inference procedures) were identical to those described above.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The data analysed in this study are publicly available through the International Brain Laboratory website (<https://www.internationalbrainlab.com/data>), under the tag “2025_Q3_Zang_et_al_Aging”. Instructions for accessing these data are provided in the GitHub repository associated with this study (https://github.com/Fenyng-Zang/Ageing_behavioral_and_neural_variability). The analysis results underlying the figures are also available in the same repository.

Code availability

Data analyses were performed using custom Python scripts in a reproducible Python environment. All code required to reproduce the figures and analyses is available at https://github.com/Fenyng-Zang/Ageing_behavioral_and_neural_variability, and has been archived on Zenodo¹⁰¹.

References

- Cremer, R. & Zeef, E. J. What kind of noise increases with age? *J. Gerontol.* **42**, 515–518 (1987).
- Crossman, E. R. & Szafran, J. Changes with age in the speed of information-intake and discrimination. *Experientia* **128**, 135 (1956).
- Salthouse, T. A. & Lichty, W. Tests of the neural noise hypothesis of age-related cognitive change. *J. Gerontol.* **40**, 443–450 (1985).
- Welford, A. T. Signal, noise, performance, and age. *Hum. Factors J. Hum. Factors Ergon. Soc.* **23**, 97–109 (1981).
- Li, S.-C. & Rieckmann, A. Neuromodulation and aging: implications of aging neuronal gain control on cognition. *Curr. Opin. Neurobiol.* **29**, 148–158 (2014).
- Koen, J. D. & Rugg, M. D. Neural dedifferentiation in the aging brain. *Trends Cogn. Sci.* **23**, 547–559 (2019).
- Boylan, M. A. et al. Greater BOLD variability is associated with poorer cognitive function in an adult lifespan sample. *Cereb. Cortex* **31**, 562–574 (2021).
- Garrett, D. D., Kovacevic, N., McIntosh, A. R. & Grady, C. L. Blood oxygen level-dependent signal variability is more than just noise. *J. Neurosci.* **30**, 4914–4921 (2010).
- Garrett, D. D., Kovacevic, N., McIntosh, A. R. & Grady, C. L. The importance of being variable. *J. Neurosci.* **31**, 4496–4503 (2011).
- Nomi, J. S., Bolt, T. S., Ezie, C. E. C., Uddin, L. Q. & Heller, A. S. Moment-to-moment BOLD signal variability reflects regional changes in neural flexibility across the lifespan. *J. Neurosci.* **37**, 5539–5548 (2017).
- Samanez-Larkin, G. R., Kuhnen, C. M., Yoo, D. J. & Knutson, B. Variability in nucleus accumbens activity mediates age-related suboptimal financial risk taking. *J. Neurosci.* **30**, 1426–1434 (2010).
- Kügler, C. F. A., Taghavy, A. & Platt, D. The event-related P300 potential analysis of cognitive human brain aging: a review. *Gerontology* **39**, 280–303 (1993).
- Waschke, L., Wöstmann, M. & Obleser, J. States and traits of neural irregularity in the age-varying human brain. *Sci. Rep.* **7**, 17381 (2017).
- Ribeiro, M. & Castelo-Branco, M. Slow fluctuations in ongoing brain activity decrease in amplitude with ageing yet their impact on task-related evoked responses is dissociable from behavior. *eLife* **11**, e75722 (2022).
- Tran, T. T., Rolle, C. E., Gazzaley, A. & Voytek, B. Linked sources of neural noise contribute to age-related cognitive decline. *J. Cogn. Neurosci.* **32**, 1813–1822 (2020).
- Voytek, B. et al. Age-related changes in 1/f neural electrophysiological noise. *J. Neurosci.* **35**, 13257–13265 (2015).
- Schmidt, F. et al. Age-related changes in “cortical” 1/f dynamics are linked to cardiac activity. *eLife* **13**, RP100605 (2024).
- Chang, Y.-M., Rosene, D. L., Killiany, R. J., Mangiamele, L. A. & Luebke, J. I. Increased action potential firing rates of layer 2/3 pyramidal cells in the prefrontal cortex are significantly related to cognitive performance in aged monkeys. *Cereb. Cortex* **15**, 409–418 (2005).
- Hua, T. et al. Functional degradation of visual cortical cells in old cats. *Neurobiol. Aging* **27**, 155–162 (2006).
- Schmolecky, M. T., Wang, Y., Pu, M. & Leventhal, A. G. Degradation of stimulus selectivity of visual cortical cells in senescent rhesus monkeys. *Nat. Neurosci.* **3**, 384–390 (2000).
- Yang, Y., Liang, Z., Li, G., Wang, Y. & Zhou, Y. Aging affects response variability of V1 and MT neurons in rhesus monkeys. *Brain Res.* **1274**, 21–27 (2009).
- Wang, M. et al. Neuronal basis of age-related working memory decline. *Nature* **476**, 210–213 (2011).
- Wang, X. et al. Aging affects correlation within the V1 neuronal population in rhesus monkeys. *Neurobiol. Aging* **76**, 1–8 (2019).
- Liu, X., Zhou, Y., Liu, J. & Xu, G. Metformin improves age-related visual cortex dysfunction in mice by reducing noise correlation in the primary visual cortex. *Front. Aging Neurosci.* **17**, 1572653 (2025).
- Churchland, M. M. et al. Stimulus onset quenches neural variability: a widespread cortical phenomenon. *Nat. Neurosci.* **13**, 369–378 (2010).
- Iurilli, G. & Datta, S. R. Population coding in an innately relevant olfactory area. *Neuron* **93**, 1180–1197.e7 (2017).
- Licata, A. M. et al. Posterior parietal cortex guides visual decisions in rats. *J. Neurosci.* **37**, 4954–4966 (2017).
- Miura, K., Mainen, Z. F. & Uchida, N. Odor representations in olfactory cortex: distributed rate coding and decorrelated population activity. *Neuron* **74**, 1087–1098 (2012).
- Poland, E., Donner, T. H., Müller, K.-M., Leopold, D. A. & Wilke, M. Thalamus exhibits less sensory variability quenching than cortex. *Sci. Rep.* **9**, 7590 (2019).
- Randall, A. D., Booth, C. & Brown, J. T. Age-related changes to Na⁺ channel gating contribute to modified intrinsic neuronal excitability. *Neurobiol. Aging* **33**, 2715–2720 (2012).

31. Simkin, D. et al. Aging-related hyperexcitability in CA3 pyramidal neurons is mediated by enhanced A-type K⁺ channel function and expression. *J. Neurosci.* **35**, 13206–13218 (2015).
32. Turner, J. G., Hughes, L. F. & Caspary, D. M. Affects of aging on receptive fields in rat primary auditory cortex layer V neurons. *J. Neurophysiol.* **94**, 2738–2747 (2005).
33. Hughes, L. F., Turner, J. G., Parrish, J. L. & Caspary, D. M. Processing of broadband stimuli across A1 layers in young and aged rats. *Hear. Res.* **264**, 79–85 (2010).
34. Maziar, A. et al. Aging differentially affects LTCC function in hippocampal CA1 and piriform cortex pyramidal neurons. *Cereb. Cortex* **33**, 1489–1503 (2023).
35. Barnes, C. A., Suster, M. S., Shen, J. & McNaughton, B. L. Multi-stability of cognitive maps in the hippocampus of old rats. *Nature* **388**, 272–275 (1997).
36. Rosenzweig, E. S. & Barnes, C. A. Impact of aging on hippocampal function: plasticity, network dynamics, and cognition. *Prog. Neurobiol.* **69**, 143–179 (2003).
37. Barnes, C. A. Secrets of aging: what does a normally aging brain look like? *F1000 Biol. Rep.* **3**, 22 (2011).
38. Wang, H.-L. et al. Age-related alterations of neuronal excitability and voltage-dependent Ca²⁺ current in a spontaneous mouse model of Alzheimer's disease. *Behav. Brain Res.* **321**, 209–213 (2017).
39. Sosulina, L. et al. Hippocampal hyperactivity in a rat model of Alzheimer's disease. *J. Neurochem.* **157**, 2128–2144 (2021).
40. Anastacio, H. T. D., Matosin, N. & Ooi, L. Neuronal hyperexcitability in Alzheimer's disease: what are the drivers behind this aberrant phenotype? *Transl. Psychiatry* **12**, 257 (2022).
41. Urai, A. E., Doiron, B., Leifer, A. M. & Churchland, A. K. Large-scale neural recordings call for new insights to link brain and behavior. *Nat. Neurosci.* **25**, 11–19 (2022).
42. International Brain Laboratory. An international laboratory for systems and computational neuroscience. *Neuron* **96**, 1213–1218 (2017).
43. International Brain Laboratory et al. Standardized and reproducible measurement of decision-making in mice. *eLife* **10**, e63711 (2021).
44. International Brain Laboratory et al. Reproducibility of in vivo electrophysiological measurements in mice. *eLife* **13**, RP100840 (2025).
45. International Brain Laboratory et al. A brain-wide map of neural activity during complex behaviour. *Nature* **645**, 177–191 (2025).
46. Cottam, N. C. et al. From circuits to lifespan: translating mouse and human timelines with neuroimaging-based tractography. *J. Neurosci.* **45**, e1429242025 (2025).
47. Jun, J. J. et al. Fully integrated silicon probes for high-density recording of neural activity. *Nature* **551**, 232–236 (2017).
48. Findling, C. et al. Brain-wide representations of prior information in mouse decision-making. *Nature* **645**, 192–200 (2025).
49. Dykiert, D., Der, G., Starr, J. M. & Deary, I. J. Age differences in intra-individual variability in simple and choice reaction time: systematic review and meta-analysis. *PLOS ONE* **7**, e45759 (2012).
50. LaPlume, A. A., Anderson, N. D., McKetton, L., Levine, B. & Troyer, A. K. When I'm 64: age-related variability in over 40,000 online cognitive test takers. *J. Gerontol. Ser. B* **77**, 104–117 (2022).
51. MacDonald, S. W. S., Nyberg, L. & Bäckman, L. Intra-individual variability in behavior: links to brain structure, neurotransmission and neuronal activity. *Trends Neurosci.* **29**, 474–480 (2006).
52. Bishop, R., Qureshi, F. & Yan, J. Age-related changes in neuronal receptive fields of primary auditory cortex in frequency, amplitude, and temporal domains. *Hear. Res.* **420**, 108504 (2022).
53. Wilson, I. A., Ikonen, S., Gallagher, M., Eichenbaum, H. & Tanila, H. Age-Associated alterations of hippocampal place cells are sub-region specific. *J. Neurosci.* **25**, 6877–6886 (2005).
54. Hwang, E. J. et al. Parietal stimulation reverses age-related decline in exploration, learning, and decision-making. Preprint at <https://doi.org/10.1101/2023.10.21.563408> (2023).
55. Leong, U.-C., Barsz, K., Allen, P. D. & Walton, J. P. Neural correlates of age-related declines in frequency selectivity in the auditory midbrain. *Neurobiol. Aging* **32**, 168–178 (2011).
56. Dean, A. F. The variability of discharge of simple cells in the cat striate cortex. *Exp. Brain Res.* **44**, 437–440 (1981).
57. Tolhurst, D. J., Movshon, J. A. & Thompson, I. D. The dependence of response amplitude and variance of cat visual cortical neurones on stimulus contrast. *Exp. Brain Res.* **41**, 414–419 (1981).
58. Tolhurst, D. J., Movshon, J. A. & Dean, A. F. The statistical reliability of signals in single neurons in cat and monkey visual cortex. *Vision Res.* **23**, 775–785 (1983).
59. Tomko, G. J. & Crapper, D. R. Neuronal variability: non-stationary responses to identical visual stimuli. *Brain Res.* **79**, 405–418 (1974).
60. Goris, R. L. T., Movshon, J. A. & Simoncelli, E. P. Partitioning neuronal variability. *Nat. Neurosci.* **17**, 858–865 (2014).
61. Kara, P., Reinagel, P. & Reid, R. C. Low response variability in simultaneously recorded retinal, thalamic, and cortical neurons. *Neuron* **27**, 635–646 (2000).
62. Schölvinc, M. L., Saleem, A. B., Benucci, A., Harris, K. D. & Carandini, M. Cortical state determines global variability and correlations in visual cortex. *J. Neurosci.* **35**, 170–178 (2015).
63. Churchland, A. K. et al. Variance as a signature of neural computations during decision making. *Neuron* **69**, 818–831 (2011).
64. International Brain Laboratory et al. A modular architecture for organizing, processing and sharing neurophysiology data. *Nat. Methods* **20**, 403–407 (2023).
65. Musall, S., Kaufman, M. T., Juavinett, A. L., Gluf, S. & Churchland, A. K. Single-trial neural dynamics are dominated by richly varied movements. *Nat. Neurosci.* **22**, 1677–1686 (2019).
66. Salkoff, D. B., Zagha, E., McCarthy, E. & McCormick, D. A. Movement and performance explain widespread cortical activity in a visual detection task. *Cereb. Cortex* **30**, 421–437 (2020).
67. Stringer, C. et al. Spontaneous behaviors drive multidimensional, brainwide activity. *Science* **364**, eaav7893 (2019).
68. Yin, C. et al. Spontaneous movements and their relationship to neural activity fluctuate with latent engagement states. *Neuron* **113**, 3048–3063.e5 (2025).
69. Greene, H. A. & Madden, D. J. Adult age differences in visual acuity, stereopsis, and contrast sensitivity. *Optom. Vis. Sci.* **64**, 749 (1987).
70. Nameda, N., Kawara, T. & Ohzu, H. Human visual spatio-temporal frequency performance as a function of age. *Optom. Vis. Sci.* **66**, 760 (1989).
71. Owsley, C. Vision and aging. *Annu. Rev. Vis. Sci.* **2**, 255–271 (2016).
72. Owsley, C., Sekuler, R. & Siemsen, D. Contrast sensitivity throughout adulthood. *Vision Res.* **23**, 689–699 (1983).
73. Owsley, C. & Sloane, M. E. Contrast sensitivity, acuity, and the perception of 'real-world' targets. *Br. J. Ophthalmol.* **71**, 791–796 (1987).
74. Lehmann, K., Schmidt, K.-F. & Löwel, S. Vision and visual plasticity in ageing mice. *Restor. Neurol. Neurosci.* **30**, 161–178 (2012).
75. Sabel, B. A. & Stein, D. G. Extensive loss of subcortical neurons in the aging rat brain. *Exp. Neurol.* **73**, 507–516 (1981).
76. Stranahan, A. M., Jiam, N. T., Spiegel, A. M. & Gallagher, M. Aging reduces total neuron number in the dorsal component of the rodent prefrontal cortex. *J. Comp. Neurol.* **520**, 1318–1326 (2012).
77. Yates, M. A., Markham, J. A., Anderson, S. E., Morris, J. R. & Juraska, J. M. Regional variability in age-related loss of neurons from the primary visual cortex and medial prefrontal cortex of male and female rats. *Brain Res.* **1218**, 1–12 (2008).

78. Markham, J. A. & Juraska, J. M. Aging and sex influence the anatomy of the rat anterior cingulate cortex. *Neurobiol. Aging* **23**, 579–588 (2002).
79. Morrison, J. H. & Baxter, M. G. The aging cortical synapse: hallmarks and implications for cognitive decline. *Nat. Rev. Neurosci.* **13**, 240–250 (2012).
80. Holt, G. R., Softky, W. R., Koch, C. & Douglas, R. J. Comparison of discharge variability in vitro and in vivo in cat visual cortex neurons. *J. Neurophysiol.* **75**, 1806–1814 (1996).
81. Nawrot, M. P. et al. Measurement of variability dynamics in cortical spike trains. *J. Neurosci. Methods* **169**, 374–390 (2008).
82. Guitart-Masip, M. et al. BOLD variability is related to dopaminergic neurotransmission and cognitive aging. *Cereb. Cortex* **26**, 2074–2083 (2016).
83. Lee, J. & Kim, H.-J. Normal aging induces changes in the brain and neurodegeneration progress: review of the structural, biochemical, metabolic, cellular, and molecular changes. *Front. Aging Neurosci.* **14**, 931536 (2022).
84. Gao, R., van den Brink, R. L., Pfeffer, T. & Voytek, B. Neuronal timescales are functionally dynamic and shaped by cortical microarchitecture. *eLife* **9**, e61277 (2020).
85. Wu, K. & Gollo, L. L. Mapping and modeling age-related changes in intrinsic neural timescales. *Commun. Biol.* **8**, 1–16 (2025).
86. Cohen, M. R. & Maunsell, J. H. R. Attention improves performance primarily by reducing interneuronal correlations. *Nat. Neurosci.* **12**, 1594–1600 (2009).
87. Jauny, G., Eustache, F. & Hinault, T. Connectivity dynamics and cognitive variability during aging. *Neurobiol. Aging* **118**, 99–105 (2022).
88. Wen, W. et al. Beta-band neural variability reveals age-related dissociations in human working memory maintenance and deletion. *PLoS Biol.* **22**, e3002784 (2024).
89. Harris, K. D. & Thiele, A. Cortical state and attention. *Nat. Rev. Neurosci.* **12**, 509–523 (2011).
90. Musall, S., Urai, A. E., Sussillo, D. & Churchland, A. K. Harnessing behavioral diversity to understand neural computations for cognition. *Curr. Opin. Neurobiol.* **58**, 229–238 (2019).
91. Wiltchko, A. B. et al. Mapping sub-second structure in mouse behavior. *Neuron* **88**, 1121–1135 (2015).
92. Wiltchko, A. B. et al. Revealing the structure of pharmacobehavioral space through motion sequencing. *Nat. Neurosci.* **23**, 1433–1443 (2020).
93. Cole, J. H., Marioni, R. E., Harris, S. E. & Deary, I. J. Brain age and other bodily ‘ages’: implications for neuropsychiatry. *Mol. Psychiatry* **24**, 266–281 (2019).
94. International Brain Laboratory. Appendix 2: IBL protocol for electrophysiology recording using Neuropixels probe. *figshare* https://figshare.com/articles/online_resource/Appendix_2_IBL_protocol_for_electrophysiology_recording_using_Neuropixels_probe/19697896/2 (2022).
95. Urai, A. E. et al. Citric acid water as an alternative to water restriction for high-yield mouse behavior. *eNeuro* **8**, ENEURO.0230-20.2020 (2021).
96. International Brain Laboratory. Behavior: Appendix 2: IBL protocol for mice training. *figshare* <https://doi.org/10.6084/m9.figshare.11634729.v3> (2020).
97. International Brain Laboratory. Spike sorting pipeline for the International Brain Laboratory. *figshare* https://figshare.com/articles/online_resource/Spike_sorting_pipeline_for_the_International_Brain_Laboratory/19705522/4 (2022).
98. International Brain Laboratory et al. ibl-sorter. <https://github.com/int-brain-lab/ibl-sorter> (2024).
99. Morey R. & Roudier J. *BayesFactor: Computation of Bayes Factors for Common Designs* (2023).
100. Vallat, R. Pingouin: statistics in Python. *J. Open Source Softw.* **3**, 1026 (2018).
101. Zang, F. & Rai, P. Age-related changes in behavioral and neural variability in a decision-making task. *Zenodo* <https://doi.org/10.5281/ZENODO.20102008> (2026).
102. Allen Institute for Brain Science. *Allen Mouse Brain Atlas [dataset]*. (2004).
103. Allen Institute for Brain Science. *Allen Reference Atlas – Mouse Brain [Brain Atlas]*. (2011).

Acknowledgements

A.E.U. thanks Graham Wildt and John Pisciotto for assistance with laboratory safety procedures, Rachel Rubino for excellent animal care advice, and Joao Couto for help with craniotomy surgeries. Nathaniel Miska contributed recordings from two aged animals. Henk van Steenberg, Bryant Jongkees, Steven Miletić, and Jacqueline Zadelaar provided valuable feedback on statistical analyses. Sander Nieuwenhuis offered helpful comments on the first draft. The IBL development team, especially Olivier Winter, provided crucial technical support with spike sorting and data release. Pranav Rai provided helpful code review. Peter Dayan and Kenneth Harris provided insightful suggestions throughout the project, and we thank the CoCoSys lab for discussions and feedback.

Author contributions

F.Z., A.K.C., and A.E.U. conceptualized the study. A.K. and A.E.U. contributed to the methodology for animal training, and A.E.U. contributed to the methodology for neural recordings. F.Z. and A.E.U. performed the formal analysis. F.Z., A.K., S.F., and A.E.U. contributed to data curation. The IBL provided software. F.Z. and A.E.U. contributed to visualization. F.Z. and A.E.U. wrote the original draft. F.Z., A.K., S.F., A.K.C., and A.E.U. reviewed and edited the manuscript. A.K.C. and A.E.U. supervised the study. F.Z., IBL, A.K.C., and A.E.U. acquired funding. Please see the Contribution Table (Supplementary Fig. 20) for additional details.

Funding

The IBL was supported by grants from the Wellcome Trust (216324) and the Simons Collaboration on the Global Brain. A.E.U. was supported by the German National Academy of Sciences Leopoldina, the International Brain Research Organization, and a Veni fellowship (VI.Veni.212.184) from the Netherlands Organisation for Scientific Research. F.Z. was supported by a PhD fellowship (202204910080) from the Chinese Scholarship Council. A.K.C. was supported by a grant from the Simons Collaboration on Plasticity and the Aging Brain.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-026-74227-1>.

Correspondence and requests for materials should be addressed to Fenying Zang or Anne E. Urai.

Peer review information *Nature Communications* thanks the anonymous reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2026

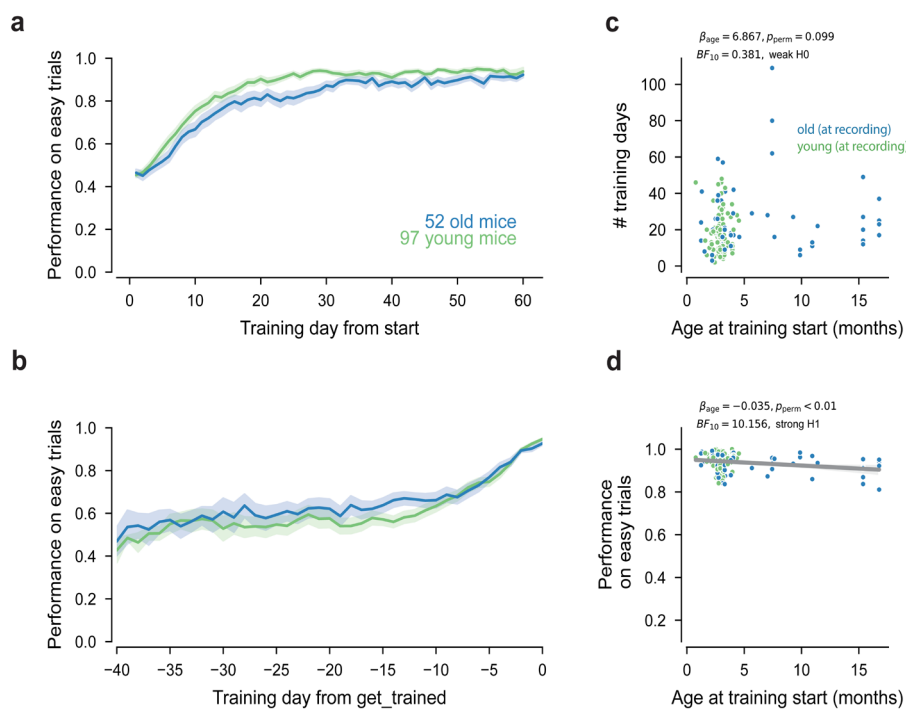
International Brain Laboratory

Larry Abbott⁴, Luigi Acerbi⁵, Valeria Aguillon-Rodriguez³, Mandana Ahmadi⁶, Jaweria Amjad⁶, Dora Angelaki⁷, Jaime Arlandis⁸, Zoe C. Ashwood⁹, Kush Banga¹⁰, Hailey Barrell¹¹, Hannah M. Bayer⁴, Brandon Benson¹², Julius Benson⁷, Jai Bhagat¹⁰, Dan Birman¹¹, Niccolò Bonacchi¹³, Kcenia Bougrova⁸, Julien Boussard⁴, Sebastian A. Bruijns¹⁴, E. Kelly Buchanan⁴, Robert Campbell¹⁵, Matteo Carandini¹⁶, Joana A. Catarino⁸, Fanny Cazettes⁸, Gaele A. Chapuis⁵, Anne K. Churchland⁴, Davide Crombie⁸, Yang Dan¹⁷, Felicia Davatolhagh², Peter Dayan¹⁴, Sophie Denève¹⁸, Eric EJ DeWitt⁸, Tatiana Engel⁹, Michele Fabbri⁴, Mayo Faulkner¹⁰, Robert Fetcho⁹, Ila Fiete¹⁹, Charles Findling⁵, Laura Freitas-Silva⁸, Surya Ganguli¹², Berk Gerçek⁵, Naureen Ghani¹⁵, Ivan Gordeliy¹⁸, Laura M. Haetzel⁹, Kenneth D. Harris¹⁰, Michael Hausser²⁰, Naoki Hiratani⁶, Sonja Hofer¹⁵, Fei Hu¹⁷, Felix Huber⁵, Julia M. Huntenburg¹⁴, Cole Hurwitz⁴, Anup Khanal², Christopher S. Krasniak³, Sanjukta Krishnagopal⁶, Michael Krumin¹⁰, Debottam Kundu¹⁴, Agnès Landemard¹⁶, Christopher Langdon⁹, Christopher Langfield⁴, Inês C. Laranjeira⁸, Peter Latham⁶, Petrina Lau²⁰, Hyun Dong Lee⁴, Ari Liu¹⁹, Zachary F. Mainen⁸, Amalia Makri-Cottingham²⁰, Hernando Martinez-Vergara¹⁵, Brenna McMannon⁹, Isaiah McRoberts⁷, Guido T. Meijer⁸, Maxwell Melin², Leenoy Meshulam²¹, Kim Miller¹¹, Nathaniel J. Miska¹⁵, Catalin Mitelut⁴, Zeinab Mohammadi⁹, Thomas Mrsic-Flogel¹⁵, Masayoshi Murakami²², Jean-Paul Noel²³, Kai Nylund¹¹, Farideh Oloomi¹⁴, Alejandro Pan Vazquez⁹, Liam Paninski⁴, Sabrina Perrenoud²⁰, Alberto Pezzotta⁶, Samuel Picard¹⁰, Jonathan W. Pillow⁹, Alexandre Pouget⁵, Carolina Quadrado¹⁰, Pranav Rai⁸, Georg Raiser⁸, Florian Rau⁸, Cyrille Rossant¹⁰, Noam Roth¹¹, Nicholas A. Roy⁹, Kamron Saniee⁴, Rylan Schaeffer¹⁹, Michael M. Schartner⁸, Yanliang Shi⁹, Karolina Z. Socha¹⁶, Cristian Soitu⁶, Nicholas A. Steinmetz¹¹, Karel Svoboda²⁴, Marsa Taheri², Charline Tessereau¹⁴, Matthew Tucker¹¹, Anne E. Urai¹✉, Erdem Varol⁴, Shuqi Wang⁴, Miles J. Wells¹⁰, Steven J. West¹⁵, Matthew R. Whiteway⁴, Charles Windolf⁴, Olivier Winter⁸, Ilana Witten⁹, Lauren E. Wool¹⁰, Zekai Xu⁶, Kenneth Yang¹¹, Yaxuan Yang¹⁵, Han Yu⁴, Anthony M. Zador³ & Yizi Zhang⁴

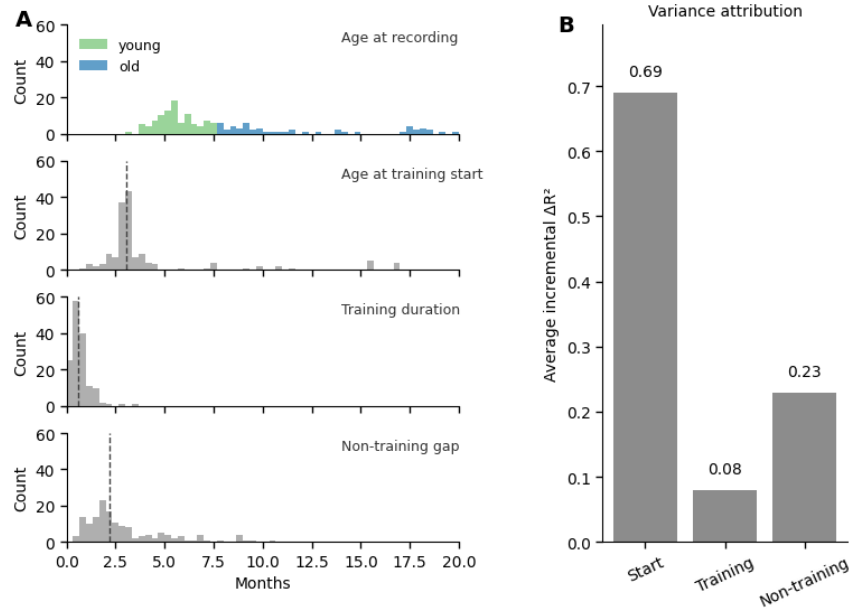
⁴Zuckerman Institute, Columbia University, New York, NY, USA. ⁵Department of Basic Neuroscience, University of Geneva, Geneva, Switzerland. ⁶Gatsby Computational Neuroscience Unit, University College London, London, UK. ⁷Center for Neural Science, New York University, New York, NY, USA. ⁸Cham-palimaud Center for the Unknown, Lisboa, Portugal. ⁹Princeton Neuroscience Institute, Princeton University, Princeton, NJ, USA. ¹⁰Institute of Neurology, University College London, London, UK. ¹¹Department of Biological Structure, University of Washington, Seattle, WA, USA. ¹²Department of Applied Physics, Stanford University, Stanford, CA, USA. ¹³William James Center for Research (WJCR), ISPA - Instituto Universitário, Lisbon, Portugal. ¹⁴Max Planck Institute for Biological Cybernetics, Tübingen, Germany. ¹⁵Sainsbury-Wellcome Centre, University College London, London, UK. ¹⁶Institute of Ophthalmology, University College London, London, UK. ¹⁷Department of Molecular and Cell Biology, University of California, Berkeley, CA, USA. ¹⁸Département D'études Cognitives, École Normale Supérieure, Paris, France. ¹⁹Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA, USA. ²⁰Wolfson Institute of Biomedical Research, University College London, London, UK. ²¹Center for Computational Neuroscience, University of Washington, Seattle, WA, USA. ²²Department of Physiology, University of Yamanashi, Kofu, Yamanashi, Japan. ²³Department of Neuroscience, University of Minnesota, Minneapolis, MN, USA. ²⁴The Allen Institute for Neural Dynamics, Seattle, WA, USA.

Supplementary Figures

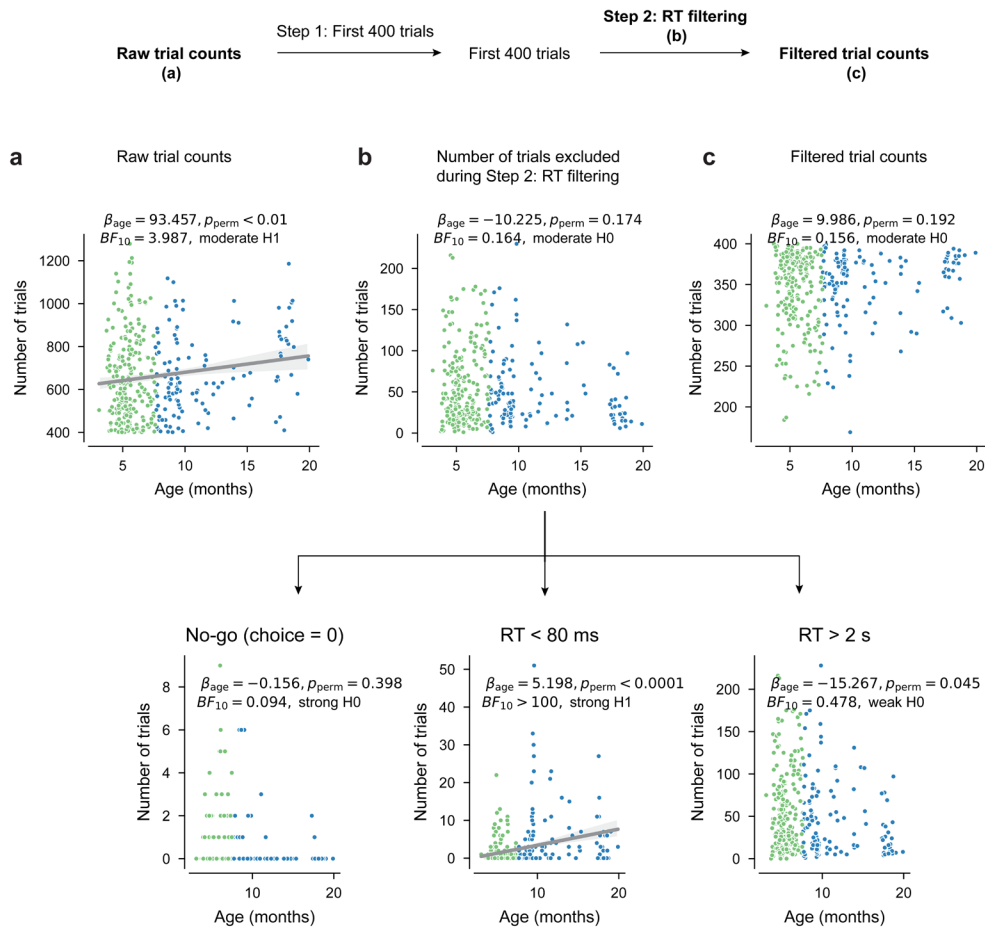
General plotting and statistical conventions for Supplementary Figures. Unless otherwise stated, age effects were estimated using general linear models with age as a continuous predictor. All statistical conclusions were based on Bayes Factors (BF_{10}); two-sided permutation p-values are unadjusted for multiple comparisons and are reported as complementary statistics (see Methods). Exact statistical values are shown in the figures or provided in the Source Data file. Regression lines are plotted only when the Bayes Factor indicates strong or moderate evidence for H1. Shaded areas and error bars indicate 95% confidence intervals.



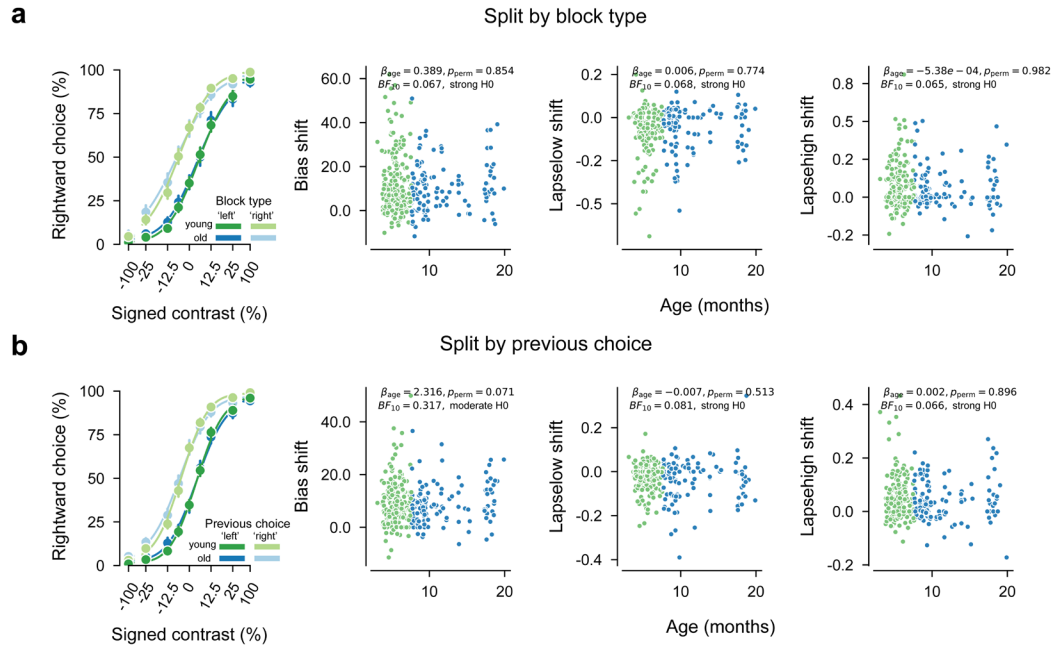
Supplementary Figure 1, related to Figure 1. Training progression in the behavioral task. All panels include the same mice unless otherwise stated: $n = 149$ mice (52 old mice and 97 young mice, age group defined at recording). **(a)** Learning curves from the start of training. Age group split was based on the age at recording, following the same convention as in other main figures. Solid lines show mean performance across mice; shaded areas indicate 95% confidence intervals. **(b)** Training curves aligned to the day animals reached the behavioral criterion ('get_trained', see the training protocol for definition¹). Solid lines show mean performance across mice; shaded areas indicate 95% confidence intervals. **(c)** Number of training days each mouse took to reach the behavioral criterion. Each point is one mouse. Conclusions are the same when we use a slightly different definition of training duration: number of days, sessions and trials required to reach the behavioural criterion or cumulative days, sessions and trials completed before the first neural recording. **(d)** Performance on easy trials on the day each mouse reached criterion. Each point is one mouse. The training criterion was identical for all mice. Observed differences reflect performance above this threshold within criterion-satisfying sessions on that day.



Supplementary Figure 2, related to Figure 1. Components contributing to variability in age at recording. All analyses used $n = 149$ mice. **(a)** Distributions (in months) of age at recording and the three components used to decompose age at recording: age at training start, training duration (number of training days), and the non-training gap between training start and the recording. Dashed lines indicate the median. **(b)** Order-invariant variance attribution using average incremental ΔR^2 across all predictor orderings in an OLS model.

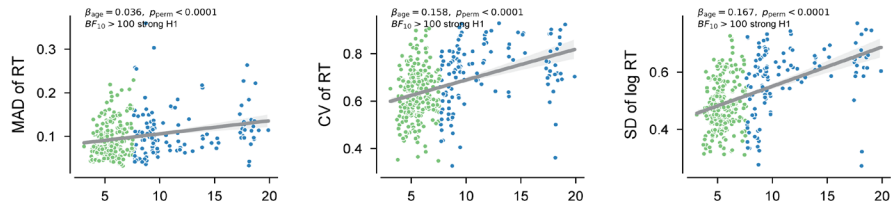


Supplementary Figure 3, related to Figure 1. Older mice performed slightly more trials in the standardized task. All panels use session-level data: $n = 367$ sessions from 149 mice. Each dot represents one session. **(a)** Number of trials as a function of age. **(b)** Number of trials excluded by RT filtering (no-go trials, $RT < 80\text{ms}$ trials, and $RT > 2\text{s}$ trials). **(c)** Number of trials included as a function of mouse age.

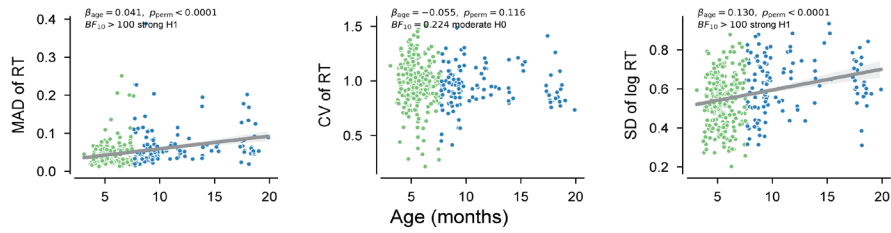


Supplementary Figure 4, related to Figure 1. No age-related differences in block- or history-dependent choice bias. All panels use session-level psychometric fits: $n = 367$ sessions from 149 mice. For psychometric curves, error bars indicate 95% confidence intervals. Each dot in the scatter plots represents one session. **(a)** Trials were split by block type (see Methods), and shifts in psychometric parameters (bias and lapse rate) were quantified between the two block conditions. Left, psychometric curves show mean rightward choice probability as a function of signed contrast, split by age group and block type. No age-related effects were observed for bias shift, low-lapse shift, or high-lapse shift. **(b)** Trials were split based on the previous choice (clockwise and counterclockwise), and shifts in psychometric parameters (bias and lapse rate) were quantified between the two types of previous choices. Left, psychometric curves show mean rightward choice probability as a function of signed contrast, split by age group and previous choice. Again, no age-related effects were observed for bias shift, low-lapse shift, or high-lapse shift.

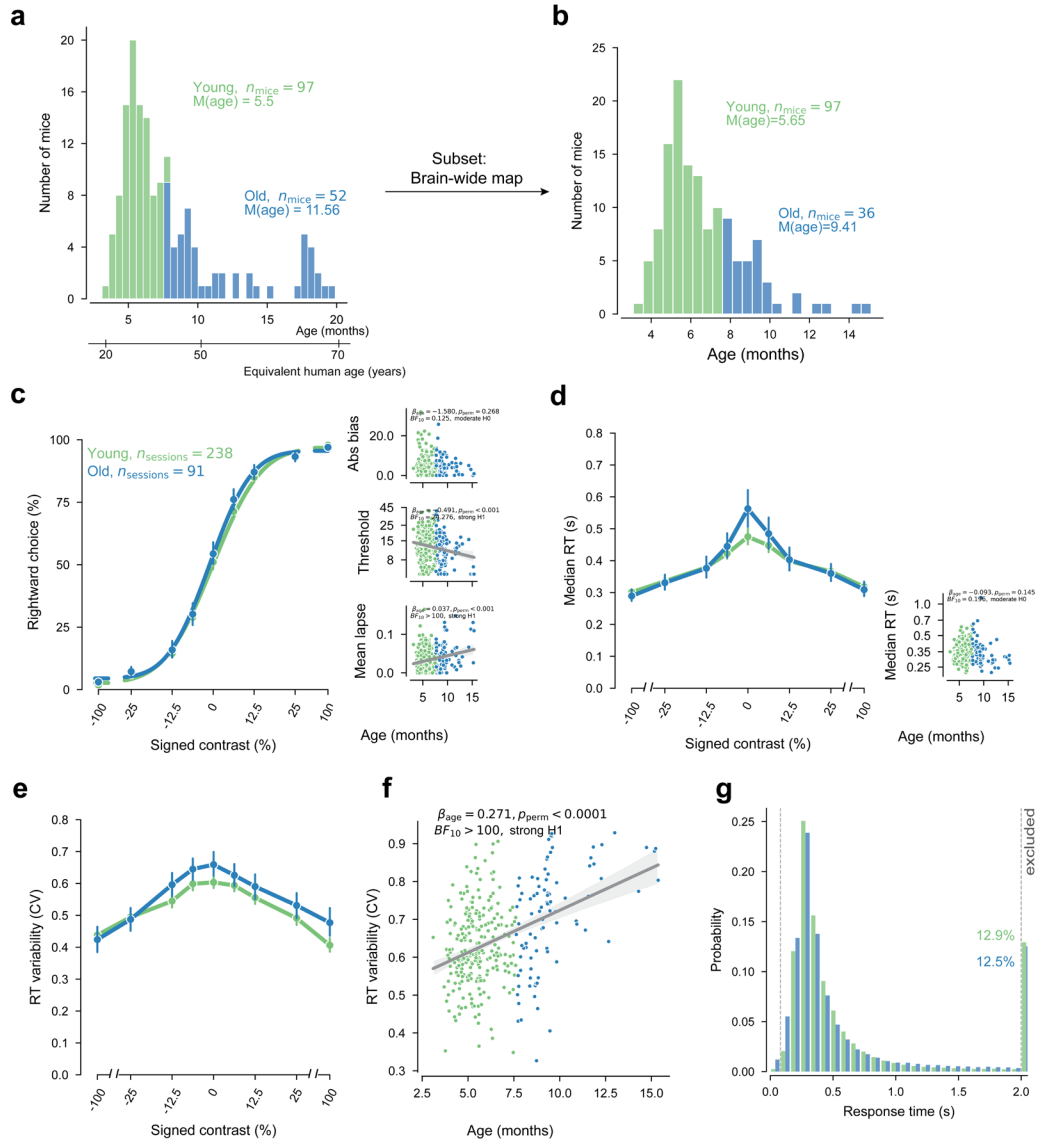
a RT = response time - stimulus onset



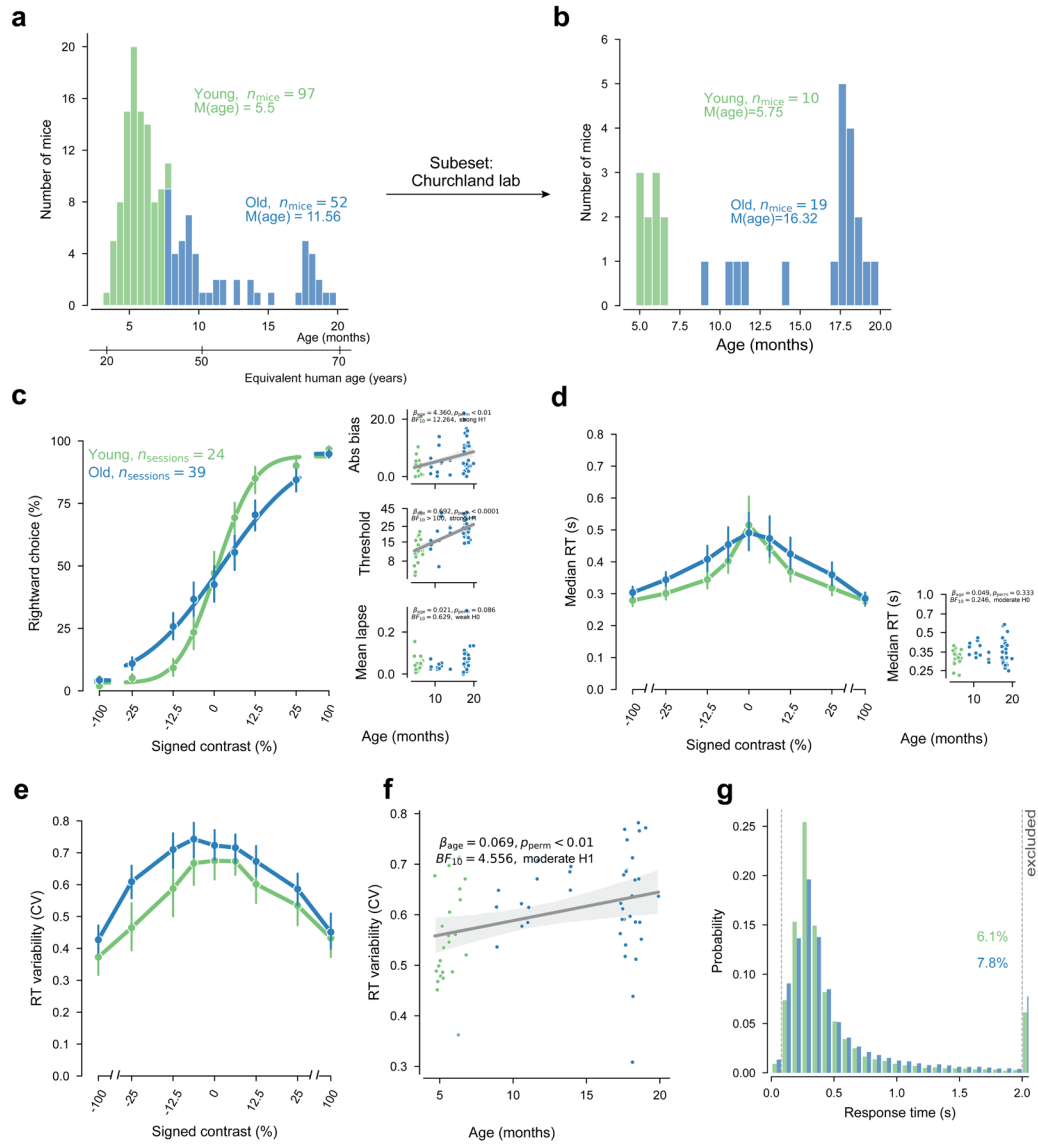
b RT = first movement time - stimulus onset



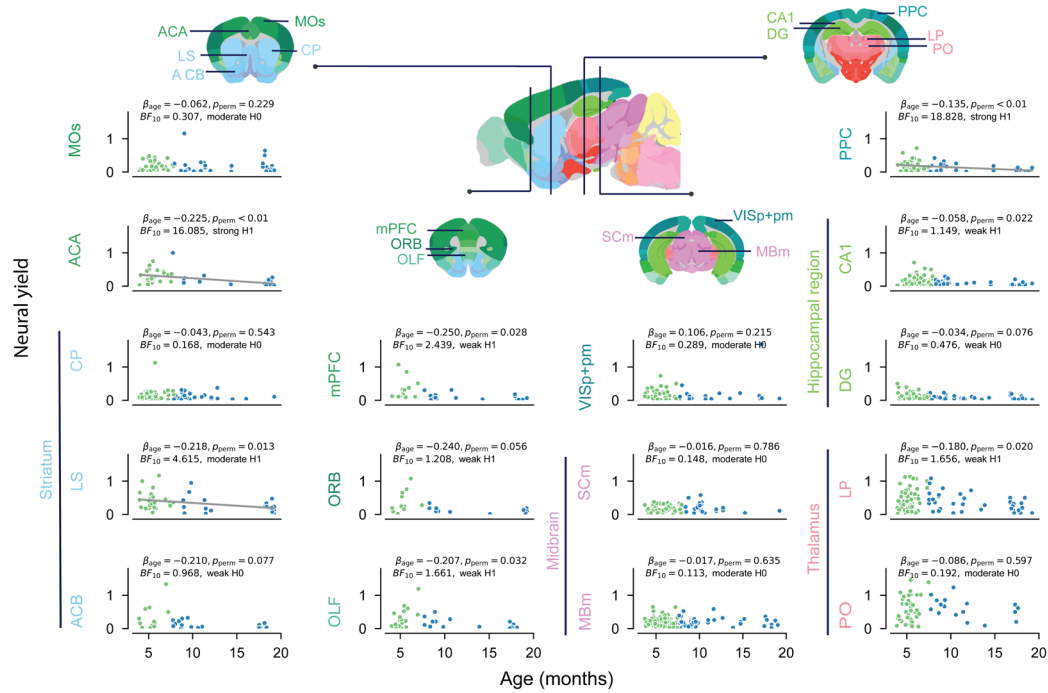
Supplementary Figure 5, related to Figure 1. RT variability measures across two RT definitions and their relationship with age. All panels use session-level data: $n = 367$ sessions from 149 mice. Each dot represents one session. Solid lines indicate fitted age effects where plotted; shaded areas indicate 95% confidence intervals. **(a)** RT was defined as the interval between stimulus onset and response time. We examined alternative measures of RT variability that capture different aspects of dispersion. Specifically, we used the median absolute deviation (MAD), a robust metric less sensitive to skew, the coefficient of variation (CV), and the standard deviation (SD) of log-transformed RTs, which stabilizes variance in positively skewed distributions. Scatterplots show the relationship between age and three variability measures: MAD of RT (left), CV of RT (middle), and SD of log-transformed RT (right). **(b)** Same as (a), but for movement initiation time, defined as the interval between stimulus onset and the first movement².



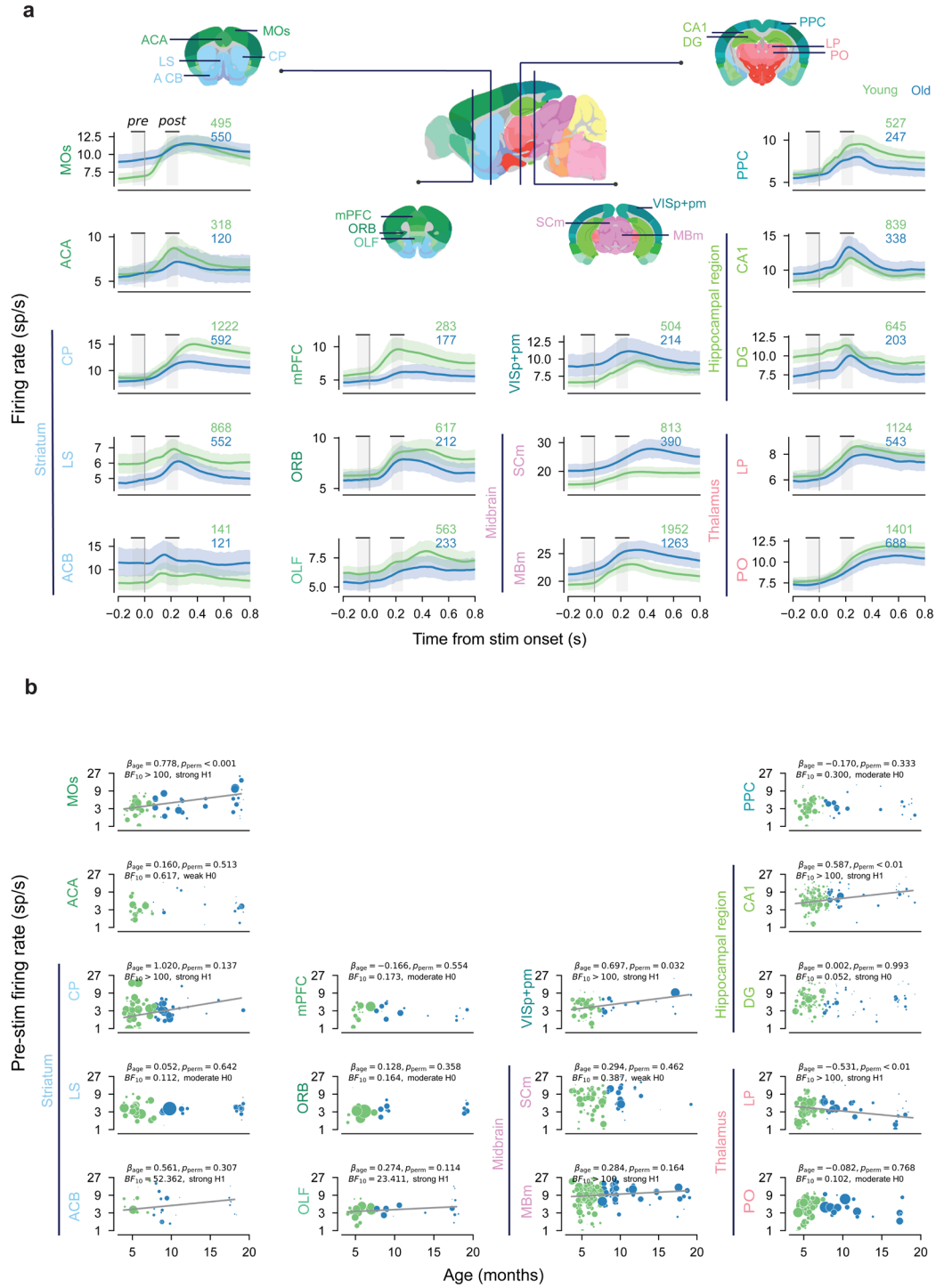
Supplementary Figure 6, related to Figure 1. Replication of Figure 1 using only mice from the original IBL “brain-wide map” dataset. Panels, analysis definitions, and plotting/statistical conventions are the same as in main Figure 1, except that analyses were restricted to mice from the original IBL “brain-wide map” dataset²: $n = 133$ mice, 329 sessions.



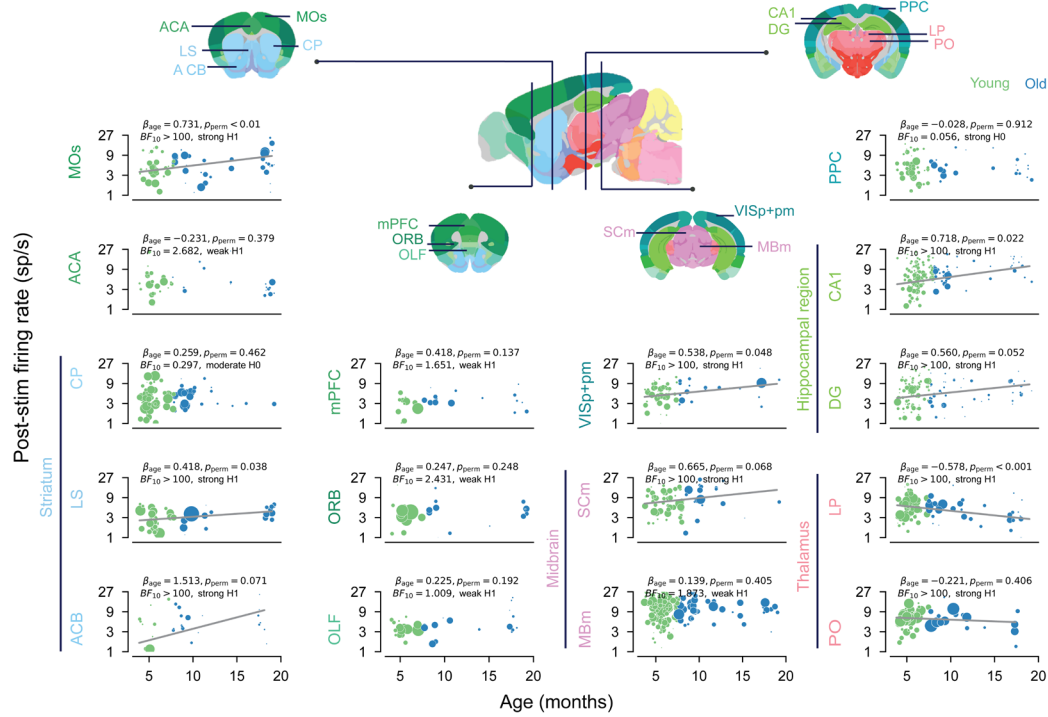
Supplementary Figure 7, related to Figure 1. Replication of Figure 1 using only sessions recorded in the Churchland lab. Panels, analysis definitions, and plotting/statistical conventions are the same as in main Figure 1, except that analyses were restricted to sessions recorded in the Churchland lab: $n = 29$ mice and 63 sessions.



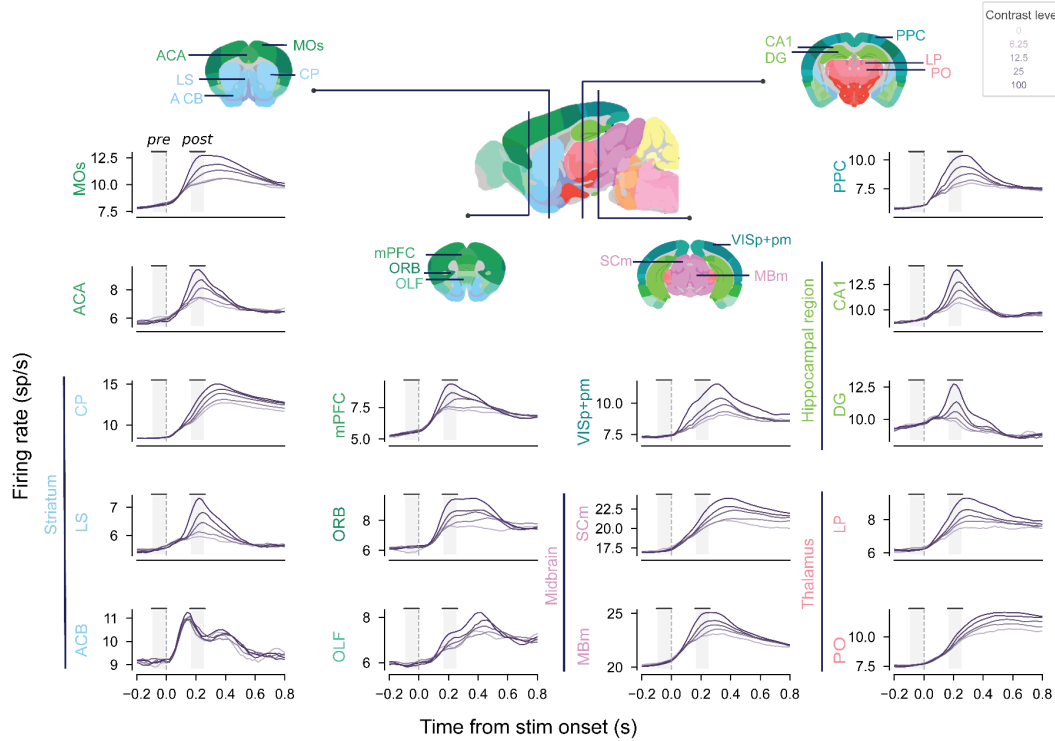
Supplementary Figure 8, related to Figure 2. Neural yield slightly decreases with age. Scatterplots show neural yield in each ROI as a function of age. Each dot corresponds to one recording insertion. Statistical analyses for neural yield were performed at the insertion level (yield is inherently defined per insertion). Neural yield is defined as the number of good neurons per good channel (along the length of the Neuropixels probe) within a given region.



Supplementary Figure 9, related to Figure 3. Regional specificity of firing rate time courses and pre-stimulus firing rates. (a) Time courses of mean firing rates in different brain regions, aligned to stimulus onset. Lines show mean firing rates across neurons; shaded areas indicate 95% confidence intervals. ROI-specific n values (number of neurons) are indicated in each panel. **(b)** The relationship between pre-stimulus (−100 to 0 ms relative to stimulus onset) firing rate and mouse age. For visualization, each dot represents the mean across neurons from one insertion, and dot size represents the number of neurons recorded in that insertion. ROI-specific n values are the same as in (a) and indicate the number of neurons included in the statistical analysis.

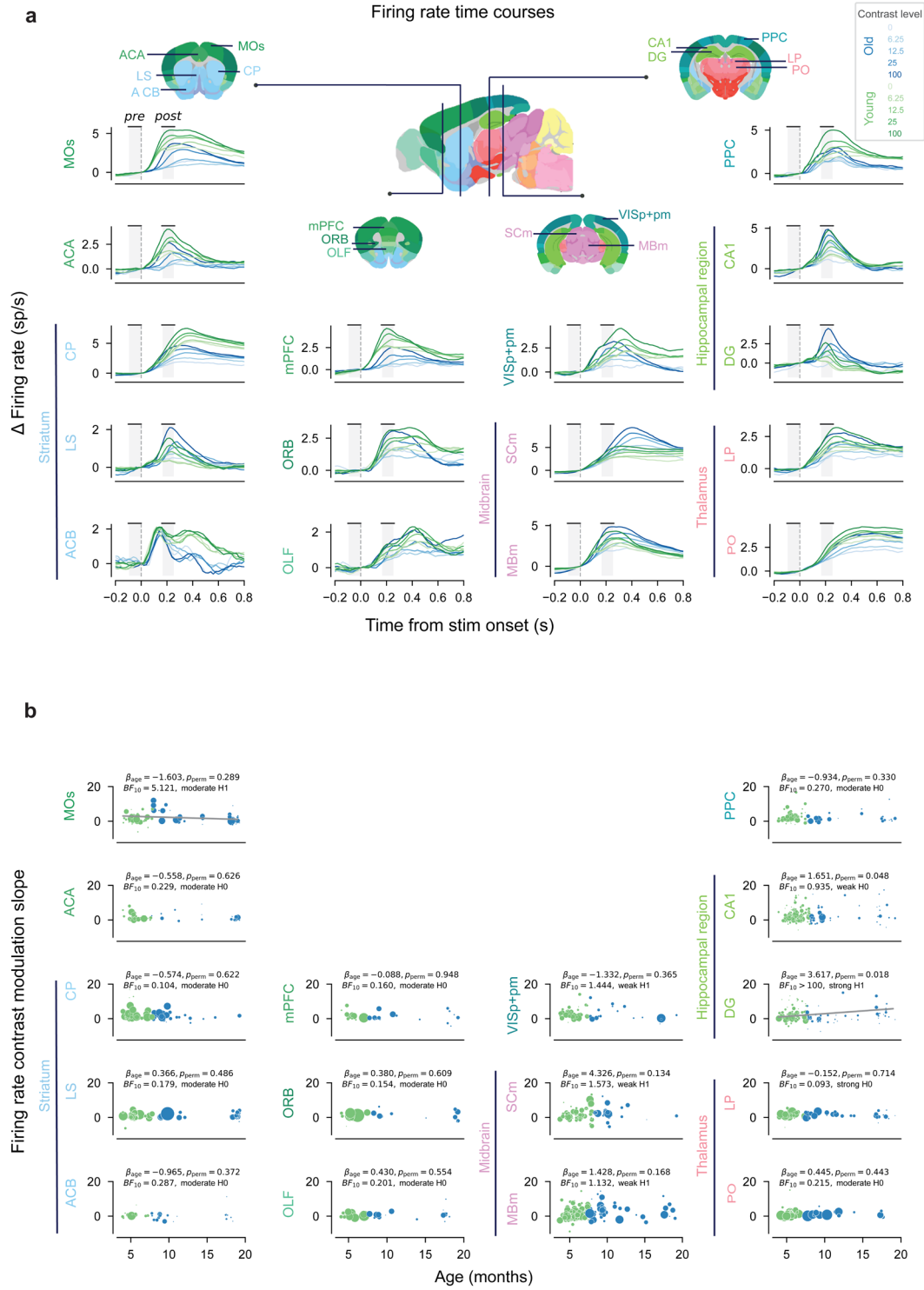


Supplementary Figure 10, related to Figure 3. Regional specificity of post-stimulus firing rates. The relationship between post-stimulus firing rate (160 to 260 ms relative to stimulus onset) and mouse age. Visualization, sample-size, plotting, and statistical conventions are the same as in Supplementary Figure 9b.



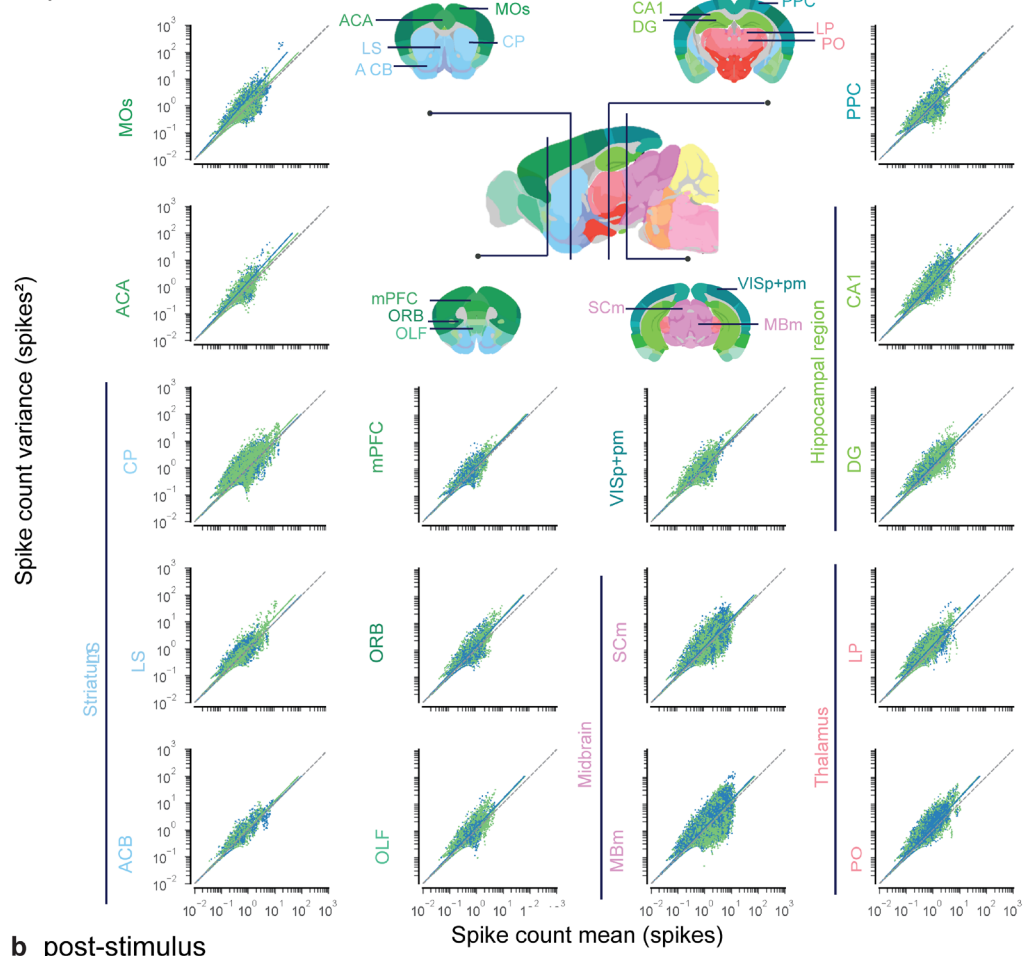
Supplementary Figure 11, related to Figure 3. Regional specificity in contrast modulation of firing rate. Time courses of the average firing rate in each ROI across different stimulus contrast levels, combining data from both age groups. Each line shows the mean firing rate across neurons for one contrast level. ROI-specific sample

sizes are the same as in Supplementary Figure 10a and indicate the number of neurons. Grey shaded regions indicate the pre- and post-stimulus windows used for summary analyses.

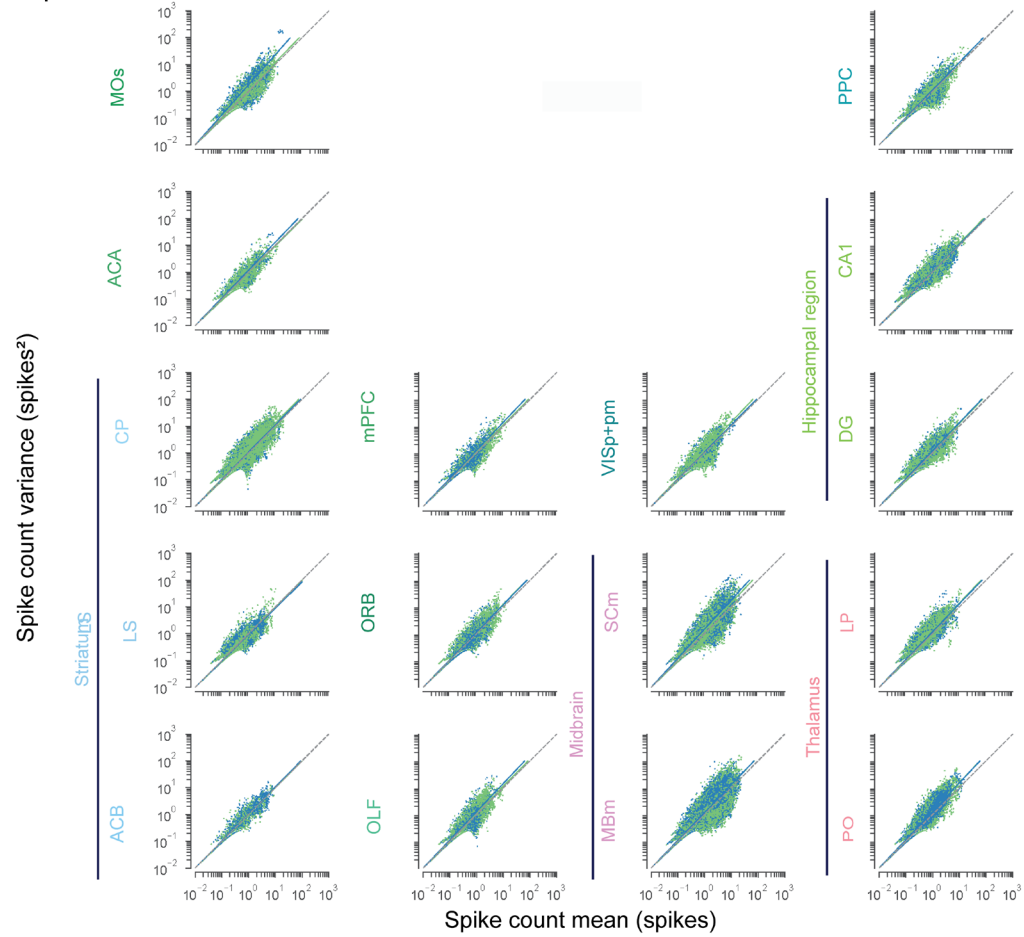


Supplementary Figure 12, related to Figure 3. Regional specificity of age-related differences in contrast modulation of firing rates. (a) Baseline-corrected mean firing rate time courses for different stimulus contrast levels in each ROI, aligned to stimulus onset. Saturation represents different contrast levels. Lines show mean firing rates across neurons. **(b)** The relationship between contrast modulation slope and mouse age. Each dot represents one insertion. For visualization, each dot represents the mean contrast modulation slope across neurons from one insertion, and dot size represents the number of neurons recorded in that insertion. ROI-specific sample sizes are the same as in Supplementary Figure 10a and indicate the number of neurons.

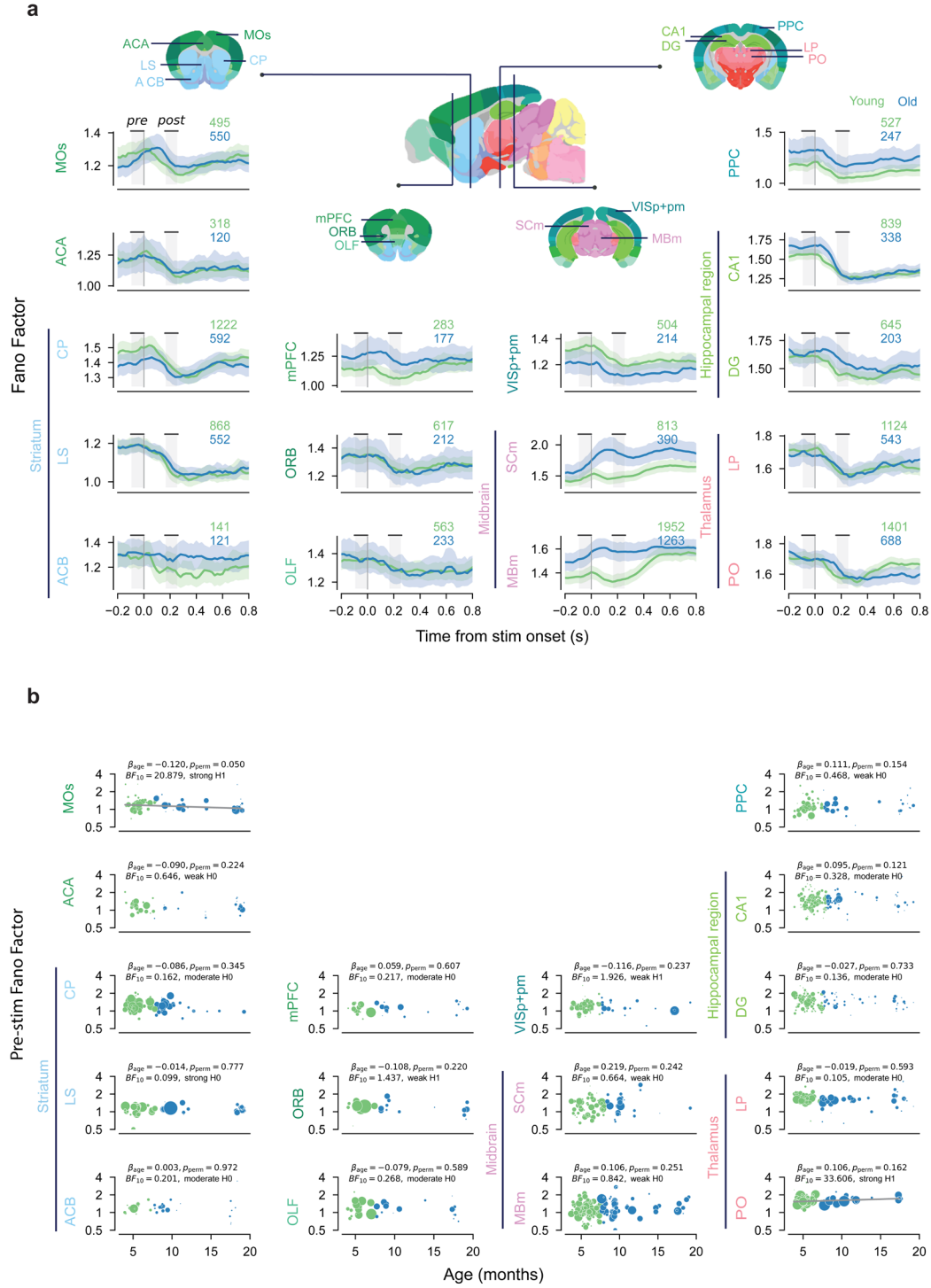
a pre-stimulus

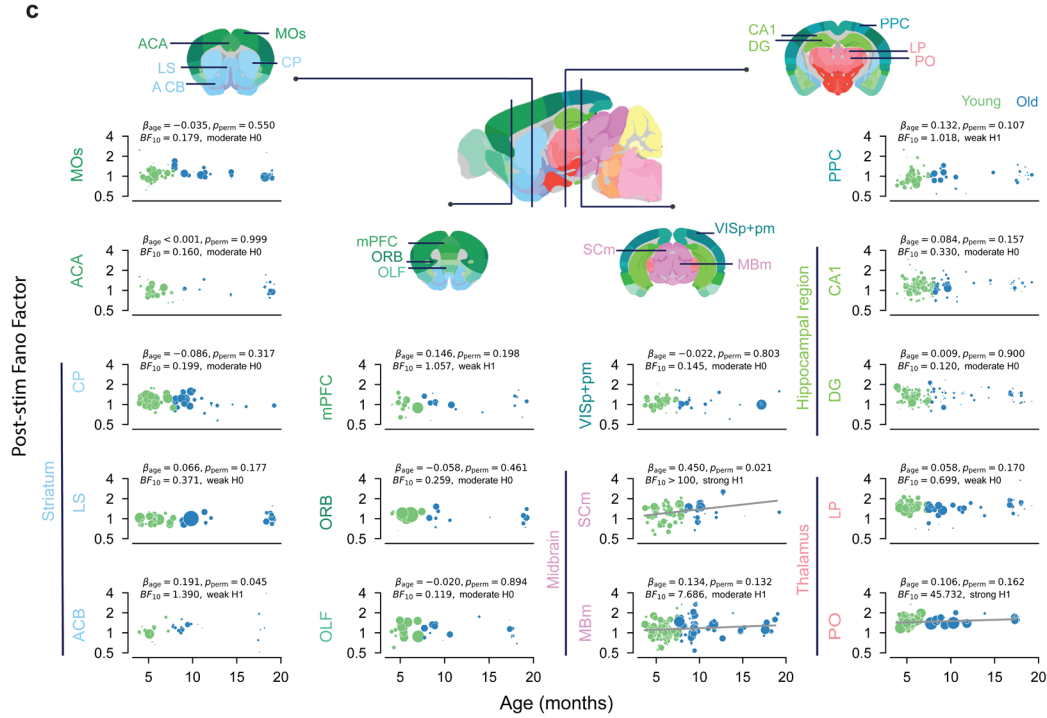


b post-stimulus

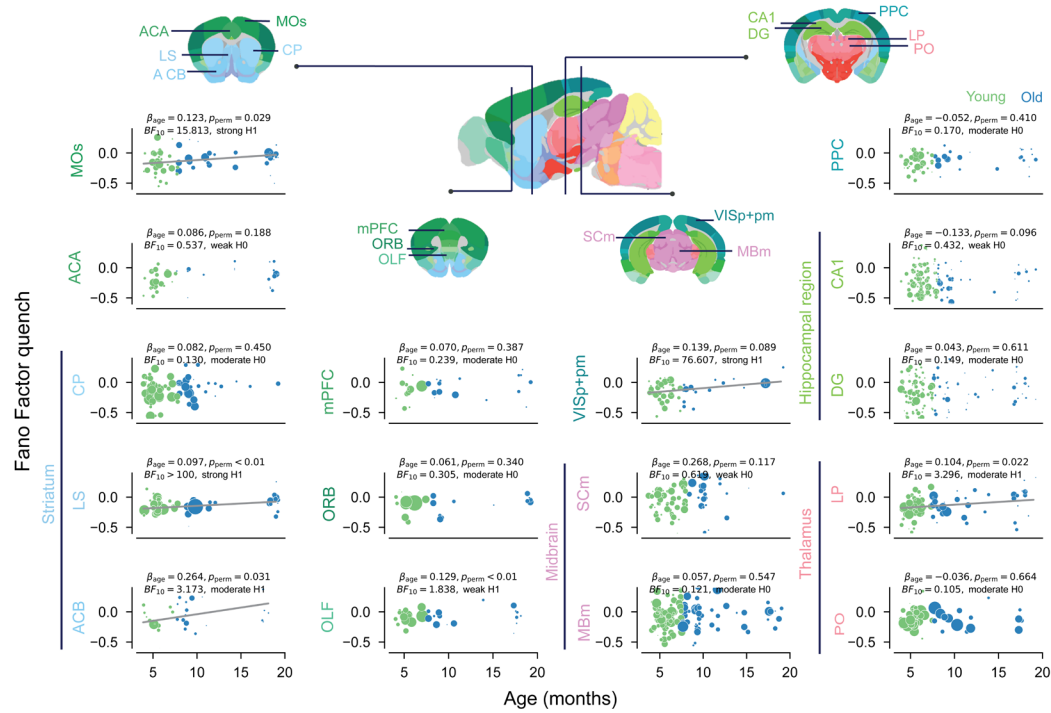


Supplementary Figure 13, related to Figure 4. The relationship between spike count variance and spike count mean. For each neuron and each stimulus level, we plotted its spike count variance (spikes²) against its spike count mean (spikes) across trials, separately for (a) the pre-stimulus time window (-100ms to 0ms) and (b) the post-stimulus time window (160ms to 260ms). Colored lines represent linear fits for each age group (blue = old, green = young), with the slope corresponding to an estimate of the Fano Factor. We observed curvilinear patterns in the lower-left part of many panels, resembling patterns found in Figure 2 of a previous study³. The origin of these patterns remains unclear and warrants further investigation.

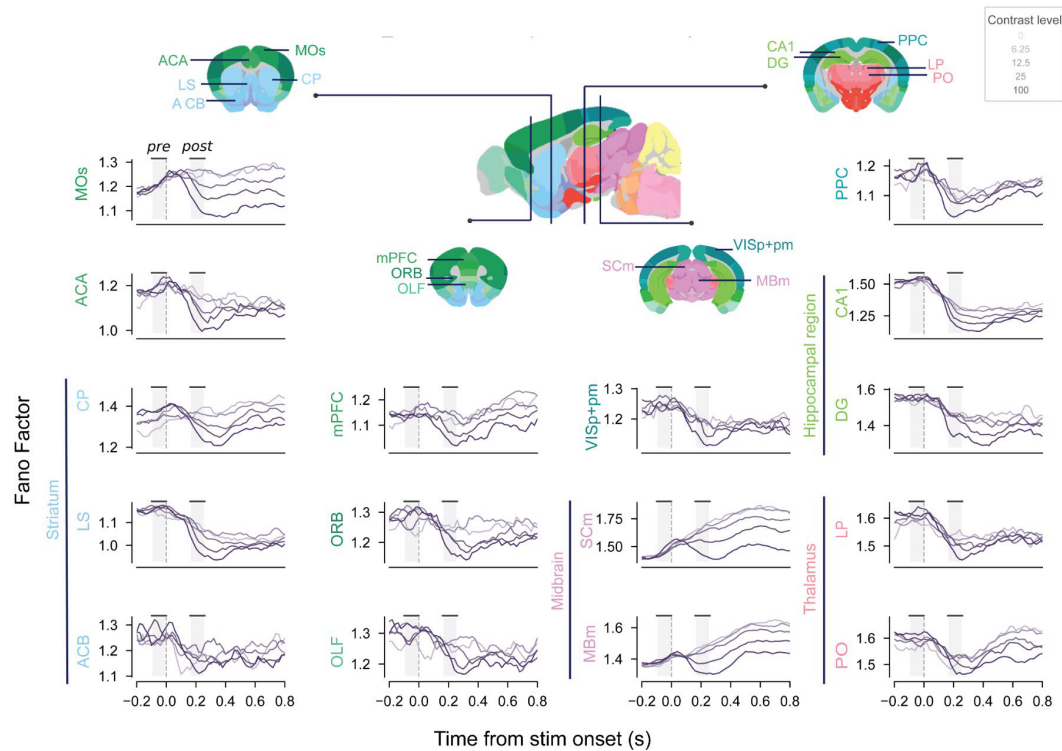




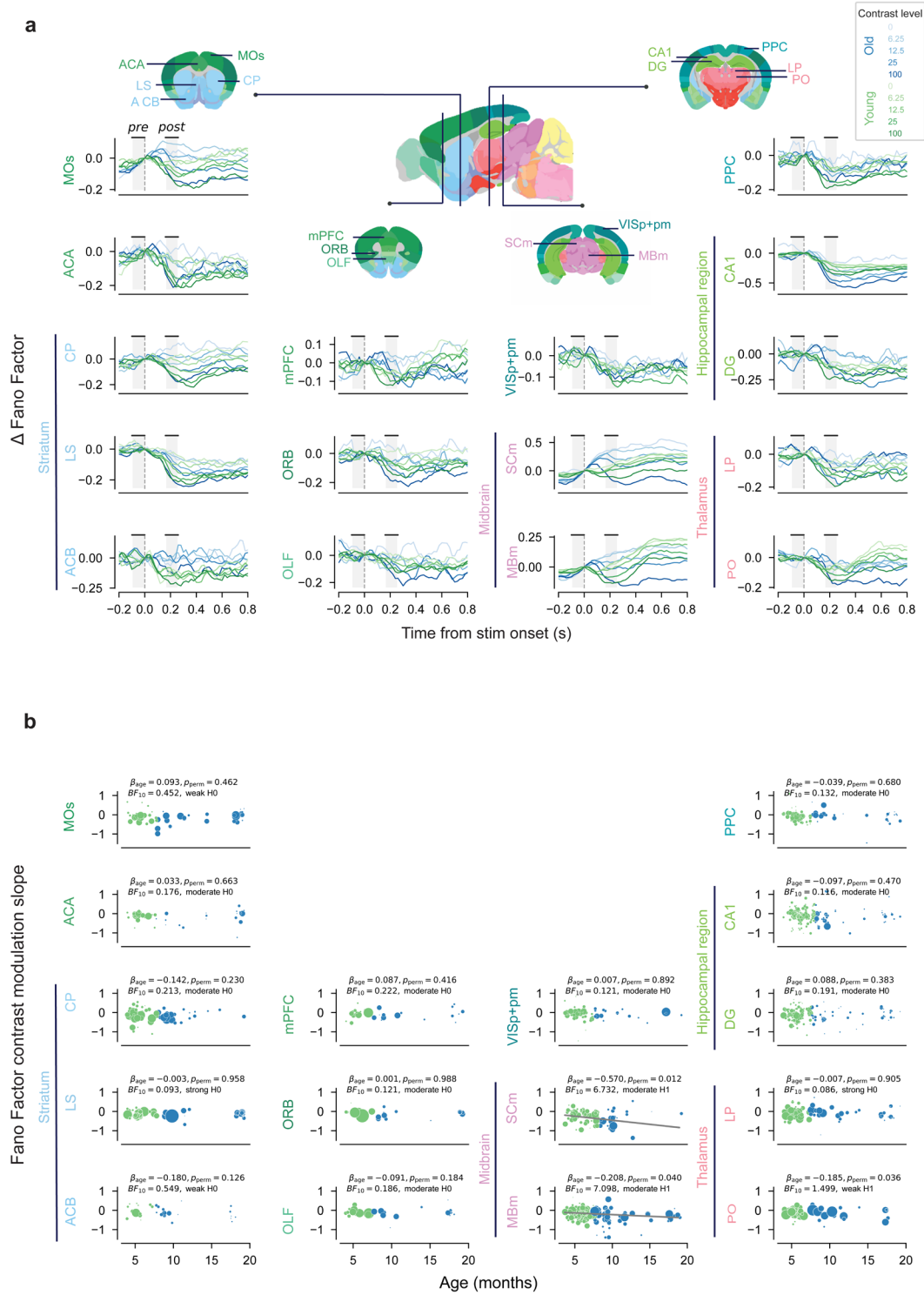
Supplementary Figure 14, related to Figure 4. Regional specificity of mean-subtracted Fano Factors. (a) Time courses of mean-subtracted Fano Factor in each ROI, aligned to stimulus onset. Thick lines show the mean Fano Factor across all neurons within each age group, with shaded areas representing the 95% confidence intervals. Sliding window width: 0.1 s; step size: 0.02 s. Numbers indicate the number of neurons in each group. **(b)** The relationship between pre-stimulus mean-subtracted Fano Factor and age. For visualization, each dot represents the mean across neurons from one insertion, and dot size represents the number of neurons recorded in that insertion. ROI-specific n values are the same as in (a) and indicate the number of neurons included in the analysis. **(c)** The relationship between post-stimulus mean-subtracted Fano Factor and age. Visualization, sample-size definition, plotting, and statistical conventions are the same as in (b).



Supplementary Figure 15, related to Figure 4. Regional specificity of mean-subtracted Fano Factor quenching. The relationship between the mean-subtracted Fano Factor quench index (calculated as post-stimulus Fano Factor minus pre-stimulus Fano Factor for each neuron) and mouse age. For visualization, each dot represents the mean quench index across neurons from one insertion, and dot size represents the number of neurons recorded in that insertion. ROI-specific n values are the same as in Supplementary Figure 14a and indicate the number of neurons included in the analysis.

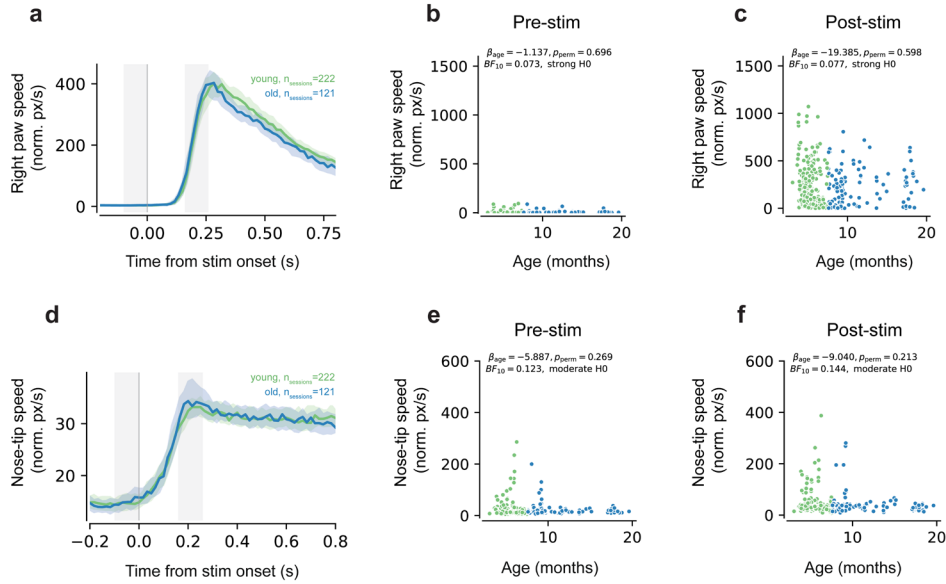


Supplementary Figure 16, related to Figure 4. Regional specificity in contrast modulation of the Fano Factor. Time courses of the mean Fano Factor for different stimulus contrast levels and ROIs, combining data from both age groups. Each line shows the mean Fano Factor across neurons for one contrast level. ROI-specific n values are the same as in Supplementary Figure 14a and indicate the number of neurons. Grey shaded regions indicate the pre- and post-stimulus windows used for summary analyses.

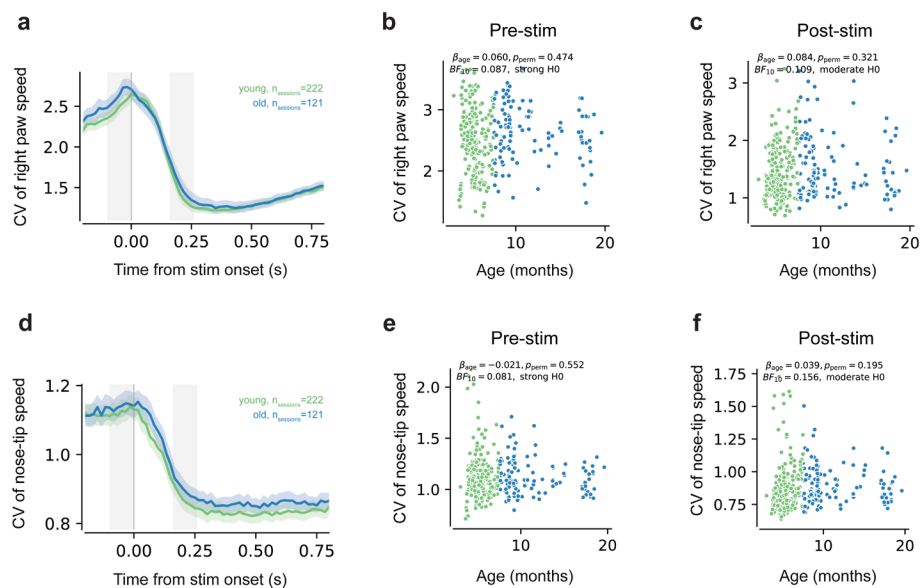


Supplementary Figure 17, related to Figure 4. Regional specificity in age-related differences in contrast modulation of the Fano Factor. (a) Baseline-corrected Fano Factor time courses, across neurons, for different stimulus contrast levels in each ROI, aligned to stimulus onset. Colors represent the age groups (blue = old, green = young). Saturation represents different contrast levels. Sliding window width: 0.1 s; step size: 0.02 s. Grey areas

indicate the pre-stimulus (-100 to 0 ms) and post-stimulus (160 to 260 ms) time windows used for summary analyses. **(b)** The relationship between Fano Factor contrast-modulation slope and mouse age. Contrast modulation slopes were computed for each neuron by fitting a linear relationship between the change in baseline-corrected Fano Factor and the different stimulus contrast levels. For visualization, each dot represents the mean contrast-modulation slope across neurons from one insertion. ROI-specific n values are the same as in Supplementary Figure 14a and indicate the number of neurons used in the statistical analysis.

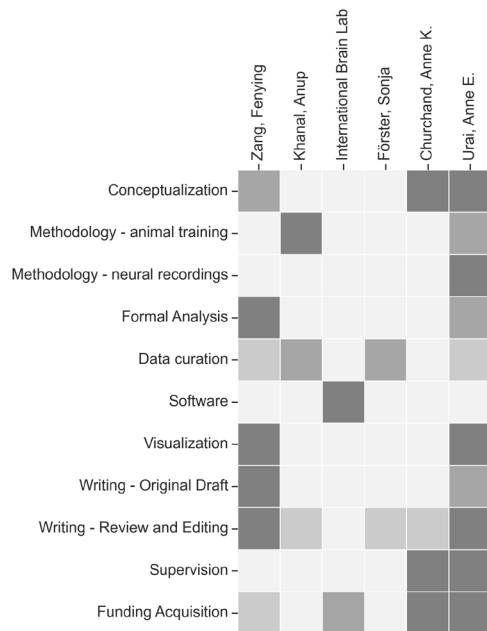


Supplementary Figure 18, related to Figure 4. Right-camera DLC movement time courses and session-level summary metrics. All panels use session-level data: n = 343 sessions (222 young and 121 old sessions). **(a)** Right paw speed time course aligned to stimulus onset (t = 0), summarized across sessions (median ± 95% confidence interval). Gray shaded regions indicate the pre-stimulus (-100 to 0 ms) and post-stimulus (160 to 260 ms) windows used to compute session-level movement summaries. **(b–c)** Session-wise median right paw speed within the pre-stimulus window (b) and post-stimulus window (c), plotted against mouse age; each dot corresponds to one session. **(d)** Nose-tip speed time course aligned to stimulus onset, summarized across sessions (median ± 95% confidence interval), with the same pre/post windows indicated. **(e–f)** Session-wise median nose-tip speed within the pre-stimulus (e) and post-stimulus (f) windows versus age; each dot corresponds to one session.



Supplementary Figure 19, related to Figure 4. Right-camera DLC movement variability time courses and session-level summary metrics. All panels use session-level data: n = 343 sessions (222 young and 121 old

sessions). **(a)** Time course of right paw speed variability, quantified as the coefficient of variation (CV) across trials and aligned to stimulus onset ($t = 0$), summarized across sessions (mean $\pm$ 95% confidence interval). Gray shaded regions indicate the pre-stimulus (-100 to 0 ms) and post-stimulus (160 to 260 ms) windows used to compute session-level movement variability summaries. **(b–c)** Session-wise right paw speed CV within the pre-stimulus window (b) and post-stimulus window (c), plotted against mouse age; each dot corresponds to one session. **(d)** Time course of nose-tip speed variability, quantified as CV and aligned to stimulus onset, summarized across sessions (mean $\pm$ 95% confidence interval), with the same pre/post windows indicated. **(e–f)** Session-wise nose-tip speed CV within the pre-stimulus (e) and post-stimulus (f) windows versus age; each dot corresponds to one session.



Supplementary Figure 20. CRediT matrix of author contributions.

Supplementary Tables

metric	β_{age}	BF_evidence_age
rt_median	-0.014	<0.1, strong H0
rt_CV	+0.158	>100, strong H1
pre_fr	+0.216	>100, strong H1
post_fr	+0.229	>100, strong H1
fr_delta_modulation	+0.373	0.28, moderate H0
pre_ff	-0.001	<0.1, strong H0
post_ff	+0.056	>100, strong H1
ff_quench	+0.070	>100, strong H1
ff_quench_modulation	-0.066	1.42, weak H1

Supplementary Table 1. Age effects from the original models (in the main analyses). For each of the key behavioral and neural metrics, the table reports β and BF_{10} for age.

metric	β_{age}	BF_evidence_age	β_{training}	BF_evidence_training
rt_median	-0.025	0.22, moderate H0	+0.483	0.49, weak H0
rt_CV	+0.140	>100, strong H1	+0.276	0.37, weak H0
pre_fr	+0.211	>100, strong H1	+0.311	<0.1, strong H0
post_fr	+0.202	>100, strong H1	+1.200	>100, strong H1
fr_delta_modulation	+0.394	0.42, weak H0	-0.946	<0.1, strong H0
pre_ff	-0.012	<0.1, strong H0	+0.449	7.81, moderate H1
post_ff	+0.045	13.6, strong H1	+0.424	8.87, moderate H1
ff_quench	+0.077	>100, strong H1	-0.297	0.25, moderate H0
ff_quench_modulation	-0.072	3.16, moderate H1	+0.310	<0.1, strong H0

Supplementary Table 2. Controls for training duration: extended models with age and training duration.

For each behavioural and neural metric, we report results on an extended model including both age at recording and training duration. In the extended models, the age term reflects the partial age effect controlling for training duration; the training term reflects any additional contribution of training duration beyond age.

metric	β_{age}	BF_evidence_age	β_{paw}	BF_evidence_paw	β_{nose}	BF_evidence_nose
pre_fr	+0.201	>100, strong H1	+0.016	0.37, weak H0	-0.028	>100, strong H1
post_fr	+0.211	>100, strong H1	+0.006	<0.1, strong H0	-0.027	>100, strong H1
fr_delta_modulation	+0.437	1.12, weak H1	-0.086	<0.1, strong H0	-0.119	3.7, moderate H1
pre_ff	+0.017	<0.1, strong H0	-0.015	6.14, moderate H1	+0.006	0.11, moderate H0
post_ff	+0.062	>100, strong H1	-0.014	7.58, moderate H1	-0.003	<0.1, strong H0
ff_quench	+0.052	11.7, strong H1	-0.002	<0.1, strong H0	-0.005	>100, strong H1
ff_quench_modulation	-0.076	5.81, moderate H1	-0.012	<0.1, strong H0	+0.014	0.45, weak H0

Supplementary Table 3. Additional controls using video-based movement (right camera): Extended model summary (Age + Paw + Nose-tip). In sessions with QC-passing right-camera recordings ($n = 343$), we extracted right-paw speed (task-relevant) and nose-tip speed (less task-specific) in pre/post windows and added them as covariates in the extended models. The table reports β and BF_{10} for age, paw, and nose-tip terms.

metric	β_{age}	BF_evidence_age	$\beta_{\text{paw_cv}}$	BF10_paw_cv	$\beta_{\text{nose_cv}}$	BF10_nose_cv
pre_fr	+0.200	>100, strong H1	+0.020	3.07, moderate H1	-0.017	1.11, weak H1
post_fr	+0.207	>100, strong H1	+0.026	13.3, strong H1	+0.023	4.83, moderate H1
fr_delta_modulation	+0.552	6.59, moderate H1	+0.130	0.77, weak H0	-0.017	<0.1, strong H0
pre_ff	+0.014	<0.1, strong H0	+0.025	>100, strong H1	-0.009	0.28, moderate H0
post_ff	+0.066	>100, strong H1	+0.023	>100, strong H1	-0.010	0.42, weak H0
ff_quench	+0.046	2.89, weak H1	+0.006	>100, strong H1	-0.007	1.55, weak H1
ff_quench_modulation	-0.081	10.8, strong H1	-0.000	<0.1, strong H0	+0.007	<0.1, strong H0

Supplementary Table 4. Additional controls using video-based movement variability (right camera): Extended model summary (Age + Paw CV+ Nose-tip CV). In sessions with QC-passing right-camera recordings ($n = 343$), we extracted right-paw and nose-tip speed variability, quantified as CV within the relevant pre- and/or post-stimulus windows, and added these as covariates in the extended models. The table reports β and BF_{10} for age, paw CV, and nose-tip CV terms. Although these covariates explained additional variance in several cases, adding these movement-variability measures as covariates to the extended models left the main age-related pattern broadly unchanged. For firing-rate contrast modulation, the evidence shifted from moderate evidence for H0 to weak evidence for H1 after including movement variability covariates, yielding weak evidence for an age effect in the adjusted model. For two additional metrics, the qualitative conclusion remained the same, but the strength of evidence changed: for Fano factor quenching, evidence for H1 was attenuated from strong to weak, whereas for Fano factor contrast modulation, evidence for H1 increased from weak to strong. All other metrics retained the same qualitative conclusion.

References

1. International Brain Laboratory. Behavior: Appendix 2: IBL protocol for mice training. Preprint at <https://doi.org/10.6084/m9.figshare.11634729.v3> (2020).
2. International Brain Laboratory *et al.* A brain-wide map of neural activity during complex behaviour. *Nature* **645**, 177–191 (2025).
3. Goris, R. L. T., Movshon, J. A. & Simoncelli, E. P. Partitioning neuronal variability. *Nat. Neurosci.* **17**, 858–865 (2014).