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

Fenying Zang[1†], Anup Khanal[2], Sonja Förster[1], International Brain Laboratory*, Anne K. Churchland[2], Anne E. Urai[1†]

[1] *Leiden University, The Netherlands* [2] *University of California Los Angeles, USA*

*A list of authors and their affiliations appears at the end of the paper.

[†] Correspondence: f.zang@fsw.leidenuniv.nl; a.e.urai@fsw.leidenuniv.nl

Abstract

Age-related cognitive decline in learning and decision-making may arise from increased variability of neural responses. Here, we investigated how ageing affects behavioral 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, 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.

Keywords: ageing; mouse; perceptual decision-making; neurophysiology; neural variability

Introduction

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 (Cremer & Zeef, 1987; Crossman & Szafran, 1956; Salthouse & Lichty, 1985; Welford, 1981). Behavioral studies at the time supported this idea, showing that the absolute threshold for detecting stimuli increases with age (Welford, 1981) and that adding random stimulus noise (simulating internal neural noise) could mimic age-related differences in task performance, particularly in visual tasks (Cremer & Zeef, 1987). 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 function (Li & Rieckmann, 2014). Such changes reduce the precision of

neural representations, implying that 'neural dedifferentiation' lies at the basis of cognitive changes in healthy ageing (Koen & Rugg, 2019).

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 regions (Boylan et al., 2021), others reported that many brain regions become less variable in older people (Garrett et al., 2010, 2011; Nomi et al., 2017). The behavioral 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 (Garrett et al., 2011), while others report increased BOLD variability in older adults, negatively affecting decision-making (Samanez-Larkin et al., 2010). Older adults' EEG activity consistently shows an increase in neural variability, as measured by metrics such as P300 latency variability (Kügler et al., 1993), weighted permutation entropy (Waschke et al., 2017), and particularly the slope of 1/f power spectra (Ribeiro & Castelo-Branco, 2022; Tran et al., 2020; Voytek et al., 2015). The latter measure, characterized by the aperiodic component of 1/f power spectral density (Voytek et al., 2015), has been interpreted as reflecting an increase in baseline neural 'noise', and correlates with age-related impairments in cognitive functions such as visual working memory (Voytek et al., 2015) and visual processing (Tran et al., 2020). However, flatter slopes of EEG power spectra were recently shown to arise from a cardiac, rather than neural, source that changes with age (Schmidt et al., 2024). 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 data (Schmidt et al., 2024).

Observations from single-neuron measurements in cortical areas show a complex picture of age-related change. Some studies report increased firing rates (Chang et al., 2005; Hua et al., 2006; Schmolesky et al., 2000; Yang et al., 2009), while others have found decreased rates with ageing (Wang et al., 2011). 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 ratio (Yang et al., 2009), and increased noise correlations (Wang et al., 2019). Although informative, these findings are limited by relatively small datasets: e.g., Yang et al. (2009) recorded 172 and 173 neurons from 3 young and 4 old monkeys, respectively. Moreover, their generalization to awake, behaving 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 (Churchland et al., 2010; Iurilli & Datta, 2017; Licata et al., 2017; Miura et al., 2012; Poland et al., 2019). Yet, how this reduction of variability in response to sensory stimuli changes with ageing remains unknown, particularly in behaving animals.

Beyond primates, 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 (e.g., Randall et al., 2012; Simkin et al., 2015; Fischer et al., 2021; Turner et al., 2005; Hughes et al., 2010; Maziar et al., 2023; Barnes et al., 1997; Rosenzweig & Barnes, 2003; Barnes, 2011). Related changes have also been reported in rodent models of neurodegenerative disease, including Alzheimer's disease, which are often characterized by circuit hyperexcitability (H.-L. Wang et al., 2017; Sosulina et al., 2021; Anastacio et al., 2022). 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 anesthetized/ex vivo preparations. As a result, they provide limited leverage 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 tasks (Urai et al., 2022). We here build on the work of the International Brain Laboratory (IBL), which has standardized training and recording pipelines for mouse decision-making behavior (International Brain Laboratory, 2017; International Brain Laboratory et al., 2021; International Brain Laboratory, Banga, et al., 2025). Recording brain-wide neural activity during a visual decision-making task revealed that decision-related signals are distributed across much of the mouse brain (International Brain Laboratory, Benson, et al., 2025). 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 brain (International Brain Laboratory, Benson, et al., 2025). Here, we build on this well-curated dataset and contribute additional recordings from older mice, allowing us to investigate age-related changes in behavior 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 analyzed extracellular Neuropixels recordings from 149 mice (aged 3–20 months; roughly corresponding to 23–70-year-old humans; Cottam et al., 2025) across 16 brain regions (in cortex, striatum, midbrain, hippocampus, and thalamus; Table 1) while animals performed a standardized perceptual decision-making task (International Brain Laboratory et al., 2020).

We found that older animals showed higher trial-to-trial variability in response times. 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 behavioral differences in older animals may be accompanied by specific changes in neural variability in response to stimulus onset.

Results

We combined a previously released dataset of extracellular Neuropixels recordings (Jun et al., 2017) from young mice (N = 130, 89 male, mean age = 6.64 months, range 3.10–15.13 months; International Brain Laboratory et al., 2023) 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 behavior and neural recordings (Figure 1a; see Methods; International Brain Laboratory et al., 2021, 2021, 2024). 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 (Figure 1b), roughly corresponding to a human age range of 23–70 years (Cottam et al., 2025)—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 = 5.50 months) and an old group (N = 52, mean age = 11.56 months), using 7.6 months (mean age in the dataset) as the age cutoff (Figure 1b). Age was treated as a continuous variable in all statistical analyses.

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

Older animals did not take longer to learn the standardized task, and animals of different ages eventually passed the standardized criteria to be considered 'trained' (Figure 1-S1). Although animals that took longer to learn were then older by the time they moved on to neural recordings, the main source of age variability at recording time arises from starting the training program at a later age (Figure 1-S2). Only behavioral data from the recording sessions were included in our main analyses.

During recording sessions, older mice performed slightly more trials until automated behavior-based stopping criteria ended the session (International Brain Laboratory et al., 2021; range: 401–1279; Figure 1-S3a). 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, consistent with prior IBL analyses (International Brain Laboratory et al., 2024; Figure 1g). Unless otherwise specified, regression lines are shown only when the Bayes Factor indicates strong or moderate evidence in favor of H1. The number of in- or excluded trials did not show a significant correlation with age (included trial counts: Figure 1-S3c; number of trials excluded by RT filtering: Figure 1-S3b).

Older mice performed slightly worse on the task. Psychometric curves fit to each session (Figure 1c) showed that three psychometric function parameters—absolute bias, threshold, and mean lapse rate—showed an increase with age (Figure 1c, right panels). We did not observe age-related changes in choice bias conditioned on blocks or previous choices (Figure 1-S4), 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 task (Findling et al., 2023), they do not seem related to the age effects observed here.

Older mice did not respond more slowly, but were more variable in their response times (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 (Figure 1d, right), their RTs were notably more variable (Figure 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 (Dykiert et al., 2012; LaPlume et al., 2022; Li & Rieckmann, 2014; MacDonald et al., 2006). To our knowledge, this is the first time this effect has been replicated in behaving rodents. This result was robust to various control analyses: other measures of variability (Figure 1-S5a) and for a definition of RT that captures the first movement initiation, rather than movement completion (Figure 1-S5b), and differences between dataset sources or lab-specific environmental factors (Figure 1-S6, S7).

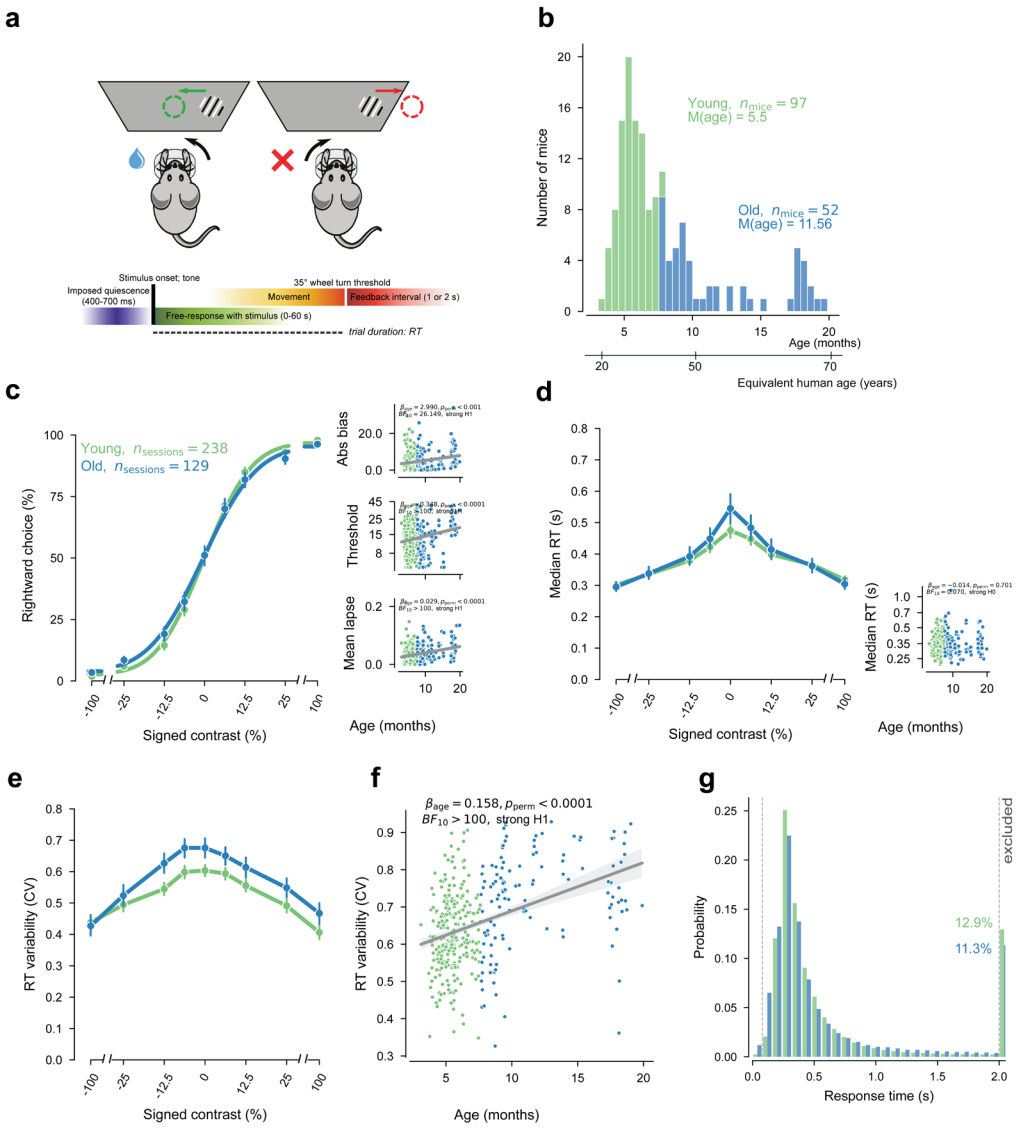


Figure 1. Older mice show larger RT variability and slightly worse performance in a standardized decision-making task. **(a)** Schematic of the task, showing a correct response (the visual stimulus was brought to the center of the screen) vs. a wrong response (the visual stimulus was moved off the screen) with resulting water reward vs. timeout. Response times were defined as the total duration between stimulus onset and completion of the response. Adapted from (International Brain Laboratory, Benson, et al., 2025). **(b)** Distribution of mouse age on the day of neural recording (green: young group, blue: old group; age cutoff: 7.6 months). Group splits were only used for visualization; all statistics used age as a continuous variable. The age range in our study roughly corresponds to a

human age range of 23–70 years (Cottam et al., 2025). **(c)** Left: Average psychometric curve across sessions for each age group. Right: The relationship between psychometric parameters and mouse age. Each dot represents one session. The slope (β) is estimated using a regression model, with accompanying p-value and Bayes Factor (see Methods). Unless otherwise noted, regression lines are plotted only when the Bayes Factor indicates 'strong or moderate H1'. Note the log-scaling on the y-axis of the threshold panel. **(d)** Left panel: Average chronometric curve across sessions for each age group. Right panel: Relationship between median RT and mouse age. Median RT was calculated across all included trials, irrespective of stimulus contrast. Note the log-scaling on the y-axis. **(e)** RT variability at different stimulus conditions for each age group. **(f)** The relationship between RT variability, measured using the coefficient of variation, and mouse age. **(g)** RT distribution by age group. The 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.

Large-scale Neuropixels recordings across the mouse lifespan

To investigate age-related changes in single-neuron activity, we analyzed 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 (Figure 2a,b; Table 1). These included 6 cortical regions, along with structures from the hippocampus, thalamus, midbrain, basal ganglia, and olfactory areas. 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 reproducibility (International Brain Laboratory, Banga, et al., 2025); 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 (Figure 2e, International Brain Laboratory et al., 2024; see Methods for details). From 242,671 recorded units (including multi-neuron activity) across these 16 regions, the quality control process identified 18,755 'good' neurons. Overall neural yield was slightly reduced in older mice (β_{age}=-0.106, p_{perm}=0.001; BF_{10} >100, strong H1), specifically in PPC, LS and ACA (Figure 2-S1).

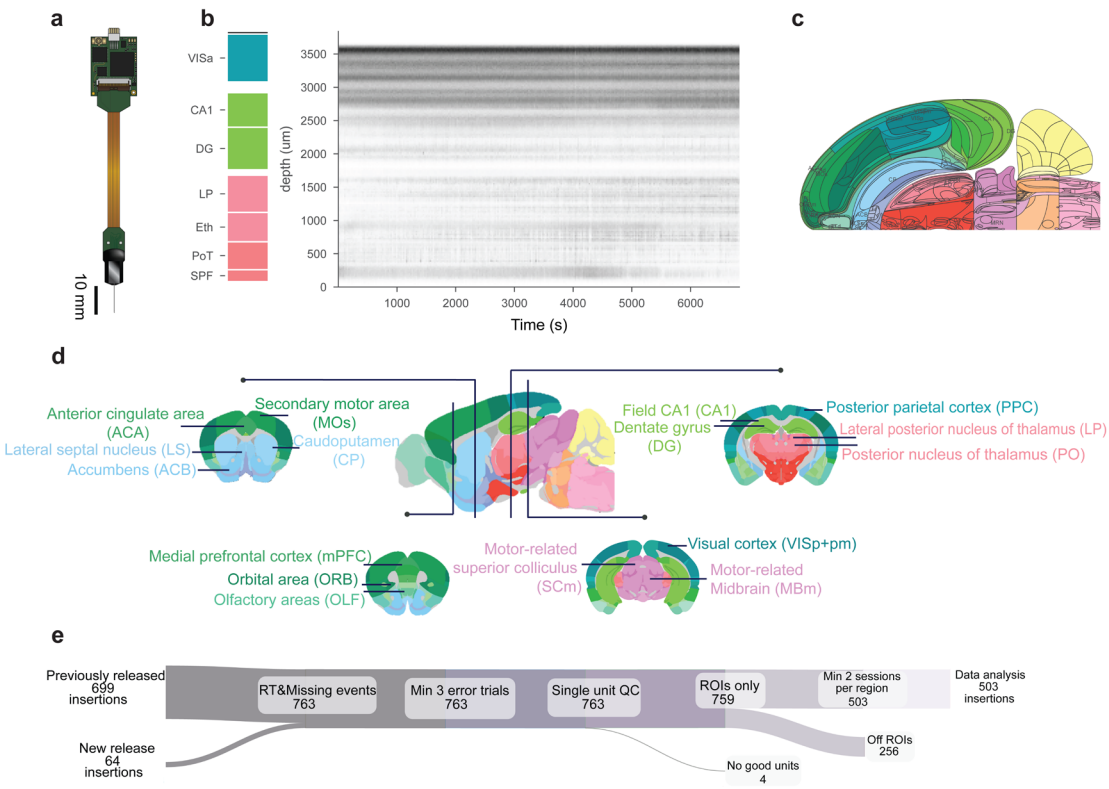


Figure 2. Large-scale neural recording across the mouse brain. (a) The Neuropixels probe used to record extracellular signals, adapted from (Jun et al., 2017). **(b)** Raster plot from one example recording probe, with aligned regions indicated on the left. Brain regions: VISa (anterior visual area), 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. **(c)** The 2D Swanson flatmap of the mouse brain. Labeled regions are ROIs as detailed in (d) and Table 1. **(d)** A 2D-sagittal mouse brain slice and four corresponding coronal slices, annotated with acronyms of ROIs, from the Allen Mouse Brain Atlas and Allen Reference Atlas – Mouse Brain. (mouse.brain-map.org and atlas.brain-map.org; Allen Institute for Brain Science, 2004, 2011) **(e)** The number of insertions remaining after each quality control step. Previously released: the public dataset from IBL; New release: dataset recorded during the present project.

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 (Hughes et al., 2010; Wilson et al., 2005; Bishop et al., 2022; see also Recanzone, 2018). Increased firing rates have also been reported in the sensory cortex of other mammals, including cat V1 (Hua et al., 2006). In non-human primates, ageing has been associated with increased firing rates in cortical areas such as macaque V1 and MT (Schmolesky et al., 2000; Yang et al., 2009), as well as in specific cell types in the prefrontal cortex (Chang et al., 2005 in monkey PFC pyramidal cells). In contrast, other studies have reported decreased firing rates with ageing, particularly in the prefrontal cortex (Wang et al., 2011 in monkey PFC). 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 (Figure 3a,b,d). 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) (Figure 3c,e, Figure 3-S1,2). Notably, the two thalamic areas LP and PO showed the reverse pattern of reduced firing rates in older animals (Figure 3c,e, Figure 3-S1,2). In PPC, an area causally implicated in age-related effects on learning and decision making (Hwang et al., 2023), 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 contrast (Figure 3f, Figure 3-S3; International Brain Laboratory, Benson, et al., 2025), but we found no significant effects of ageing on this global contrast modulation (Figure 3g). 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 (Hua et al., 2006; Schmolesky et al., 2000), and aged mice show declines in frequency selectivity within the auditory midbrain (Leong et al., 2011). 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 information (Koen & Rugg, 2019). 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 significant age effects on contrast modulation overall (Figure 3g). Region-wise estimates are shown in the corresponding Swanson map (Figure 3h) 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 (Figure 3h; Figure 3-S4), 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.

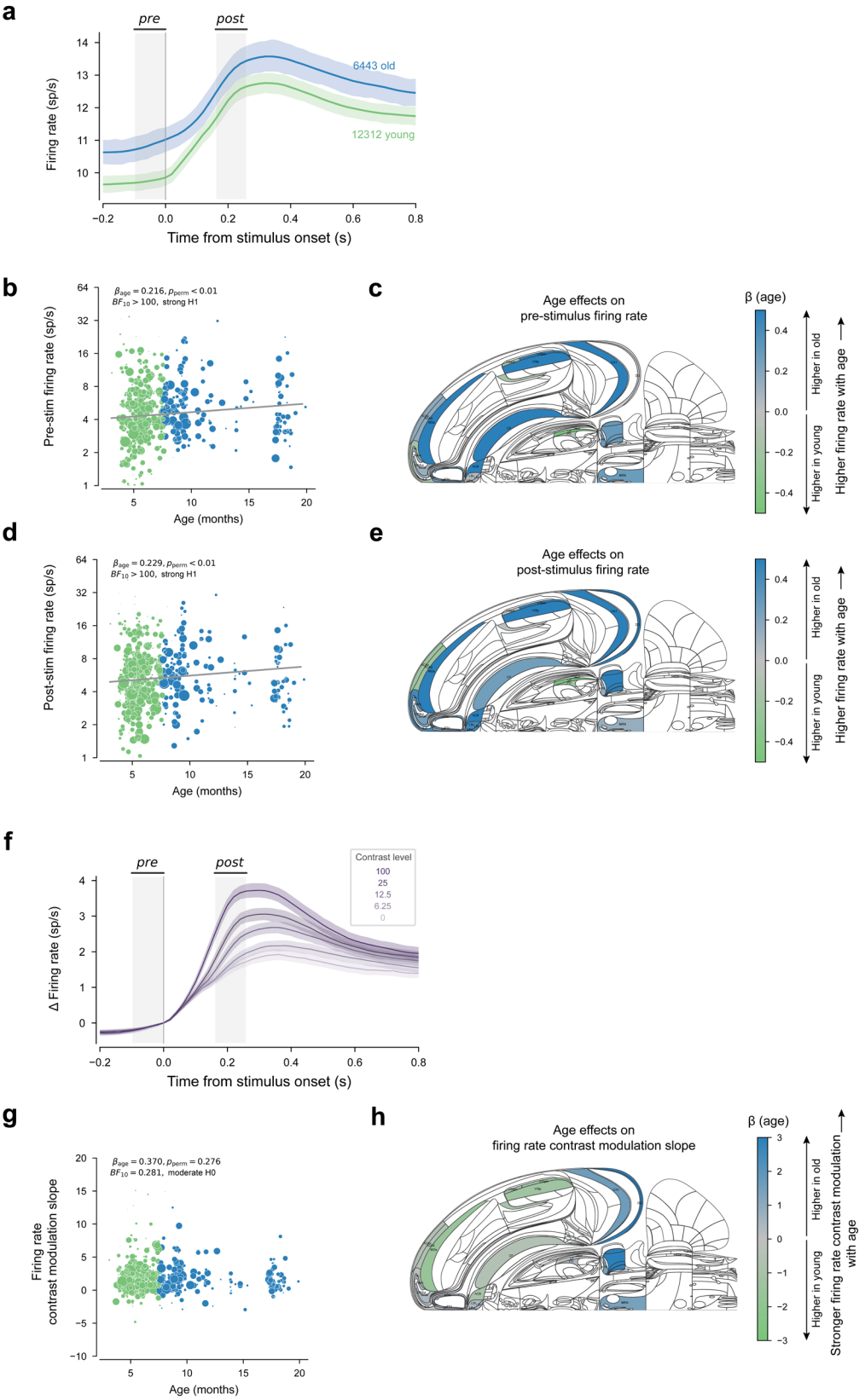

a
pre
post
6443 old
12312 young
Firing rate (sp/s)
Time from stimulus onset (s)
b
$\beta_{age} = 0.216, p_{perm} < 0.01$
$BF_{10} > 100$, strong H1
Pre-stim firing rate (sp/s)
Age (months)
c
Age effects on
pre-stimulus firing rate
β (age)
Higher in old
Higher in young
Higher firing rate with age
d
$\beta_{age} = 0.229, p_{perm} < 0.01$
$BF_{10} > 100$, strong H1
Post-stim firing rate (sp/s)
Age (months)
e
Age effects on
post-stimulus firing rate
Higher in old
Higher in young
Higher firing rate with age
f
pre
post
Contrast level
100
25
12.5
6.25
0
Δ Firing rate (sp/s)
Time from stimulus onset (s)
g
$\beta_{age} = 0.370, p_{perm} = 0.276$
$BF_{10} = 0.281$, moderate H0
Firing rate
contrast modulation slope
Age (months)
h
Age effects on
firing rate contrast modulation slope
β (age)
Higher in old
Higher in young
Stronger firing rate contrast modulation
with age

Figure 3. Global firing rate and contrast modulation. (a) Time courses of overall firing rates, aligned to stimulus onset. The number of neurons averaged within each age group is shown on the top right. Thick lines represent the mean firing rate in each group (blue = old, green = young). Shaded areas indicate 95% confidence intervals obtained by pooling neurons within each age group and are provided for visualization only. Statistical inferences about age effects are based on permutation tests and Bayes factors as described in the Methods. Sliding window width: 0.1 s; step size: 0.02 s. White areas indicate the pre (-100ms, 0ms) and post (160ms, 260ms) time windows separately. **(b)** The relationship between pre-stimulus firing rate and mouse age. Note the log-scaling on the y-axis. Each dot represents one insertion, with the size of the dots showing the number of neurons recorded in the insertion. Unless stated otherwise, scatter plots for neural metrics show insertion-level summaries for visualization, whereas statistical analyses and fitted regression lines are based on single-neuron metrics. **(c)** A Swanson map showing region-specific age effects on pre-stimulus firing rate. The color bar indicates the slope of age. Grey areas were not covered in our recordings. **(d)(e)** Same as (b)(c), but for post-stimulus firing rate. **(f)** Firing rate time courses across different contrast levels. Saturation represents different contrast levels. **(g)(h)** Same as (b)(c), 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. Because the omnibus analysis does not support a global age effect, this Swanson map is shown for completeness and is not interpreted in detail in the main text.

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 (e.g., Dean, 1981; Tolhurst et al., 1981, 1983; Tomko & Crapper, 1974). 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, Figure 4-S1), studies using Fano Factor have revealed significant regional differences in spiking variability across brain regions Goris et al., 2014; Kara et al., 2000; Poland et al., 2019; Schölvinck et al., 2015), as well as phenomena like 'quenching', where variability decreases at stimulus onset (Churchland et al., 2010). Reporting neural variability as Fano factor also facilitates direct comparison with prior work in this line of research (e.g., Yang et al., 2009; X. Wang et al., 2019; Liu et al., 2025). To control for biases in Fano Factor estimates due to condition-specific differences in mean firing rates, we applied a mean-subtraction correction method (Churchland et al., 2011; see Methods).

Time courses showed different effects of ageing on the Fano Factor between pre- and post-stimulus periods (Figure 4a). Ageing did not affect pre-stimulus Fano Factors (Figure 4b), suggesting that baseline neural variability remains stable with age. Following stimulus presentation, however, older animals exhibited increased Fano Factors (Figure 4d). This effect was most robustly observed in the midbrain (SCm, MBm) and thalamic area PO (Figure 4e, Figure 4-S2c). 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 anesthetized rhesus monkeys that reported increased Fano Factors in V1 and MT neurons of older animals (Yang et al., 2009). 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 (Churchland et al., 2010; Iurilli & Datta, 2017; Licata et al., 2017; Miura et al., 2012; Poland et al., 2019), 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 (Figure 4f). Our region-specific analyses revealed age-related attenuation in variability quenching in visual and motor cortex (VISp+pm, MOs), striatum (LS, ACB) and thalamic area LP (Figure 4g, Figure 4-S3).

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 (Figure 4-S4). 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 (Figure 4i). 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) (Figure 4j, Figure 4-S5b).

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 partly 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 (Figure 1-S2), and we did not observe slower learning in older mice (Figure 1-S1), training duration in the IBL task varies systematically across animals and laboratories (International Brain Laboratory et al., 2023). 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 (Table S1, S2): 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 (Table S2). 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 (Musall, Kaufman, et al., 2019), including in well-trained animals (Yin et al., 2025), 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) (Figure 4-S6 a,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 (Figure 4-S6, b-c,e-f). We next incorporated both movement measures (paw speed and nose-tip speed) as covariates in extended models for each neural metric (Table S3). 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 → weak H1; Fano Factor contrast modulation: weak H1 → moderate H1). All other metrics retained the same qualitative conclusion.

Together, these analyses suggest that while movement explains additional variance in firing-rate and variability metrics, the reported age-related differences persist after accounting for the two video-derived movement features and are therefore unlikely to be explained by these movement differences alone.

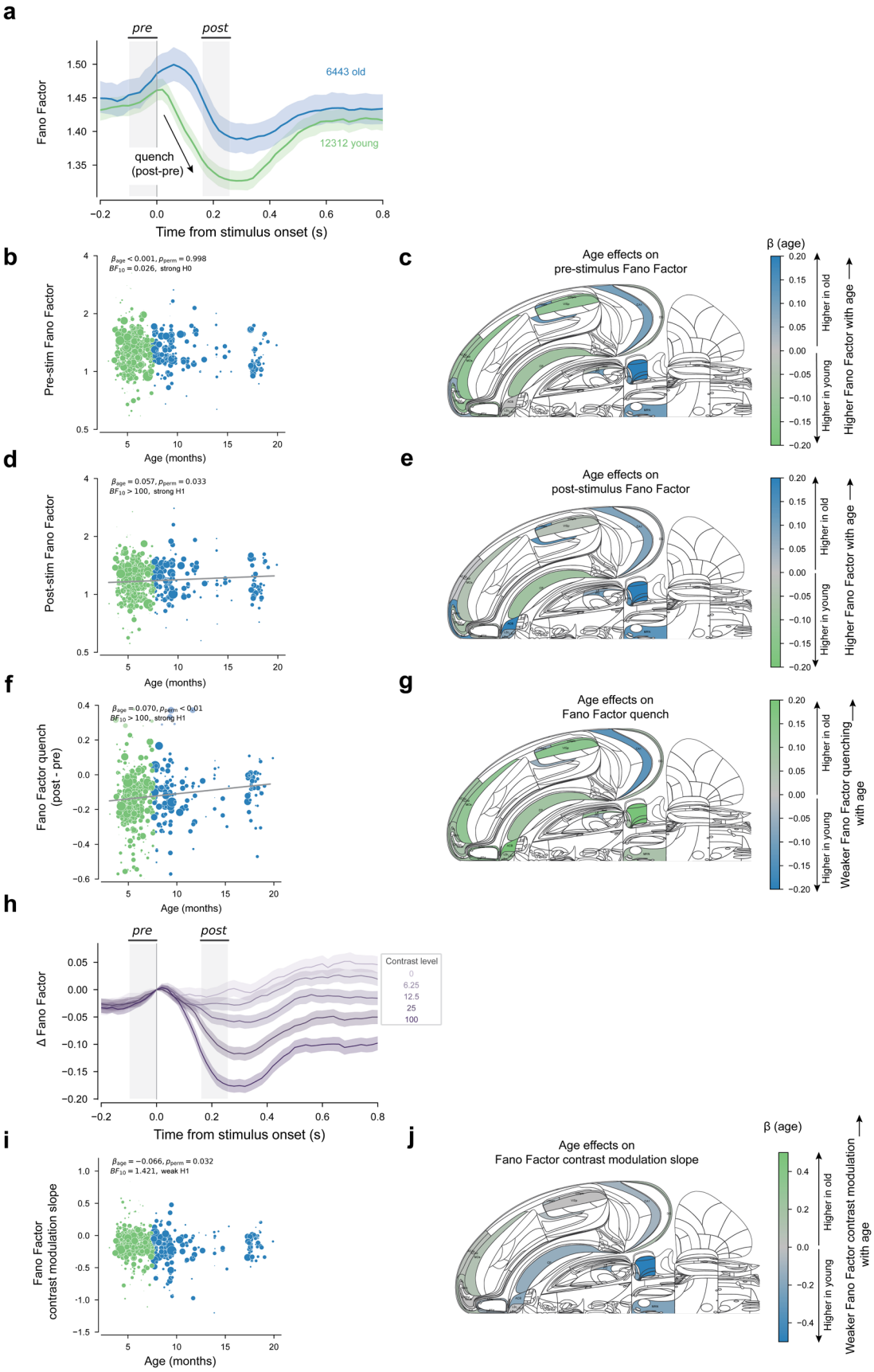

a
pre
post
Fano Factor
6443 old
12312 young
quench
(post-pre)
Time from stimulus onset (s)
b
Pre-stim Fano Factor
Age (months)
c
Age effects on
pre-stimulus Fano Factor
β (age)
Higher in old
Higher in young
Higher Fano Factor with age
d
Post-stim Fano Factor
Age (months)
e
Age effects on
post-stimulus Fano Factor
Higher in old
Higher in young
Higher Fano Factor with age
f
Fano Factor quench
(post - pre)
Age (months)
g
Age effects on
Fano Factor quench
Higher in old
Higher in young
Weaker Fano Factor quenching
with age
h
pre
post
Δ Fano Factor
Contrast level
0
6.25
12.5
25
100
Time from stimulus onset (s)
i
Fano Factor
contrast modulation slope
Age (months)
j
Age effects on
Fano Factor contrast modulation slope
β (age)
Higher in old
Higher in young
Weaker Fano Factor contrast modulation
with age

Figure 4 Global mean-subtracted Fano Factor and contrast modulation. (a) Time courses of the overall mean-subtracted Fano Factor, aligned to stimulus onset. The number of neurons averaged within each age group is shown on the top right. Thick lines represent the mean Fano Factor in each group (blue = old, green = young). Shaded areas indicate 95% confidence intervals obtained by pooling neurons within each age group and are provided for visualization only. Grey areas indicate the pre-stimulus (-100ms, 0ms) and post-stimulus (160ms, 260ms) time window separately. **(b)** The relationship between pre-stimulus (-100ms, 0ms) Fano Factor and mouse age. Each dot represents one insertion. The size of the dots shows the number of neurons recorded in the insertion. **(c)** The Swanson map showing region-specific age effects on pre-stimulus Fano Factor. The color indicates the slope of age. Grey areas were not covered in our recordings. **(d)(e)** Same as (b)(c), but showing post-stimulus Fano Factor. **(f)(g)** Same as (b) and (c), but showing Fano Factor quenching (post-stimulus Fano Factor - pre-stimulus Fano Factor). **(h)** Contrast modulation of Fano Factor quench. Saturation represents different contrast levels. **(i)(j)** Same as (b)(c), but for contrast modulation slope. Contrast modulation slopes are computed for each neuron by fitting a linear relationship between the change in baseline-corrected Fano Factor and the different stimulus contrast levels.

Discussion

In this study, we investigated age-related changes in behavior and neural activity using large-scale extracellular recordings in mice. Older mice performed slightly worse (showing increases in absolute bias, perceptual thresholds, and mean lapse rate), and showed increased variability in their response times. 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 (e.g., Greene & Madden, 1987; Nameda et al., 1989; Owsley, 2016; Owsley et al., 1983; Owsley & Sloane, 1987). However, visual deficits in C57BL/6 mouse strains start manifesting most dramatically after 20 months, beyond the age range of this study (Lehmann et al., 2012). 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 behavioral 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 (Dykiert et al., 2012; Li & Rieckmann, 2014; MacDonald et al., 2006). To our knowledge, our study provides the first explicit demonstration of this cross-species consistency in a mouse model, highlighting a potentially fundamental aspect of behavioral ageing conserved across 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 (Sabel & Stein, 1981; Stranahan et al., 2012; Yates et al., 2008). More broadly, ageing is accompanied by cytoarchitectural and structural remodeling, including changes in cellular organization, dendritic morphology, and synaptic integrity (Markham & Juraska, 2002; Morrison & Baxter, 2012). 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 (Hua et al., 2006 in cat V1; Schmolesky et al., 2000 in macaque V1) and specific cell types in the prefrontal cortex (Chang et al., 2005 in monkey PFC pyramidal cells). Similarly, we found an increase in post-stimulus Fano Factors with age, consistent with previous findings from anesthetized rhesus monkeys (Yang et al., 2009). Future work could also examine complementary measures of neural variability beyond Fano Factor to test the robustness and specificity of these effects (e.g., spike train irregularity; Holt et al., 1996; Nawrot et al., 2008). 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 behavioral 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 (Churchland et al., 2010; Poland et al., 2019), the PPC of rats (Licata et al., 2017), and the olfactory cortex of both mice (Iurilli & Datta, 2017) and rats (Miura et al., 2012), although thalamic nuclei did not show stimulus-induced quenching of neural variability (Poland et al., 2019). 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 response times. 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 behavior.

Changes in single-neuron variability may arise from various biological mechanisms. Age-related degradation of inhibitory and neuromodulatory systems (Guitart-Masip et al., 2016; Lee & Kim, 2022; Li & Rieckmann, 2014; MacDonald et al., 2006) may cause reduced gain modulation, shorter intrinsic neural timescales (Gao et al., 2020; Wu & Gollo, 2024), and lower temporal precision of single-neuron responses. These could in turn be reflected in pairwise noise correlations (Cohen & Maunsell, 2009; Wang et al., 2019), neural population dynamics (Jauny et al., 2022; Voytek et al., 2015; Wen et al., 2024), and cortical state (e.g., as reflected in multi-unit activity power spectra; Harris & Thiele, 2011; Schölvinck et al., 2015). Future work on such population-level measures may further reveal how age-related neural variability across scales impacts behavior.

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 behavioral 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 behavior. While we controlled for video-derived movement covariates (paw, nose-tip), future work could further refine such integrated predictors using richer behavioural characterizations (see Musall, Urai, et al., 2019), for example motion-energy–based video features or higher-dimensional, data-driven representations of behavioural modules (Wiltschko et al., 2015, 2020). Accordingly, a range of biological, behavioral, and life history factors could be further integrated to compute an individual animal's 'brain age' (Cole et al., 2019) as opposed to their chronological age, which may reveal higher predictive power in explaining between-animal differences in behavioral 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 animals (International Brain Laboratory, Benson, et al., 2025). 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 format (International Brain Laboratory et al., 2023) 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 (International Brain Laboratory, Benson, et al., 2025, the 'brain-wide map' dataset) from the International Brain Laboratory (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 (International Brain Laboratory, 2022a). 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 1411117 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 protocols (International Brain Laboratory et al., 2021). 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 (Figure 1–S2). After quality control, 149 mice were included in the formal analysis. For visualization, mice were categorized based on their age at recording into a young group (N = 97, mean age 5.50 months) and an old group (N = 52, mean age 11.56 months), using 7.6 months (mean age in the dataset) as the age cutoff (Figure 1b). Age was treated as a continuous variable in all statistical analyses.

Behavioral task

Mice were trained to perform a standardized visual decision-making task (International Brain Laboratory et al., 2021). 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 seconds. 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 center of the screen. They had up to 60 seconds 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 (positive for right stimuli, negative for left 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 behavioral task is further described in (International Brain Laboratory et al., 2021). The detailed protocol for animal training can be found in the Methods section of (International Brain Laboratory et al., 2021) and the IBL protocol for mice training (International Brain Laboratory, 2020).

Behavioral 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:

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

P is the probability of rightward choice;
c is the contrast level;
γ, λ, σ, μ 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 IBL Protocol (International Brain Laboratory, 2020). 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 seconds, i.e., timeout, or when the rotary encoder indicated the stimulus reached ±35º azimuth). A chronometric curve was computed for each session.
To assess RT variability, we computed the $Coefficient\ of\ Variation\ (CV)$ for each session.

$$CV = \frac{\sigma}{\mu}$$

Where:
$\sigma\ = standard\ deviation$
$\mu\ =\ mean\ of\ response\ times$

Neuropixels recordings

Extracellular recordings were acquired using Neuropixels probes (Figure 2a) (Jun et al., 2017). For a detailed description of the animal surgery, apparatus, and recording procedure, see Appendix 2 and 3 in (International Brain Laboratory et al., 2022). Up to two probes were inserted per recording session.

Spike sorting

All sessions were spike sorted using *ibl-sorter* (2.35.0) (International Brain Laboratory, 2022b), the same version used in (International Brain Laboratory, Meshulam, et al., 2025). The code is available at (International Brain Laboratory et al., 2024/2024).

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 those established by (International Brain Laboratory, Benson, et al., 2025), with additional requirements for single-neuron quality. These additions were made to ensure that our computations were based on stable neurons, to which the Fano Factor is particularly sensitive. The quality/inclusion/exclusion/quality control 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) The sessions have to pass the hardware tests (definitions can be found 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 (see the 'Recording Inclusion metrics and Guidelines for Optimal Reproducibility' (RIGOR) from (International Brain Laboratory,

Banga, et al., 2025)) (5) Insertions must have resolved alignments (see Appendix 6 of (International Brain Laboratory, Banga, et al., 2025) for definition). (6) Additionally, insertions must 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), feedback_times (time of feedback), stimOn_times (stimulus onset), or firstMovement_times (the time of the first detected movement with sufficient amplitude); and (2) trials with RTs outside the range of 0.08–2s (International Brain Laboratory, Benson, et al., 2025).

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 RIGOR (International Brain Laboratory, Banga, et al., 2025); (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).

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	ORBl, 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

Table 1. ROI definitions. This table details the composition of each defined ROI, including its abbreviation and the constituent brain regions.

ROIs include a combination of areas from the Beryl parcellation (International Brain Laboratory, Benson, et al., 2025). 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 labs; International Brain Laboratory et al., 2022), 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.

QC	**level**	**Previously released dataset** (International Brain Laboratory, Benson, et al., 2025)		**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

Table 2. Quality control process. 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.

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:

$$Fano\ Factor\ = \frac{\sigma^2}{\mu}$$

Where:
$\sigma^2\ = variance\ of\ spike\ count\ across\ trials$
$\mu\ =\ mean\ spike\ count\ across\ trials$

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 method from (Churchland et al., 2011) to pool trials while controlling for condition-specific differences in mean firing rates. Specifically, for each trial, we subtracted the mean spike count of its condition to obtain a residual. Fano Factors were then computed from the variance of these residuals using a sliding-window approach (width = 0.1 s, step size = 0.02 s) for each neuron.

$$FF_{\langle N_i \rangle} = \frac{Var[Z_{\cup}]}{\sum_{i=1}^{M} \frac{n_i}{n_{\cup}} \overline{N_i}}$$

It is defined as the variance of the union of residuals from all conditions $Z_{\cup}$, divided by the weighted average of the mean spike count for the conditions. where $n_{\cup}$ is the total number of trials across all M conditions, n_i and $\overline{N_i}$ 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 used Bayes Factors and complementary p-values to test the effect of age on each behavioral and neural metric. Behavioral 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 analyzed 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 a statistically significant relationship between each metric and mouse age, we used the slope of age (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. All main inferences about age effects are based on Bayes Factors computed on these summary statistics and permutation tests with age-label shuffling.

The models were specified as follows:

Behavioral metrics:

```
For median RT, psychometric threshold:
     log metric ~ age_years
For RT variability, absolute bias, mean lapse:
     metric ~ age_years
```

Neural metrics:

```
neural yield ~ age_years + C(cluster_region)
(pre, post) log firing rate ~ age_years + C(cluster_region) +
abs_signed_contrast + n_trials
FR_contrast_modulation_slope ~ age_years + C(cluster_region)
(pre, post) log Fano Factor ~ age_years + C(cluster_region) + n_trials
Fano Factor_quench ~ age_years + C(cluster_region) + n_trials
FF_contrast_modulation_slope ~ age_years + C(cluster_region)
```

We computed Bayes Factors (BF_{10}), using the *BayesFactor* package (Morey & Rouder, 2023) and *Pingouin* (Vallat, 2018), 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 behavioral or neural metric) and <1 favoring the null.

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 behavioral metrics; n = 2,000 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.

The two approaches offer complementary strengths: permutation tests account for the hierarchical data structure by shuffling neurons based on sessions, providing robust frequentist inference, while Bayes Factors provide a graded measure of evidence for or against an effect. 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).

For visualization purposes, regression lines are plotted only when the BF_{10} provides strong (BF_{10}>10) or moderate (10 ≥ BF_{10}> 3) evidence in favor 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.

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 standardised 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_years, 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

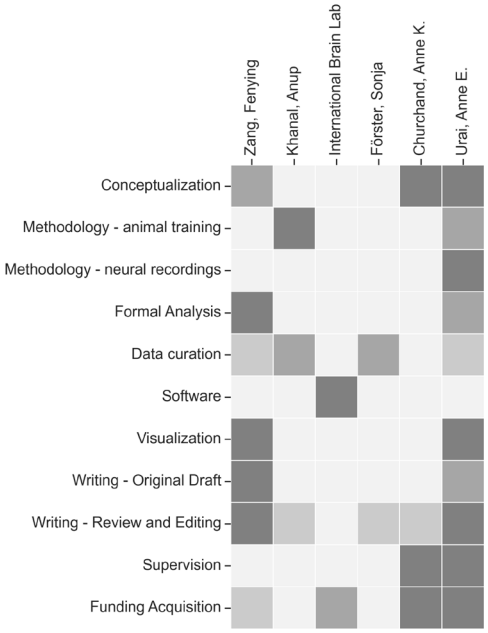
We used IBL right-camera videos to derive movement covariates from DeepLabCut tracking of two representative keypoints: right paw (task-relevant) and nose tip (task-irrelevant) (Figure 4-S6a,d). From the tracked x/y coordinates, we computed frame-to-frame speed (coordinates normalized by camera resolution and scaled by the camera sampling rate), and aligned the resulting speed time courses to stimulus onset. We summarized movement at the session level by computing median speed within the same pre- and post-stimulus windows used for neural metrics (pre: −0.1 to 0 s; post: 0.16 to 0.26 s), yielding four covariates (pre/post paw speed; pre/post nose-tip speed).

Sessions were required to pass IBL extended QC for the right camera (videoRight and dlcRight). This yielded n = 343 sessions with valid right-camera movement measures. We then fitted extended models for each neural metric by adding the movement covariates to the age-only models. Movement covariates were time-window matched: metrics defined in a single window used the corresponding pre- or post-stimulus movement measure, whereas metrics spanning both windows (e.g., Fano Factor quenching) included both pre- and post-stimulus movement covariates. All other modelling choices (predictors, covariates, log-transform conventions, and inference procedures) were identical to those described above.

Data and code availability

All data are publicly available through the IBL website (https://www.internationalbrainlab.com/data). Instructions and code are available at https://github.com/Fenying-Zang/Ageing_behavioral_and_neural_variability.

Author contributions



Competing interests

The authors declare no competing interests.

Acknowledgements

AEU thanks Graham Wildt and John Pisciotta 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 Steenbergen, Bryant Jongkees, Steven Miletić, and Jacqueline Zadelaar provided valuable feedback on statistical analyses. Sander Nieuwenhuis offered helpful comments on the first draft. We are grateful to the IBL development team, especially Olivier Winter, for technical support with spike sorting. Peter Dayan and Kenneth Harris provided insightful suggestions throughout the project.

The IBL was supported by grants from the Wellcome Trust (216324) and the Simons Collaboration on the Global Brain. AEU 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. FZ was supported by a PhD fellowship (202204910080) from the Chinese Scholarship Council. AKC was supported by a grant from the Simons Collaboration on Plasticity and the Aging Brain.

Consortia

Larry Abbot[3], Luigi Acerbi[4], Valeria Aguillon-Rodriguez[5], Mandana Ahmadi[6], Jaweria Amjad[6], Dora Angelaki[7], Jaime Arlandis[8], Zoe C. Ashwood[9], Kush Banga[10], Hailey Barrell[11], Hannah M. Bayer[3], Brandon Benson[12], Julius Benson[7], Jai Bhagat[10], Dan Birman[11], Niccolò Bonacchi[8], Kcenia Bougrova[8], Julien Boussard[3], Sebastian A. Bruijns[13], E Kelly Buchanan[3], Robert

Campbell[14], Matteo Carandini[15], Joana A. Catarino[8], Fanny Cazettes[8], Gaelle A. Chapuis[4], Anne K. Churchland[2], Davide Crombie[8], Yang Dan[16], Felicia Davatolhagh[2], Peter Dayan[13], Sophie Denève[17], Eric EJ DeWitt[8], Tatiana Engel[9], Michele Fabbri[3], Mayo Faulkner[10], Robert Fetcho[9], Ila Fiete[18], Charles Findling[4], Laura Freitas-Silva[8], Surya Ganguli[12], Berk Gercek[4], Naureen Ghani[14], Ivan Gordeliy[17], Laura M. Haetzel[9], Kenneth D. Harris[10], Michael Hausser[19], Naoki Hiratani[6], Sonja Hofer[14], Fei Hu[16], Felix Huber[4], Julia M. Huntenburg[13], Cole Hurwitz[3], Anup Khanal[2], Christopher S. Krasniak[5,20], Sanjukta Krishnagopal[6], Michael Krumin[10], Debottam Kundu[13], Agnès Landemard[15], Christopher Langdon[9], Christopher Langfield[3], Inês C. Laranjeira[8], Peter Latham[6], Petrina Lau[19], Hyun Dong Lee[3], Ari Liu[18], Zachary F. Mainen[8], Amalia Makri-Cottington[19], Hernando Martinez-Vergara[14], Brenna McMannon[9], Isaiah McRoberts[7], Guido T. Meijer[8], Maxwell Melin[2], Leenoy Meshulam[21], Kim Miller[11], Nathaniel J. Miska[14], Catalin Mitelut[3], Zeinab Mohammadi[9], Thomas Mrsic-Flogel[14], Masayoshi Murakami[22], Jean-Paul Noel[7], Kai Nylund[11], Farideh Oloomi[13], Alejandro Pan Vazquez[9], Liam Paninski[3], Sabrina Perrenoud[19], Alberto Pezzotta[6], Samuel Picard[10], Jonathan W. Pillow[9], Alexandre Pouget[4], Carolina Quadrado[10], Pranav Rai[8], Georg Raiser[8], Florian Rau[8], Cyrille Rossant[10], Noam Roth[11], Nicholas A. Roy[9], Kamron Saniee[3], Rylan Schaeffer[18], Michael M. Schartner[8], Yanliang Shi[9], Karolina Z. Socha[15], Cristian Soitu[5], Nicholas A. Steinmetz[11], Karel Svoboda[23], Marsa Taheri[2], Charline Tessereau[13], Matthew Tucker[11], Anne E. Urai[1], Erdem Varol[3], Shuqi Wang[3], Miles J. Wells[10], Steven J. West[14], Matthew R. Whiteway[3], Charles Windolf[3], Olivier Winter[8], Ilana Witten[9], Lauren E. Wool[10], Zekai Xu[6], Kenneth Yang[11], Yaxuan Yang[14], Han Yu[3], Anthony M. Zador[5] & Yizi Zhang[3]

Affiliations (consortium members)

[1] Leiden University, The Netherlands
[2] University of California Los Angeles, USA
[3] Zuckerman Institute, Columbia University, New York, NY, USA
[4] Department of Basic Neuroscience, University of Geneva, Geneva, Switzerland
[5] Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, USA
[6] Gatsby Computational Neuroscience Unit, University College London, London, UK
[7] Center for Neural Science, New York University, New York, NY, USA
[8] Champalimaud Center for the Unknown, Lisboa, Portugal
[9] Princeton Neuroscience Institute, Princeton University, Princeton, NJ, USA
[10] Institute of Neurology, University College London, London, UK
[11] Department of Biological Structure, University of Washington, Seattle, WA, USA
[12] Department of Applied Physics, Stanford University, Stanford, CA, USA
[13] Max Planck Institute for Biological Cybernetics, Tübingen, Germany
[14] Sainsbury-Wellcome Centre, University College London, London, UK
[15] Institute of Opthalmology, University College London, London, United Kingdom
[16] Department of Molecular and Cell Biology, University of California, Berkeley, CA, USA
[17] Département D'études Cognitives, École Normale Supérieure, Paris, France
[18] Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA, USA
[19] Wolfson Institute of Biomedical Research, University College London, London, United Kingdom
[20] Watson School of Biological Science, Cold Spring Harbor, NY, USA
[21] Center for Computational Neuroscience, University of Washington, Seattle, WA, USA
[22] Department of Physiology, University of Yamanashi, Kofu, Yamanashi, Japan
[23] The Allen Institute for Neural Dynamics, Seattle, Washington, USA

Bibliography

Allen Institute for Brain Science. (2004). *Allen Mouse Brain Atlas* [Database]. https://mouse.brain-map.org

Allen Institute for Brain Science. (2011). *Allen Reference Atlas – Mouse Brain* [Brain atlas]. https://atlas.brain-map.org/

Anastacio, H. T. D., Matosin, N., & Ooi, L. (2022). Neuronal hyperexcitability in Alzheimer's disease: What are the drivers behind this aberrant phenotype? *Translational Psychiatry*, *12*(1), 257. https://doi.org/10.1038/s41398-022-02024-7

Barnes, C. A. (2011). Secrets of aging: What does a normally aging brain look like? *F1000 Biology Reports*, *3*, 22. https://doi.org/10.3410/B3-22

Barnes, C. A., Suster, M. S., Shen, J., & McNaughton, B. L. (1997). Multistability of cognitive maps in the hippocampus of old rats. *Nature*, *388*(6639), 272–275. https://doi.org/10.1038/40859

Boylan, M. A., Foster, C. M., Pongpipat, E. E., Webb, C. E., Rodrigue, K. M., & Kennedy, K. M. (2021). Greater BOLD Variability is Associated With Poorer Cognitive Function in an Adult Lifespan Sample. *Cerebral Cortex*, *31*(1), 562–574. https://doi.org/10.1093/cercor/bhaa243

Chang, Y.-M., Rosene, D. L., Killiany, R. J., Mangiamele, L. A., & Luebke, J. I. (2005). Increased Action Potential Firing Rates of Layer 2/3 Pyramidal Cells in the Prefrontal Cortex are Significantly Related to Cognitive Performance in Aged Monkeys. *Cerebral Cortex*, *15*(4), 409–418. https://doi.org/10.1093/cercor/bhh144

Churchland, Anne. K., Kiani, R., Chaudhuri, R., Wang, X.-J., Pouget, A., & Shadlen, M. N. (2011). Variance as a Signature of Neural Computations during Decision Making. *Neuron*, *69*(4), 818–831. https://doi.org/10.1016/j.neuron.2010.12.037

Churchland, M. M., Yu, B. M., Cunningham, J. P., Sugrue, L. P., Cohen, M. R., Corrado, G. S., Newsome, W. T., Clark, A. M., Hosseini, P., Scott, B. B., Bradley, D. C., Smith, M. A., Kohn, A., Movshon, J. A., Armstrong, K. M., Moore, T., Chang, S. W., Snyder, L.

H., Lisberger, S. G., … Shenoy, K. V. (2010). Stimulus onset quenches neural variability: A widespread cortical phenomenon. *Nature Neuroscience*, *13*(3), Article 3. https://doi.org/10.1038/nn.2501

Cohen, M. R., & Maunsell, J. H. R. (2009). Attention improves performance primarily by reducing interneuronal correlations. *Nature Neuroscience*, *12*(12), Article 12. https://doi.org/10.1038/nn.2439

Cole, J. H., Marioni, R. E., Harris, S. E., & Deary, I. J. (2019). Brain age and other bodily 'ages': Implications for neuropsychiatry. *Molecular Psychiatry*, *24*(2), 266–281. https://doi.org/10.1038/s41380-018-0098-1

Cottam, N. C., Ofori, K., Stoll, K. T., Bryant, M., Rogge, J. R., Hekmatyar, K., Sun, J., & Charvet, C. J. (2025). From Circuits to Lifespan: Translating Mouse and Human Timelines with Neuroimaging-Based Tractography. *Journal of Neuroscience*, *45*(12). https://doi.org/10.1523/JNEUROSCI.1429-24.2025

Cremer, R., & Zeef, E. J. (1987). What Kind of Noise Increases With Age? *Journal of Gerontology*, *42*(5), 515–518. https://doi.org/10.1093/geronj/42.5.515

Crossman, E. R. & Szafran, J. (1956). Changes with age in the speed of information-intake and discrimination. *Experientia*, 128–135.

Dean, A. F. (1981). The variability of discharge of simple cells in the cat striate cortex. *Experimental Brain Research*, *44*(4), 437–440. https://doi.org/10.1007/BF00238837

Dykiert, D., Der, G., Starr, J. M., & Deary, I. J. (2012). Age Differences in Intra-Individual Variability in Simple and Choice Reaction Time: Systematic Review and Meta-Analysis. *PLOS ONE*, *7*(10), e45759. https://doi.org/10.1371/journal.pone.0045759

Fischer, T. F., Czerniak, A. S., Weiß, T., Schoeder, C. T., Wolf, P., Seitz, O., Meiler, J., & Beck-Sickinger, A. G. (2021). Ligand-binding and -scavenging of the chemerin receptor GPR1. *Cellular and Molecular Life Sciences*, *78*(17), 6265–6281. https://doi.org/10.1007/s00018-021-03894-8

Gao, R., van den Brink, R. L., Pfeffer, T., & Voytek, B. (2020). Neuronal timescales are functionally dynamic and shaped by cortical microarchitecture. *eLife*, *9*, e61277.

https://doi.org/10.7554/eLife.61277

Garrett, D. D., Kovacevic, N., McIntosh, A. R., & Grady, C. L. (2010). Blood Oxygen Level-Dependent Signal Variability Is More than Just Noise. *Journal of Neuroscience*, *30*(14), 4914–4921. https://doi.org/10.1523/JNEUROSCI.5166-09.2010

Garrett, D. D., Kovacevic, N., McIntosh, A. R., & Grady, C. L. (2011). The importance of being variable. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *31*(12), 4496–4503. https://doi.org/10.1523/JNEUROSCI.5641-10.2011

Goris, R. L. T., Movshon, J. A., & Simoncelli, E. P. (2014). Partitioning neuronal variability. *Nature Neuroscience*, *17*(6), Article 6. https://doi.org/10.1038/nn.3711

Greene, H. A., & Madden, D. J. (1987). Adult Age Differences in Visual Acuity, Stereopsis, and Contrast Sensitivity. *Optometry and Vision Science*, *64*(10), 749.

Guitart-Masip, M., Salami, A., Garrett, D., Rieckmann, A., Lindenberger, U., & Bäckman, L. (2016). BOLD Variability is Related to Dopaminergic Neurotransmission and Cognitive Aging. *Cerebral Cortex*, *26*(5), 2074–2083. https://doi.org/10.1093/cercor/bhv029

Harris, K. D., & Thiele, A. (2011). Cortical state and attention. *Nature Reviews Neuroscience*, *12*(9), 509–523. https://doi.org/10.1038/nrn3084

Holt, G. R., Softky, W. R., Koch, C., & Douglas, R. J. (1996). Comparison of discharge variability in vitro and in vivo in cat visual cortex neurons. *Journal of Neurophysiology*, *75*(5), 1806–1814. https://doi.org/10.1152/jn.1996.75.5.1806

Hua, T., Li, X., He, L., Zhou, Y., Wang, Y., & Leventhal, A. G. (2006). Functional degradation of visual cortical cells in old cats. *Neurobiology of Aging*, *27*(1), 155–162. https://doi.org/10.1016/j.neurobiolaging.2004.11.012

Hughes, L. F., Turner, J. G., Parrish, J. L., & Caspary, D. M. (2010). Processing of Broadband Stimuli Across A1 Layers in Young and Aged Rats. *Hearing Research*, *264*(1–2), 79–85. https://doi.org/10.1016/j.heares.2009.09.005

Hwang, E. J., Korde, S., Han, Y., Sambangi, J., Lian, B., Owusu-Ofori, A., Diasamidze, M.,

Wong, L. M., Pickering, N., & Begin, S. (2023). *Parietal stimulation reverses age-related decline in exploration, learning, and decision-making* (p. 2023.10.21.563408). bioRxiv. https://doi.org/10.1101/2023.10.21.563408

International Brain Laboratory. (2017). An International Laboratory for Systems and Computational Neuroscience. *Neuron*, *96*(6), 1213–1218. https://doi.org/10.1016/j.neuron.2017.12.013

International Brain Laboratory. (2020). *Behavior: Appendix 2: IBL protocol for mice training*. figshare. https://doi.org/10.6084/m9.figshare.11634729.v3

International Brain Laboratory. (2022a, May 4). *Appendix 2: IBL protocol for electrophysiology recording using Neuropixels probe* [Online resource]. figshare. Figshare. https://doi.org/10.6084/m9.figshare.19697896.v2

International Brain Laboratory. (2022b, May 4). *Spike sorting pipeline for the International Brain Laboratory* [Online resource]. figshare. Figshare. https://doi.org/10.6084/m9.figshare.19705522.v4

International Brain Laboratory, Aguillon-Rodriguez, V., Angelaki, D., Bayer, H., Bonacchi, N., Carandini, M., Cazettes, F., Chapuis, G., Churchland, A. K., Dan, Y., Dewitt, E., Faulkner, M., Forrest, H., Haetzel, L., Häusser, M., Hofer, S. B., Hu, F., Khanal, A., Krasniak, C., … Zador, A. M. (2021). Standardized and reproducible measurement of decision-making in mice. *eLife*, *10*, e63711. https://doi.org/10.7554/eLife.63711

International Brain Laboratory, Banga, K., Benson, J., Bhagat, J., Biderman, D., Birman, D., Bonacchi, N., Bruijns, S. A., Buchanan, K., Campbell, R. A. A., Carandini, M., Chapuis, G. A., Churchland, A. K., Davatolhagh, M. F., Lee, H. D., Faulkner, M., Gerçek, B., Hu, F., Huntenburg, J., … Zhang, Y. (2025). Reproducibility of in vivo electrophysiological measurements in mice. *eLife*, *13*, RP100840. https://doi.org/10.7554/eLife.100840

International Brain Laboratory, Banga, K., Faulkner, M., Langfield, C., Rossant, C., & Winter, O. (2024). *Ibl-sorter* (Version 1.8.0) [Python]. https://github.com/int-brain-lab/ibl-sorter (Original work published 2024)

International Brain Laboratory, Benson, B., Benson, J., Birman, D., Bonacchi, N., Carandini, M., Catarino, J. A., Chapuis, G. A., Churchland, A. K., Dan, Y., Dayan, P., DeWitt, E. E., Engel, T. A., Fabbri, M., Faulkner, M., Fiete, I. R., Findling, C., Freitas-Silva, L., Gercek, B., … Witten, I. B. (2025). A Brain-Wide Map of Neural Activity during Complex Behaviour. *Nature*, In press.

International Brain Laboratory, Bonacchi, N., Chapuis, G. A., Churchland, A. K., DeWitt, E. E. J., Faulkner, M., Harris, K. D., Huntenburg, J. M., Hunter, M., Laranjeira, I. C., Rossant, C., Sasaki, M., Schartner, M. M., Shen, S., Steinmetz, N. A., Walker, E. Y., West, S. J., Winter, O., & Wells, M. J. (2023). A modular architecture for organizing, processing and sharing neurophysiology data. *Nature Methods*, *20*(3), 403–407. https://doi.org/10.1038/s41592-022-01742-6

International Brain Laboratory, Meshulam, L., Angelaki, D., Benson, B., Benson, J., Birman, D., Bruijns, S. A., Carandini, M., Catarino, J. A., Chapuis, G. A., Churchland, A. K., Dan, Y., Davatolhagh, F., Dayan, P., DeWitt, E. E., Engel, T. A., Fabbri, M., Faulkner, M., Fiete, I. R., … Witten, I. B. (2025). A brain-wide map of neural activity during complex behaviour. *Nature*, *645*(8079), 177–191. https://doi.org/10.1038/s41586-025-09235-0

Iurilli, G., & Datta, S. R. (2017). Population Coding in an Innately Relevant Olfactory Area. *Neuron*, *93*(5), 1180-1197.e7. https://doi.org/10.1016/j.neuron.2017.02.010

Jauny, G., Eustache, F., & Hinault, T. (2022). Connectivity dynamics and cognitive variability during aging. *Neurobiology of Aging*, *118*, 99–105. https://doi.org/10.1016/j.neurobiolaging.2022.07.001

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

Kara, P., Reinagel, P., & Reid, R. C. (2000). Low Response Variability in Simultaneously

Recorded Retinal, Thalamic, and Cortical Neurons. *Neuron*, *27*(3), 635–646. https://doi.org/10.1016/S0896-6273(00)00072-6

Koen, J. D., & Rugg, M. D. (2019). Neural Dedifferentiation in the Aging Brain. *Trends in Cognitive Sciences*, *23*(7), 547–559. https://doi.org/10.1016/j.tics.2019.04.012

Kügler, C. F. A., Taghavy, A., & Platt, D. (1993). The Event-Related P300 Potential Analysis of Cognitive Human Brain Aging: A Review. *Gerontology*, *39*(5), 280–303. https://doi.org/10.1159/000213544

LaPlume, A. A., Anderson, N. D., McKetton, L., Levine, B., & Troyer, A. K. (2022). When I'm 64: Age-Related Variability in Over 40,000 Online Cognitive Test Takers. *The Journals of Gerontology: Series B*, *77*(1), 104–117. https://doi.org/10.1093/geronb/gbab143

Lee, J., & Kim, H.-J. (2022). Normal Aging Induces Changes in the Brain and Neurodegeneration Progress: Review of the Structural, Biochemical, Metabolic, Cellular, and Molecular Changes. *Frontiers in Aging Neuroscience*, *14*. https://doi.org/10.3389/fnagi.2022.931536

Lehmann, K., Schmidt, K.-F., & Löwel, S. (2012). Vision and visual plasticity in ageing mice. *Restorative Neurology and Neuroscience*, *30*(2), 161–178. https://doi.org/10.3233/RNN-2012-110192

Leong, U.-C., Barsz, K., Allen, P. D., & Walton, J. P. (2011). Neural correlates of age-related declines in frequency selectivity in the auditory midbrain. *Neurobiology of Aging*, *32*(1), 168–178. https://doi.org/10.1016/j.neurobiolaging.2009.01.006

Li, S.-C., & Rieckmann, A. (2014). Neuromodulation and aging: Implications of aging neuronal gain control on cognition. *Current Opinion in Neurobiology, SI: Neuromodulation*, *29*, 148–158. https://doi.org/10.1016/j.conb.2014.07.009

Licata, A. M., Kaufman, M. T., Raposo, D., Ryan, M. B., Sheppard, J. P., & Churchland, A. K. (2017). Posterior Parietal Cortex Guides Visual Decisions in Rats. *Journal of Neuroscience*, *37*(19), 4954–4966. https://doi.org/10.1523/JNEUROSCI.0105-17.2017

Liu, X., Zhou, Y., Liu, J., & Xu, G. (2025). Metformin improves age-related visual cortex dysfunction in mice by reducing noise correlation in the primary visual cortex. *Frontiers in Aging Neuroscience*, *17*. https://doi.org/10.3389/fnagi.2025.1572653

MacDonald, S. W. S., Nyberg, L., & Bäckman, L. (2006). Intra-individual variability in behavior: Links to brain structure, neurotransmission and neuronal activity. *Trends in Neurosciences*, *29*(8), 474–480. https://doi.org/10.1016/j.tins.2006.06.011

Markham, J. A., & Juraska, J. M. (2002). Aging and sex influence the anatomy of the rat anterior cingulate cortex. *Neurobiology of Aging*, *23*(4), 579–588. https://doi.org/10.1016/s0197-4580(02)00004-0

Maziar, A., Critch, T. N. R. H. Y., Ghosh, S., Rajani, V., Flynn, C. M., Qin, T., Reinhardt, C., Man, K. N. M., Lee, A., Hell, J. W., & Yuan, Q. (2023). Aging differentially affects LTCC function in hippocampal CA1 and piriform cortex pyramidal neurons. *Cerebral Cortex*, *33*(4), 1489–1503. https://doi.org/10.1093/cercor/bhac152

Miura, K., Mainen, Z. F., & Uchida, N. (2012). Odor Representations in Olfactory Cortex: Distributed Rate Coding and Decorrelated Population Activity. *Neuron*, *74*(6), 1087–1098. https://doi.org/10.1016/j.neuron.2012.04.021

Morey R & Rouder J. (2023). *BayesFactor: Computation of Bayes Factors for Common Designs* (Version R package version 0.9.12-4.6) [Computer software]. https://rdrr.io/cran/BayesFactor/

Morrison, J. H., & Baxter, M. G. (2012). The ageing cortical synapse: Hallmarks and implications for cognitive decline. *Nature Reviews Neuroscience*, *13*(4), Article 4. https://doi.org/10.1038/nrn3200

Musall, S., Kaufman, M. T., Juavinett, A. L., Gluf, S., & Churchland, A. K. (2019). Single-trial neural dynamics are dominated by richly varied movements. *Nature Neuroscience*, *22*(10), 1677–1686. https://doi.org/10.1038/s41593-019-0502-4

Musall, S., Urai, A. E., Sussillo, D., & Churchland, A. K. (2019). Harnessing behavioral diversity to understand neural computations for cognition. *Current Opinion in Neurobiology, Computational Neuroscience*, *58*, 229–238.

https://doi.org/10.1016/j.conb.2019.09.011

Nameda, N., Kawara, T., & Ohzu, H. (1989). Human Visual Spatio-Temporal Frequency Performance as a Function of Age. *Optometry and Vision Science*, *66*(11), 760.

Nawrot, M. P., Boucsein, C., Rodriguez Molina, V., Riehle, A., Aertsen, A., & Rotter, S. (2008). Measurement of variability dynamics in cortical spike trains. *Journal of Neuroscience Methods, Methods for Computational Neuroscience*, *169*(2), 374–390. https://doi.org/10.1016/j.jneumeth.2007.10.013

Nomi, J. S., Bolt, T. S., Ezie, C. E. C., Uddin, L. Q., & Heller, A. S. (2017). Moment-to-Moment BOLD Signal Variability Reflects Regional Changes in Neural Flexibility across the Lifespan. *Journal of Neuroscience*, *37*(22), 5539–5548. https://doi.org/10.1523/JNEUROSCI.3408-16.2017

Owsley, C. (2016). Vision and Aging. *Annual Review of Vision Science*, *2*(Volume 2, 2016), 255–271. https://doi.org/10.1146/annurev-vision-111815-114550

Owsley, C., Sekuler, R., & Siemsen, D. (1983). Contrast sensitivity throughout adulthood. *Vision Research*, *23*(7), 689–699. https://doi.org/10.1016/0042-6989(83)90210-9

Owsley, C., & Sloane, M. E. (1987). Contrast sensitivity, acuity, and the perception of “real-world” targets. *British Journal of Ophthalmology*, *71*(10), 791–796. https://doi.org/10.1136/bjo.71.10.791

Poland, E., Donner, T. H., Müller, K.-M., Leopold, D. A., & Wilke, M. (2019). Thalamus exhibits less sensory variability quenching than cortex. *Scientific Reports*, *9*(1), Article 1. https://doi.org/10.1038/s41598-019-43934-9

Randall, A. D., Booth, C., & Brown, J. T. (2012). Age-related changes to Na+ channel gating contribute to modified intrinsic neuronal excitability. *Neurobiology of Aging*, *33*(11), 2715–2720. https://doi.org/10.1016/j.neurobiolaging.2011.12.030

Ribeiro, M., & Castelo-Branco, M. (2022). 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. https://doi.org/10.7554/eLife.75722

Rosenzweig, E. S., & Barnes, C. A. (2003). Impact of aging on hippocampal function:

Plasticity, network dynamics, and cognition. *Progress in Neurobiology*, *69*(3), 143–179. https://doi.org/10.1016/s0301-0082(02)00126-0

Sabel, B. A., & Stein, D. G. (1981). Extensive loss of subcortical neurons in the aging rat brain. *Experimental Neurology*, *73*(2), 507–516. https://doi.org/10.1016/0014-4886(81)90284-3

Salthouse, T. A., & Lichty, W. (1985). Tests of the Neural Noise Hypothesis of Age-Related Cognitive Change. *Journal of Gerontology*, *40*(4), 443–450. https://doi.org/10.1093/geronj/40.4.443

Samanez-Larkin, G. R., Kuhnen, C. M., Yoo, D. J., & Knutson, B. (2010). Variability in Nucleus Accumbens Activity Mediates Age-Related Suboptimal Financial Risk Taking. *The Journal of Neuroscience*, *30*(4), 1426–1434. https://doi.org/10.1523/JNEUROSCI.4902-09.2010

Schmidt, F., Danböck, S. K., Trinka, E., Klein, D. P., Demarchi, G., & Weisz, N. (2024). Age-related changes in “cortical” 1/f dynamics are linked to cardiac activity. *eLife*, *13*. https://doi.org/10.7554/eLife.100605.1

Schmolesky, M. T., Wang, Y., Pu, M., & Leventhal, A. G. (2000). Degradation of stimulus selectivity of visual cortical cells in senescent rhesus monkeys. *Nature Neuroscience*, *3*(4), Article 4. https://doi.org/10.1038/73957

Schölvinck, M. L., Saleem, A. B., Benucci, A., Harris, K. D., & Carandini, M. (2015). Cortical State Determines Global Variability and Correlations in Visual Cortex. *The Journal of Neuroscience*, *35*(1), 170–178. https://doi.org/10.1523/JNEUROSCI.4994-13.2015

Simkin, D., Hattori, S., Ybarra, N., Musial, T. F., Buss, E. W., Richter, H., Oh, M. M., Nicholson, D. A., & Disterhoft, J. F. (2015). Aging-Related Hyperexcitability in CA3 Pyramidal Neurons Is Mediated by Enhanced A-Type K+ Channel Function and Expression. *Journal of Neuroscience*, *35*(38), 13206–13218. https://doi.org/10.1523/JNEUROSCI.0193-15.2015

Sosulina, L., Mittag, M., Geis, H.-R., Hoffmann, K., Klyubin, I., Qi, Y., Steffen, J., Friedrichs, D., Henneberg, N., Fuhrmann, F., Justus, D., Keppler, K., Cuello, A. C., Rowan, M.

J., Fuhrmann, M., & Remy, S. (2021). Hippocampal hyperactivity in a rat model of Alzheimer's disease. *Journal of Neurochemistry*, *157*(6), 2128–2144. https://doi.org/10.1111/jnc.15323

Stranahan, A. M., Jiam, N. T., Spiegel, A. M., & Gallagher, M. (2012). Aging Reduces Total Neuron Number in the Dorsal Component of the Rodent Prefrontal Cortex. *The Journal of Comparative Neurology*, *520*(6), 1318–1326. https://doi.org/10.1002/cne.22790

Tolhurst, D. J., Movshon, J. A., & Dean, A. F. (1983). The statistical reliability of signals in single neurons in cat and monkey visual cortex. *Vision Research*, *23*(8), 775–785. https://doi.org/10.1016/0042-6989(83)90200-6

Tolhurst, D. J., Movshon, J. A., & Thompson, I. D. (1981). The dependence of response amplitude and variance of cat visual cortical neurones on stimulus contrast. *Experimental Brain Research*, *41*(3–4), 414–419. https://doi.org/10.1007/BF00238900

Tomko, G. J., & Crapper, D. R. (1974). Neuronal variability: Non-stationary responses to identical visual stimuli. *Brain Research*, *79*(3), 405–418. https://doi.org/10.1016/0006-8993(74)90438-7

Tran, T. T., Rolle, C. E., Gazzaley, A., & Voytek, B. (2020). Linked sources of neural noise contribute to age-related cognitive decline. *Journal of Cognitive Neuroscience*, *32*(9), 1813–1822. https://doi.org/10.1162/jocn_a_01584

Turner, J. G., Hughes, L. F., & Caspary, D. M. (2005). Affects of Aging on Receptive Fields in Rat Primary Auditory Cortex Layer V Neurons. *Journal of Neurophysiology*, *94*(4), 2738–2747. https://doi.org/10.1152/jn.00362.2005

Urai, A. E., Doiron, B., Leifer, A. M., & Churchland, A. K. (2022). Large-scale neural recordings call for new insights to link brain and behavior. *Nature Neuroscience*, *25*(1), Article 1. https://doi.org/10.1038/s41593-021-00980-9

Vallat, R. (2018). Pingouin: Statistics in Python. *Journal of Open Source Software*, *3*(31), 1026. https://doi.org/10.21105/joss.01026

Voytek, B., Kramer, M. A., Case, J., Lepage, K. Q., Tempesta, Z. R., Knight, R. T., & Gazzaley, A. (2015). Age-Related Changes in 1/f Neural Electrophysiological Noise. *Journal of Neuroscience*, *35*(38), 13257–13265. https://doi.org/10.1523/JNEUROSCI.2332-14.2015

Wang, H.-L., Xian, X.-H., Song, Q.-Y., Pang, C., Wang, J.-L., Wang, M.-W., & Li, W.-B. (2017). Age-related alterations of neuronal excitability and voltage-dependent Ca2+ current in a spontaneous mouse model of Alzheimer's disease. *Behavioural Brain Research*, *321*, 209–213. https://doi.org/10.1016/j.bbr.2017.01.009

Wang, M., Gamo, N. J., Yang, Y., Jin, L. E., Wang, X.-J., Laubach, M., Mazer, J. A., Lee, D., & Arnsten, A. F. T. (2011). Neuronal basis of age-related working memory decline. *Nature*, *476*(7359), Article 7359. https://doi.org/10.1038/nature10243

Wang, X., Zhang, B., Wang, H., Liu, J., Xu, G., & Zhou, Y. (2019). Aging affects correlation within the V1 neuronal population in rhesus monkeys. *Neurobiology of Aging*, *76*, 1–8. https://doi.org/10.1016/j.neurobiolaging.2018.11.025

Waschke, L., Wöstmann, M., & Obleser, J. (2017). States and traits of neural irregularity in the age-varying human brain. *Scientific Reports*, *7*(1), 17381. https://doi.org/10.1038/s41598-017-17766-4

Welford, A. T. (1981). Signal, Noise, Performance, and Age. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *23*(1), 97–109. https://doi.org/10.1177/001872088102300109

Wen, W., Grover, S., Hazel, D., Berning, P., Baumgardt, F., Viswanathan, V., Tween, O., & Reinhart, R. M. G. (2024). Beta-band neural variability reveals age-related dissociations in human working memory maintenance and deletion. *PLOS Biology*, *22*(9), e3002784. https://doi.org/10.1371/journal.pbio.3002784

Wiltschko, A. B., Johnson, M. J., Iurilli, G., Peterson, R. E., Katon, J. M., Pashkovski, S. L., Abraira, V. E., Adams, R. P., & Datta, S. R. (2015). Mapping Sub-Second Structure in Mouse Behavior. *Neuron*, *88*(6), 1121–1135. https://doi.org/10.1016/j.neuron.2015.11.031

Wiltschko, A. B., Tsukahara, T., Zeine, A., Anyoha, R., Gillis, W. F., Markowitz, J. E., Peterson, R. E., Katon, J., Johnson, M. J., & Datta, S. R. (2020). Revealing the structure of pharmacobehavioral space through motion sequencing. *Nature Neuroscience*, *23*(11), 1433–1443. https://doi.org/10.1038/s41593-020-00706-3

Wu, K., & Gollo, L. L. (2024). *Mapping and Modeling Age-Related Changes in Intrinsic Neural Timescales* (p. 2024.09.10.612380). bioRxiv. https://doi.org/10.1101/2024.09.10.612380

Yang, Y., Liang, Z., Li, G., Wang, Y., & Zhou, Y. (2009). Aging affects response variability of V1 and MT neurons in rhesus monkeys. *Brain Research*, *1274*, 21–27. https://doi.org/10.1016/j.brainres.2009.04.015

Yates, M. A., Markham, J. A., Anderson, S. E., Morris, J. R., & Juraska, J. M. (2008). Regional variability in age-related loss of neurons from the primary visual cortex and medial prefrontal cortex of male and female rats. *Brain Research*, *1218*, 1–12. https://doi.org/10.1016/j.brainres.2008.04.055

Yin, C., Melin, M. D., Rojas-Bowe, G., Sun, X. R., Couto, J., Gluf, S., Kostiuk, A., Musall, S., & Churchland, A. K. (2025). Spontaneous movements and their relationship to neural activity fluctuate with latent engagement states. *Neuron*, *113*(18), 3048-3063.e5. https://doi.org/10.1016/j.neuron.2025.06.001

Supplementary figures

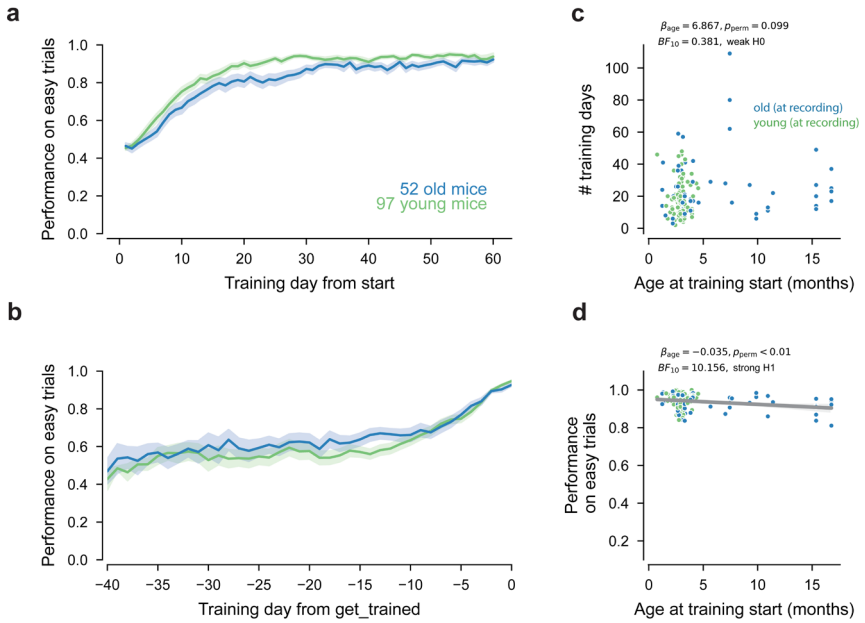


Figure 1 S1. Training progression in the behavioral task. (a) Learning curves from the start of training. Age group split was based on the age at recording, following the same convention as in other main figures. **(b)** Training curves aligned to the day animals reached the behavioral criterion ('get_trained', see the training protocol for definition; International Brain Laboratory et al., 2021). **(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. The training criterion was identical for all mice. Observed differences reflect performance above this threshold within criterion-satisfying sessions on that day.

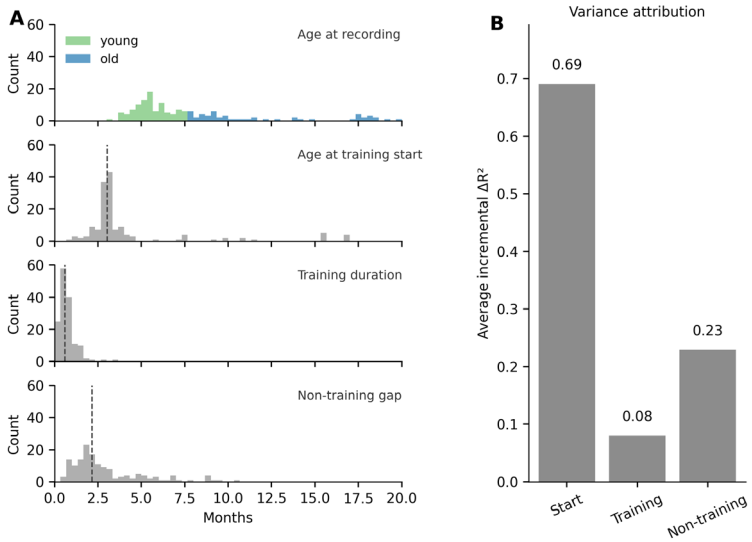


Figure 1 S2. **Components contributing to variability in age at recording. (a)** Distributions (in months) of age at recording and the three components used to decompose age at recording: age at training start, training duration (number of training days), and the non-training gap between training start and the recording. Dash lines indicate the median. **(b)** Order-invariant variance attribution using average incremental ΔR^2 across all predictor orderings in an OLS model.

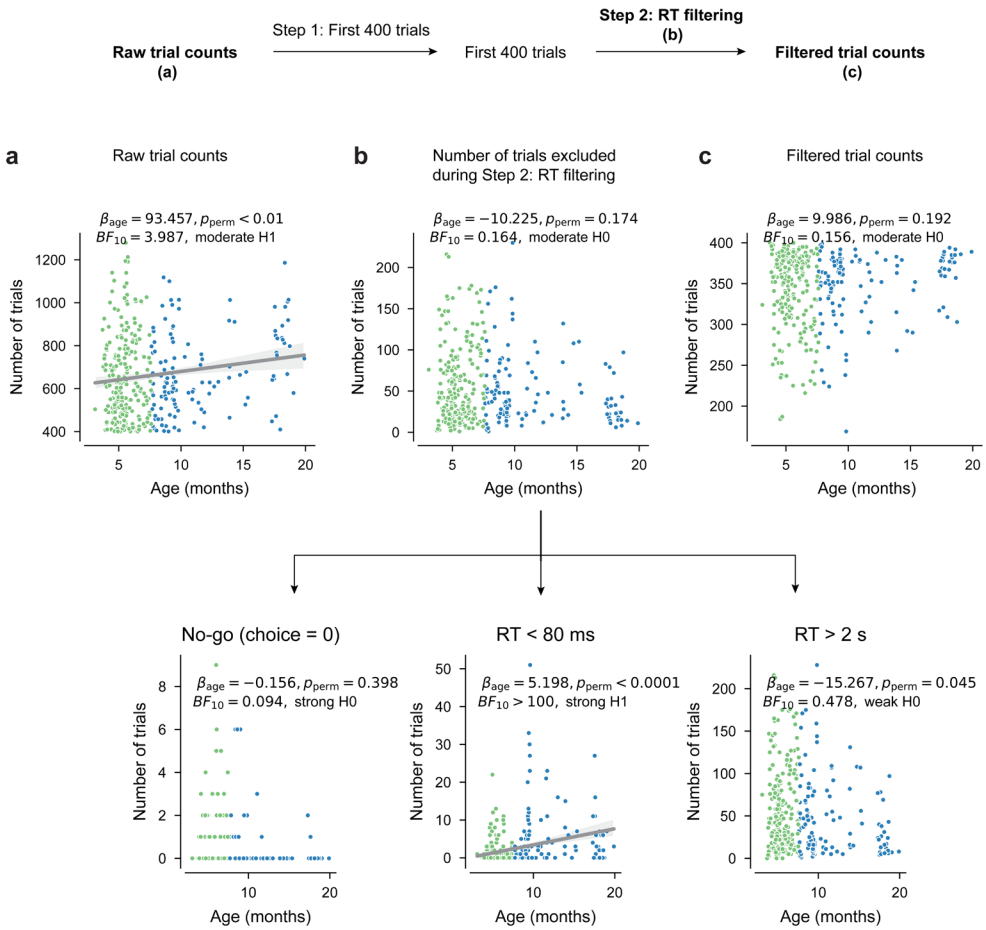


Figure 1 S3. Older mice performed slightly more trials in the standardized task. (a) Number of trials as a function of age. Each dot represents one session. **(b)** Number of trials excluded by RT filtering (no-go trials, RT < 80ms trials, and RT > 2s trials). **(c)** Number of trials included as a function of mouse age.

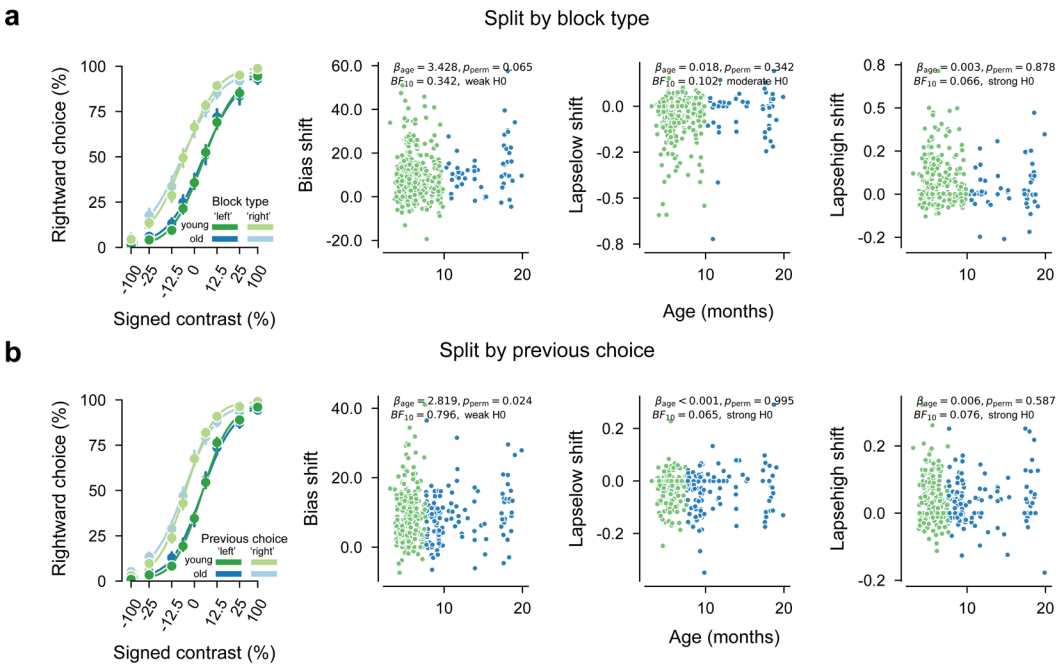


Figure 1 S4. No age-related differences in block- or history-dependent choice bias. （a）Trials were split by block type (see Methods), and shifts in psychometric parameters (bias and lapse rate) were quantified between the two block conditions. No age-related effects were observed for bias shift, low-lapse 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 2 types of previous choices. Again, no age-related effects were observed for bias shift, low-lapse shift, or high-lapse shift.

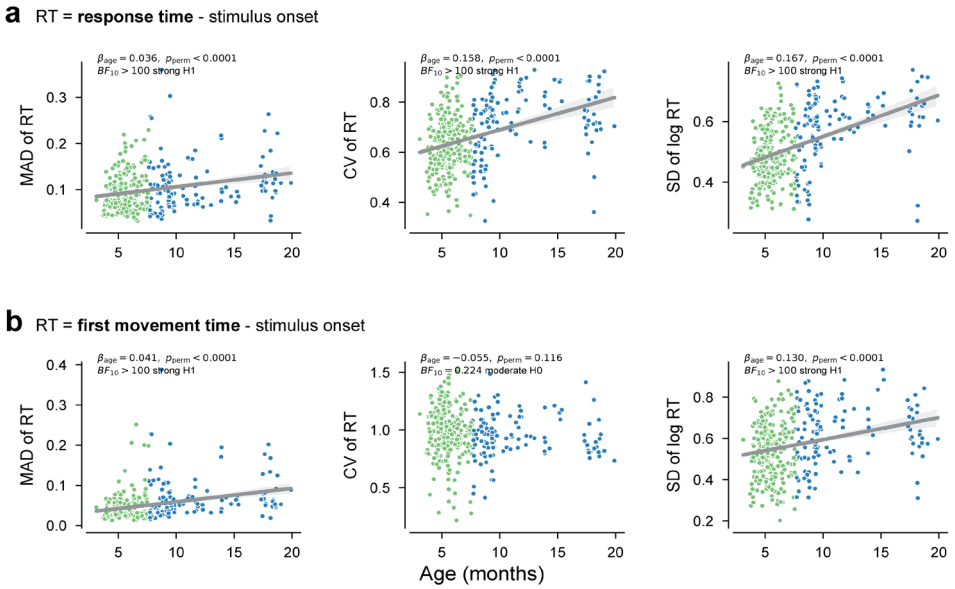


Figure 1 S5. **RT variability measures across two RT definitions and their relationship with age**. **(a)** Response Time was defined as response time minus stimulus onset. We examined alternative measures of RT variability that capture different aspects of dispersion. Specifically, we used the median absolute deviation (MAD), a robust metric less sensitive to skew, 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). Each dot represents one session. **(b)**

Same as (a), but for movement initiation time, defined as the interval between stimulus onset and the first movement (International Brain Laboratory, Benson, et al., 2025).

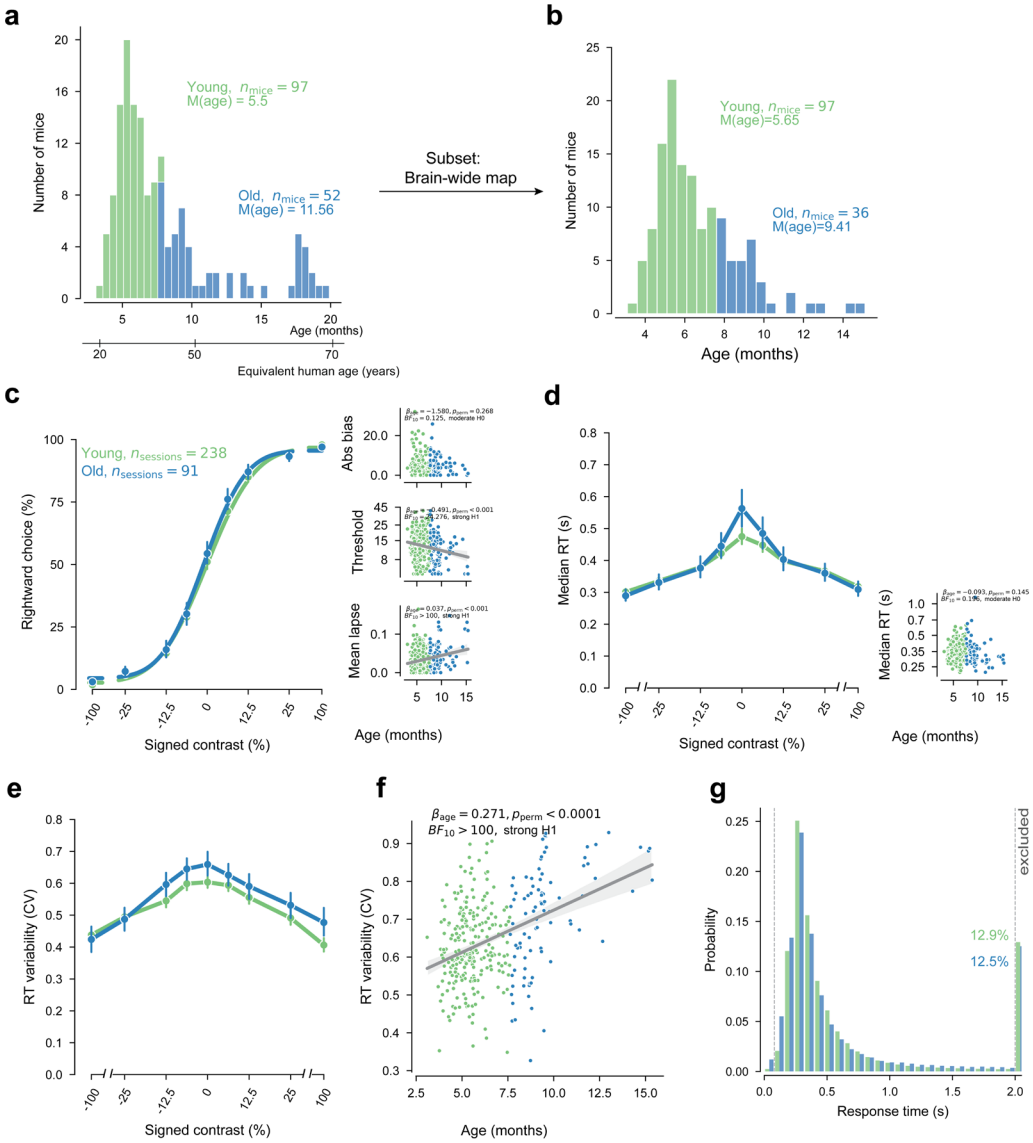


Figure 1 S6. Replication of Figure 1 using only mice from the original IBL "brain-wide map" dataset. Age-related effects on median RT and RT variability are similar in direction and magnitude to those obtained in the full dataset.

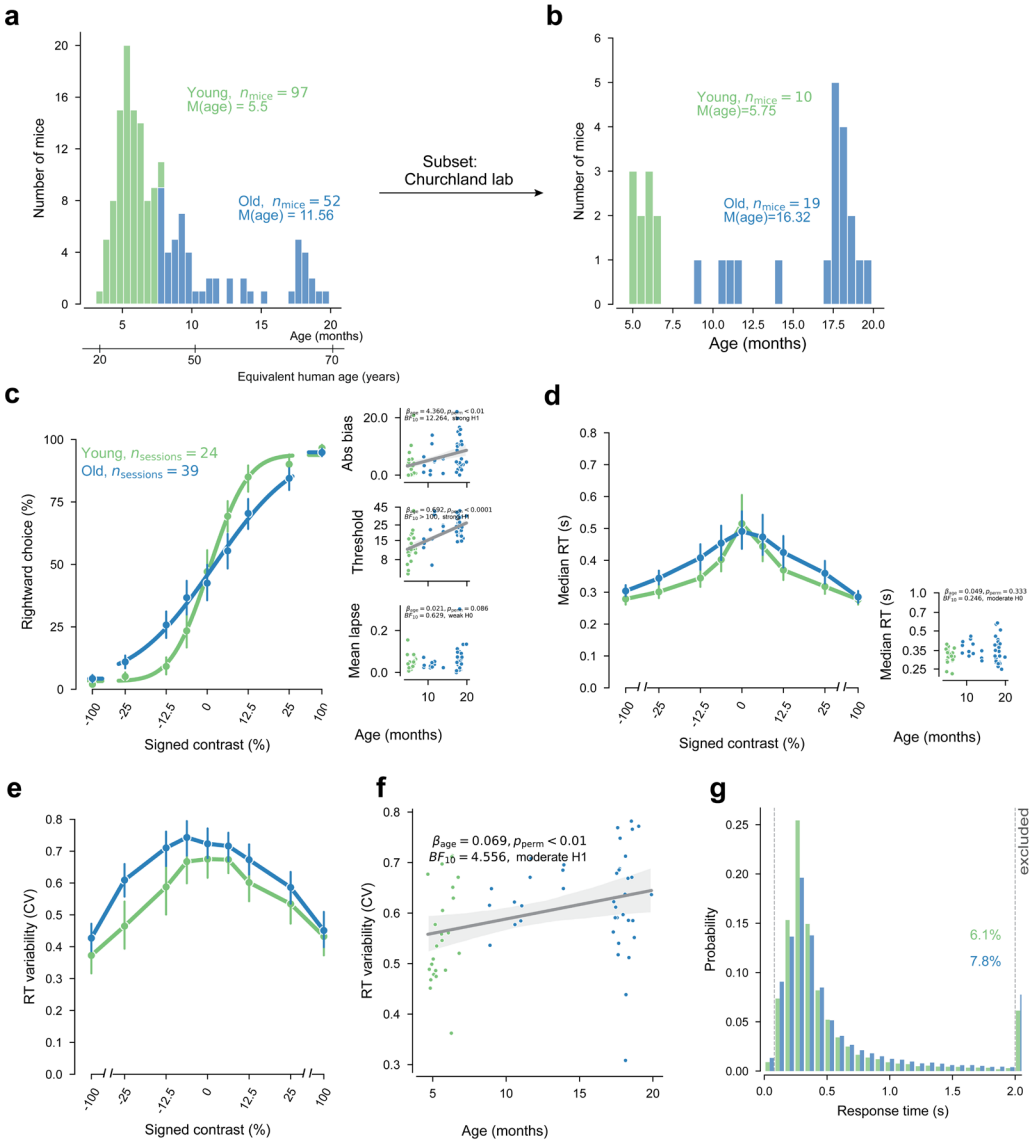


Figure 1 S7. Replication of Figure 1 using only sessions recorded in the Churchland lab.

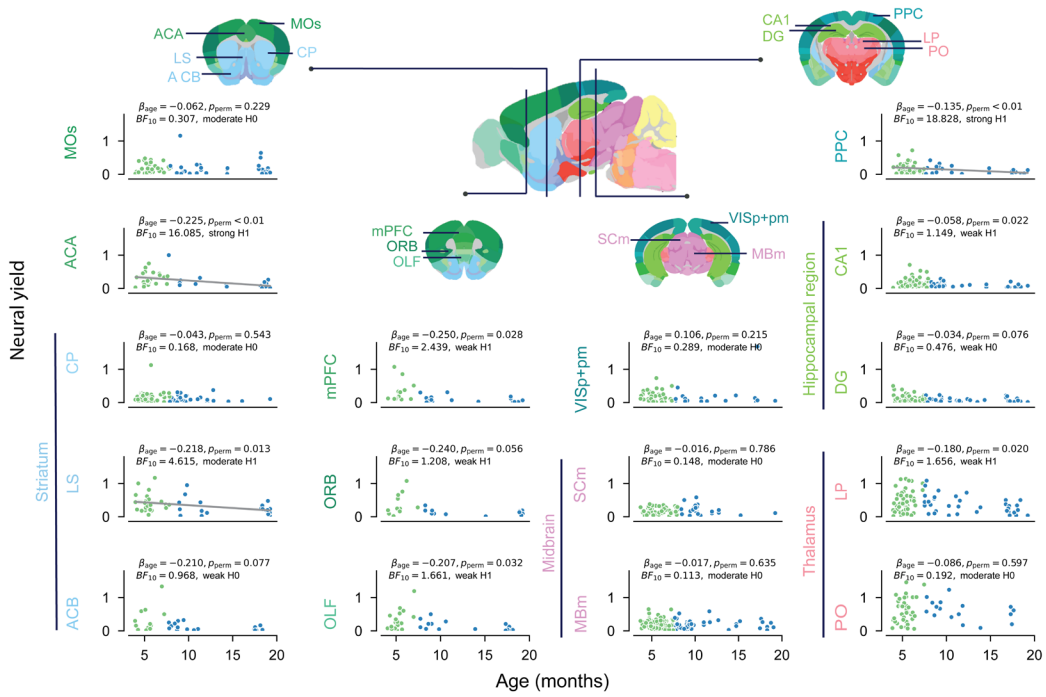


Figure 2 S1. Neural yield slightly decreases with age. Scatterplots show neural yield in each ROI as a function of age. For each panel, the estimated age slope (β), permutation test p-value, and Bayes Factor are indicated in the top left. Each dot corresponds to one recording insertion. Neural yield is defined as the number of good neurons per good channel (along the length of the Neuropixels probe) within a given region. Statistical analyses for neural yield were performed at the insertion level (yield is inherently defined per insertion).

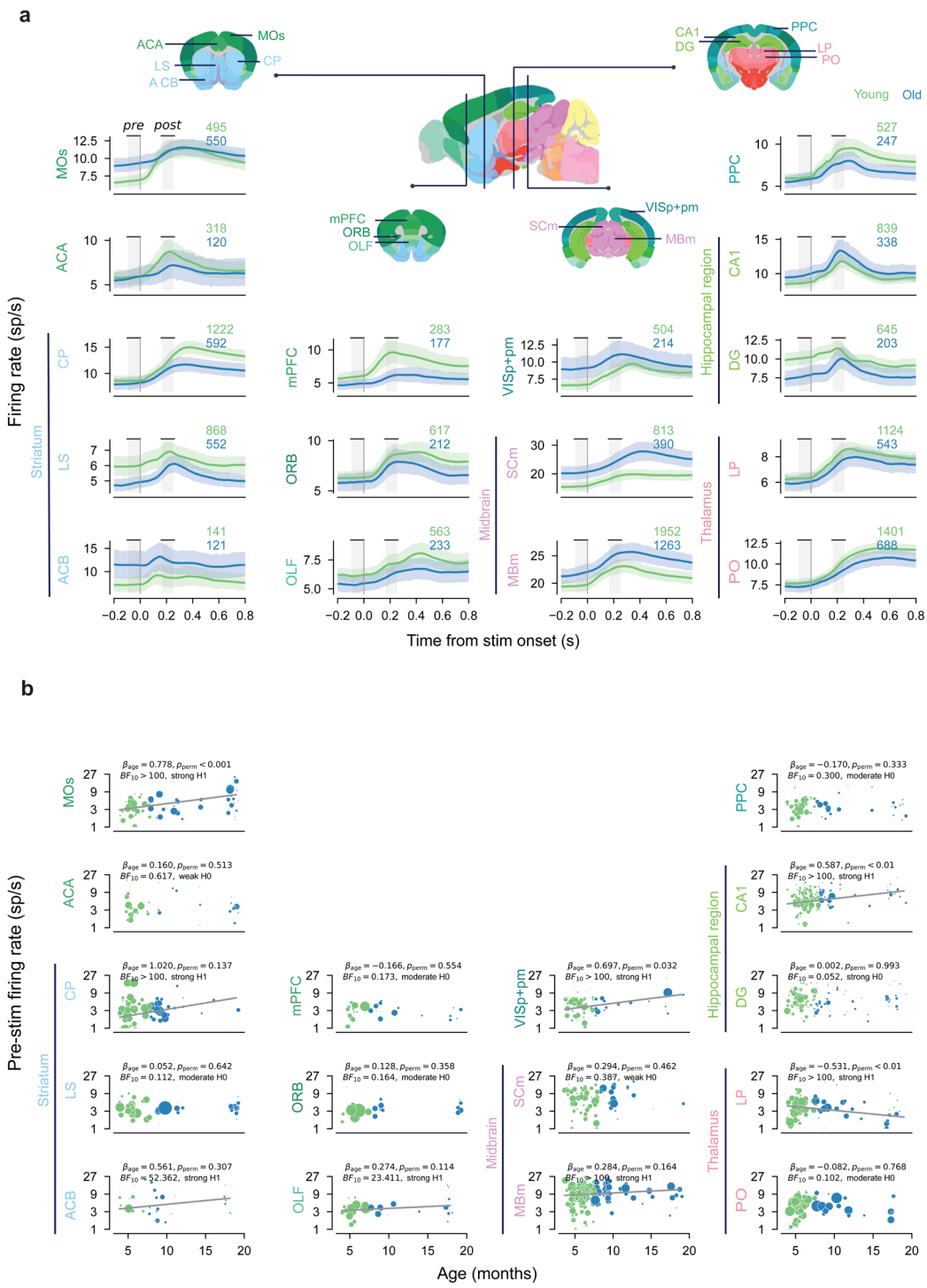


Figure 3 S1. Regional specificity of firing rate time courses and pre-stimulus firing rates. (a) Time courses of mean firing rates in different brain regions, aligned to stimulus onset. **(b)** The relationship between pre-stimulus ([-100ms, 0ms] relative to stimulus onset) firing rate and mouse age. Each dot represents one insertion.

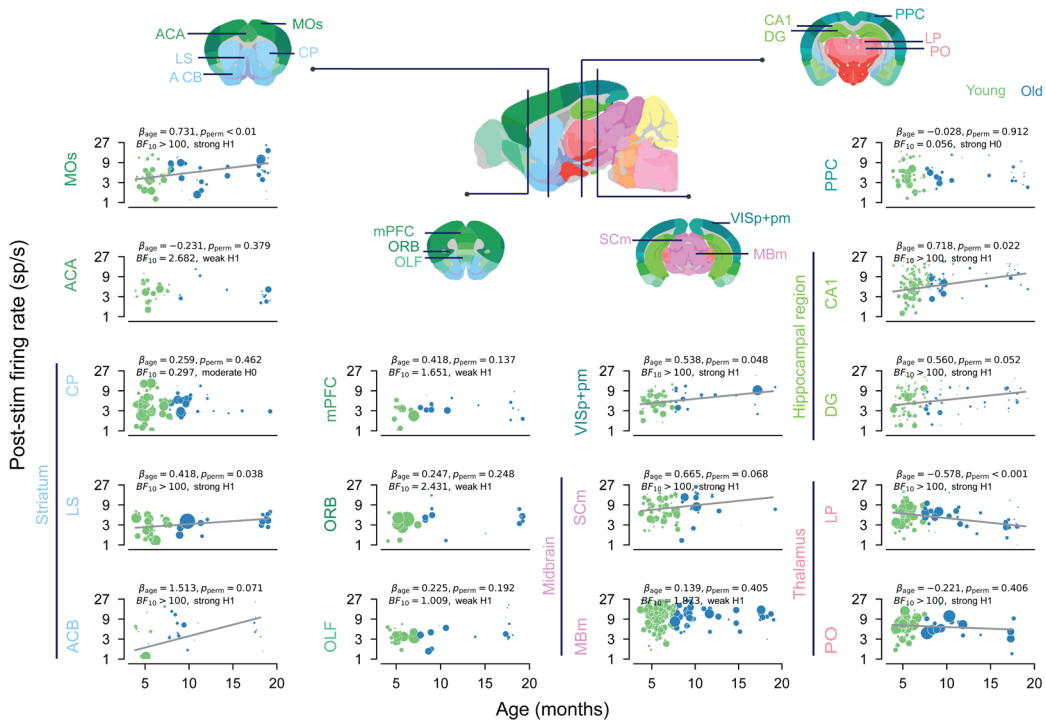


Figure 3 S2. Regional specificity of post-stimulus firing rates. The relationship between post-stimulus firing rate and mouse age. Each dot represents one insertion.

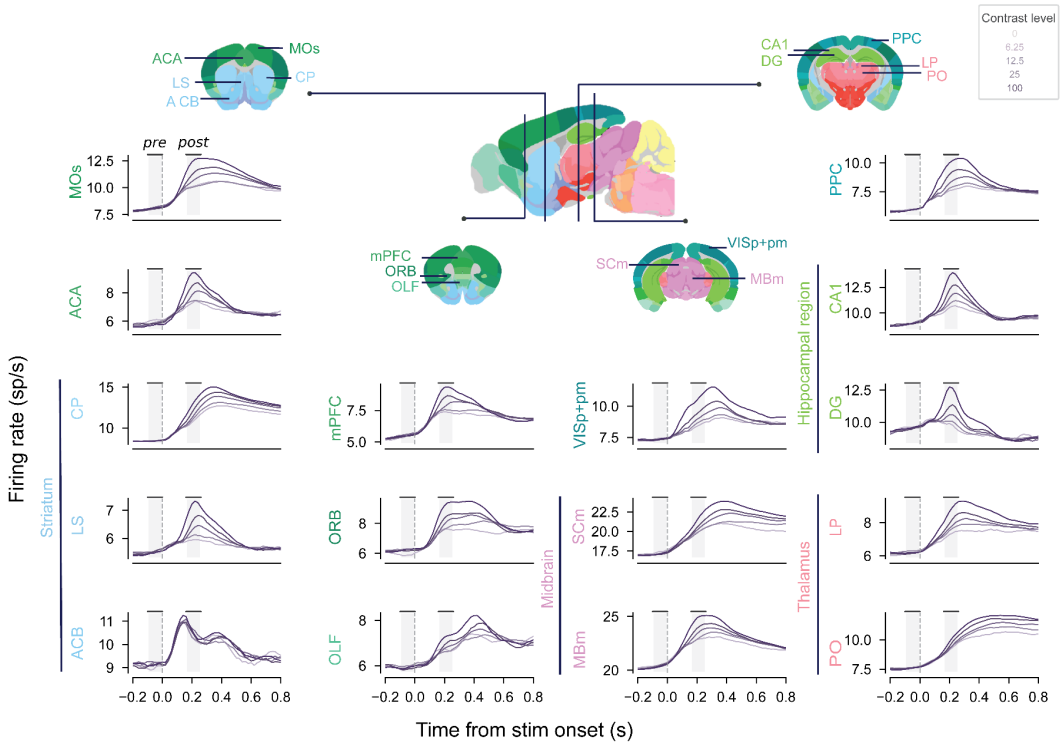


Figure 3 S3. Regional specificity in contrast modulation of firing rate. Time courses of the average firing rate in each ROI across different stimulus contrast levels, combining data from both age groups.

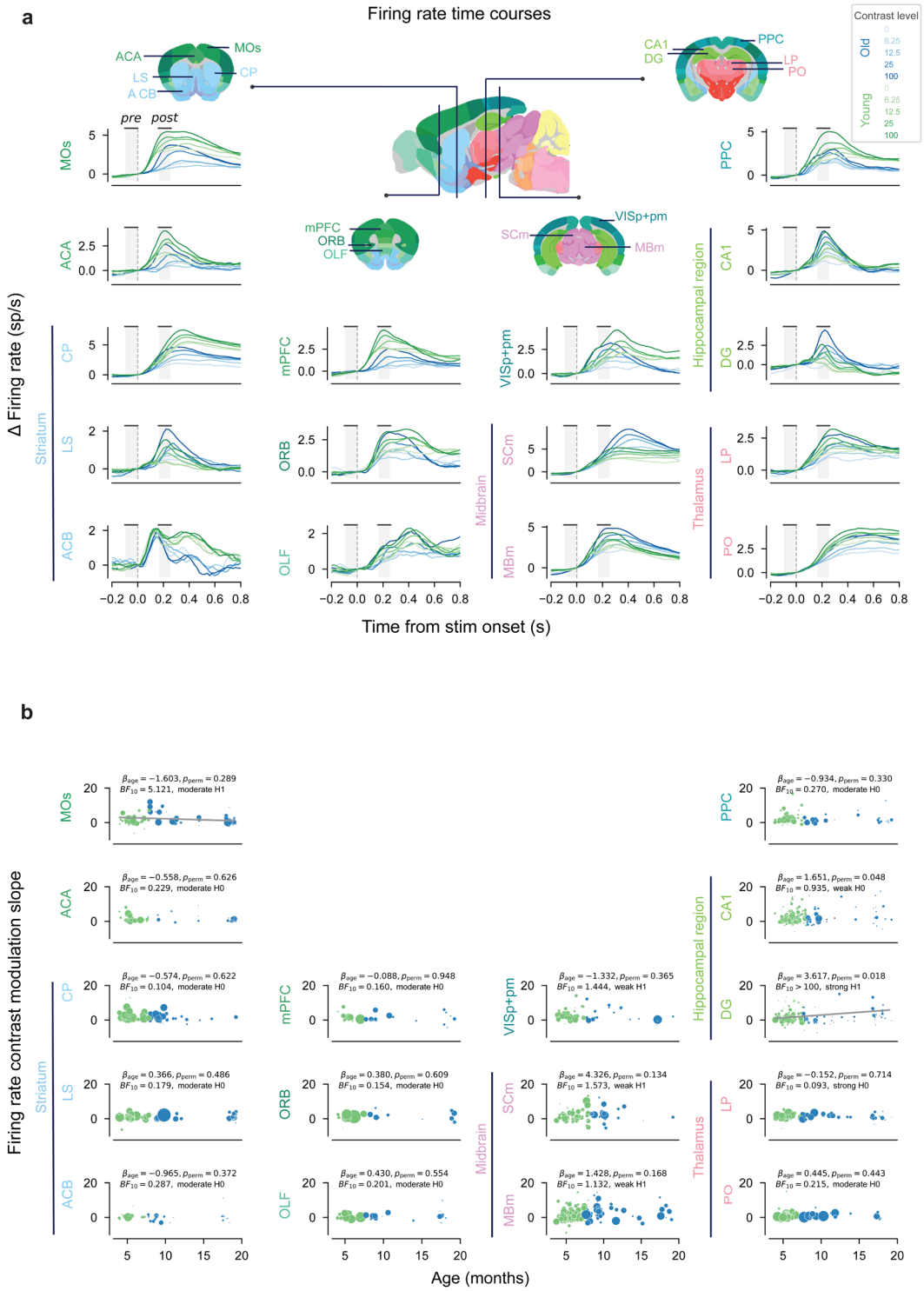


Figure 3 S4. Regional specificity of age-related differences in contrast modulation of firing rates. (a) Baseline-corrected mean firing rate time courses for different stimulus contrast levels in each ROI, aligned to stimulus onset. Blue and green represent different age groups. Saturation represents different contrast levels. **(b)** The relationship between contrast modulation slope and mouse age. Each dot represents one insertion.

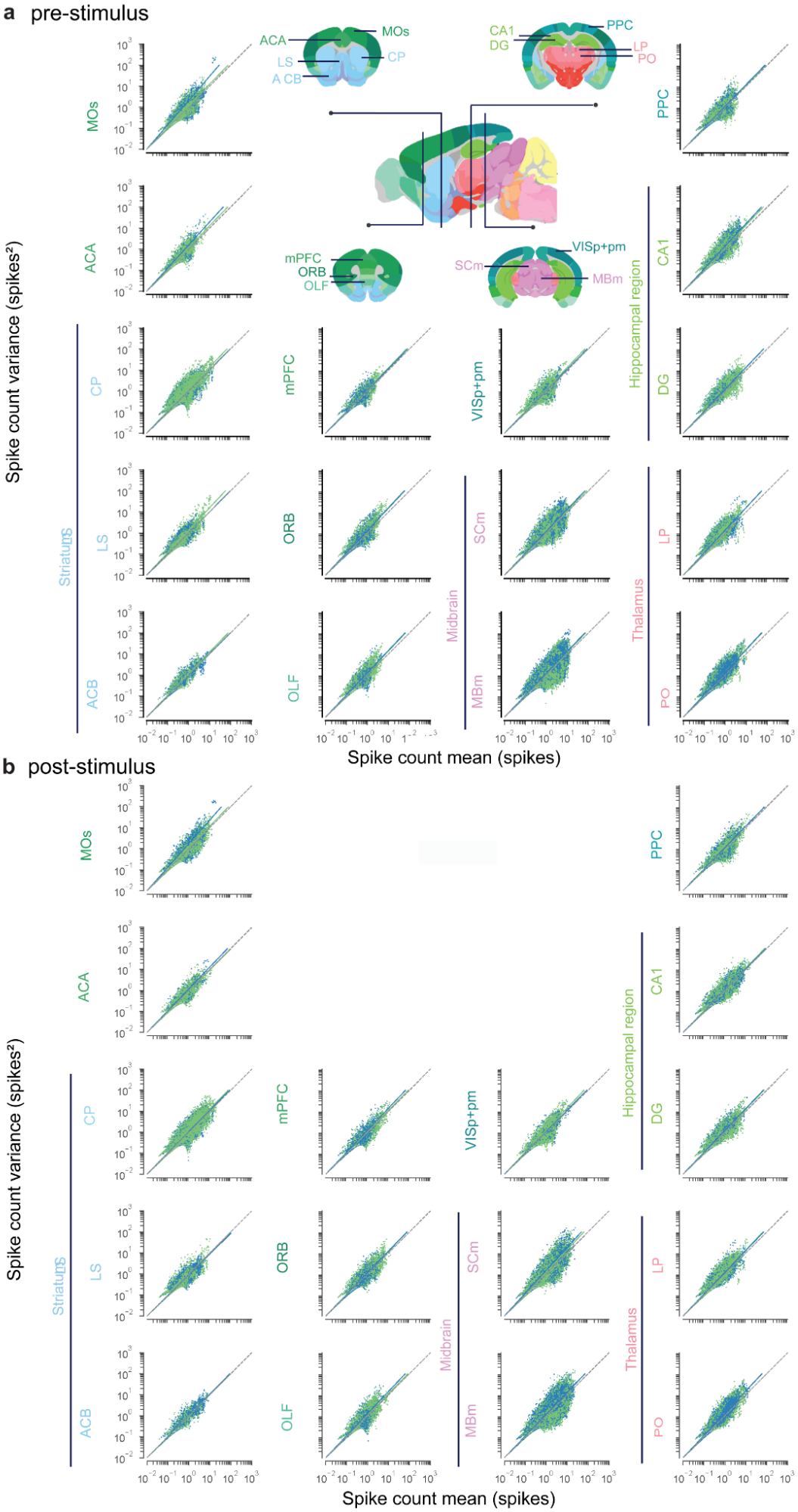

a pre-stimulus
b post-stimulus
Spike count variance (spikes²)
Spike count mean (spikes)
MOs
ACA
CP
LS
ACB
mPFC
ORB
OLF
VISp+pm
SCm
MBm
PPC
CA1
DG
LP
PO
Striatum
Midbrain
Thalamus
Hippocampal region

Figure 4 S1. The relationship between spike count variance and spike count mean. For each neuron and each stimulus level, we plotted its spike count variance (spikes^2) against its spike count mean (spikes) across trials, separately for (a) the pre-stimulus time window (-100ms, 0ms) and (b) the post-stimulus time window (160ms, 260ms). Colored lines represent linear fits for each age group (blue = old, green = young), with the slope corresponding to the Fano Factor. We observed curvilinear patterns in the lower-left part of many panels, resembling patterns found in Figure 2 of Goris et al. (2014). The origin of these patterns remains unclear and warrants further investigation.

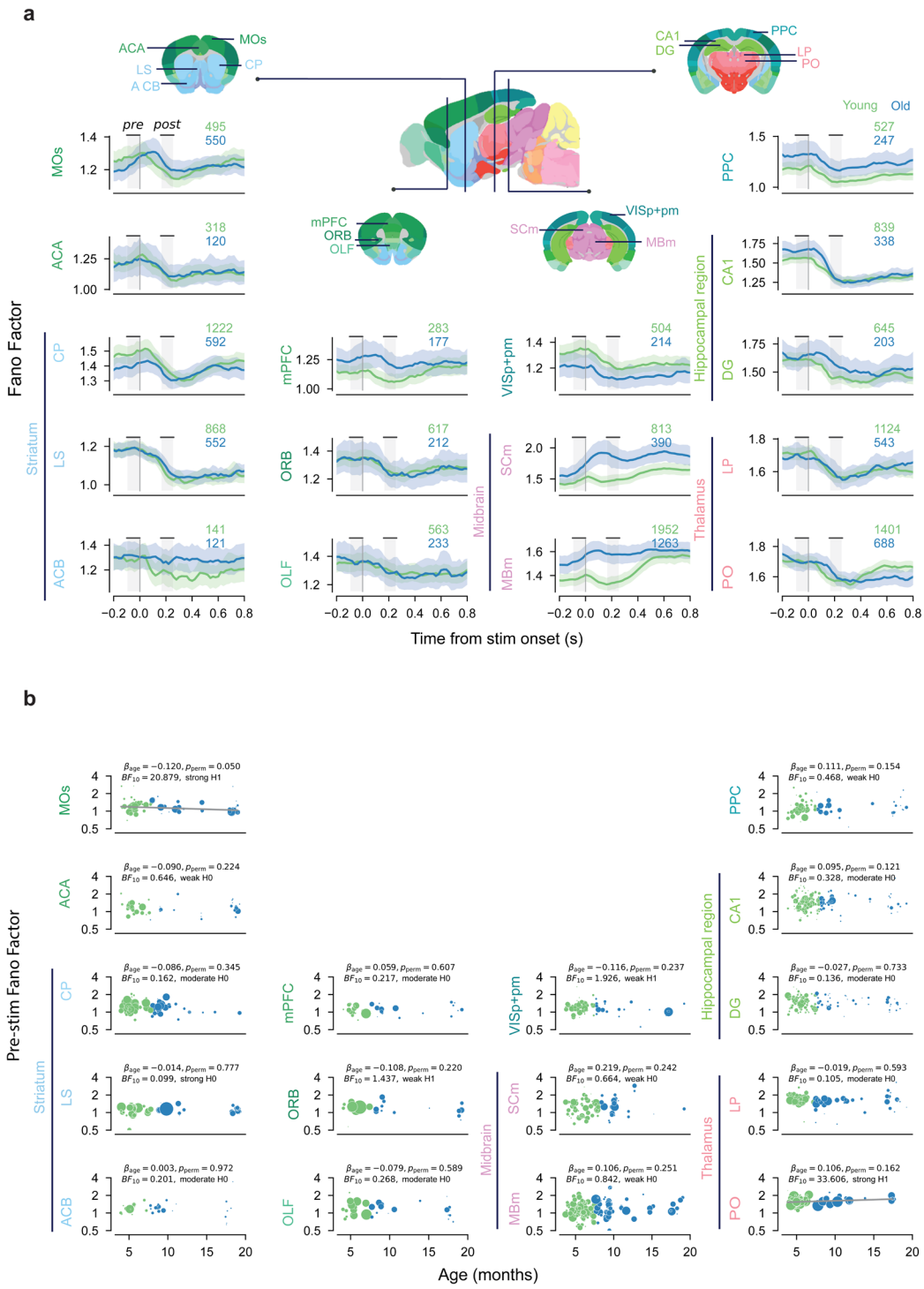


Figure 4 S2a-b. Regional specificity of mean-subtracted Fano Factors. (a) The mean-subtracted Fano Factor time courses for each ROI, aligned to stimulus onset. Colors represent the age groups (blue = old, green = young). Thick lines show the mean Fano Factor across all data points within each age group, with shaded areas representing the 95% confidence interval. 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. Each dot represents one insertion. The size of the dots shows the number of neurons recorded in the insertion.

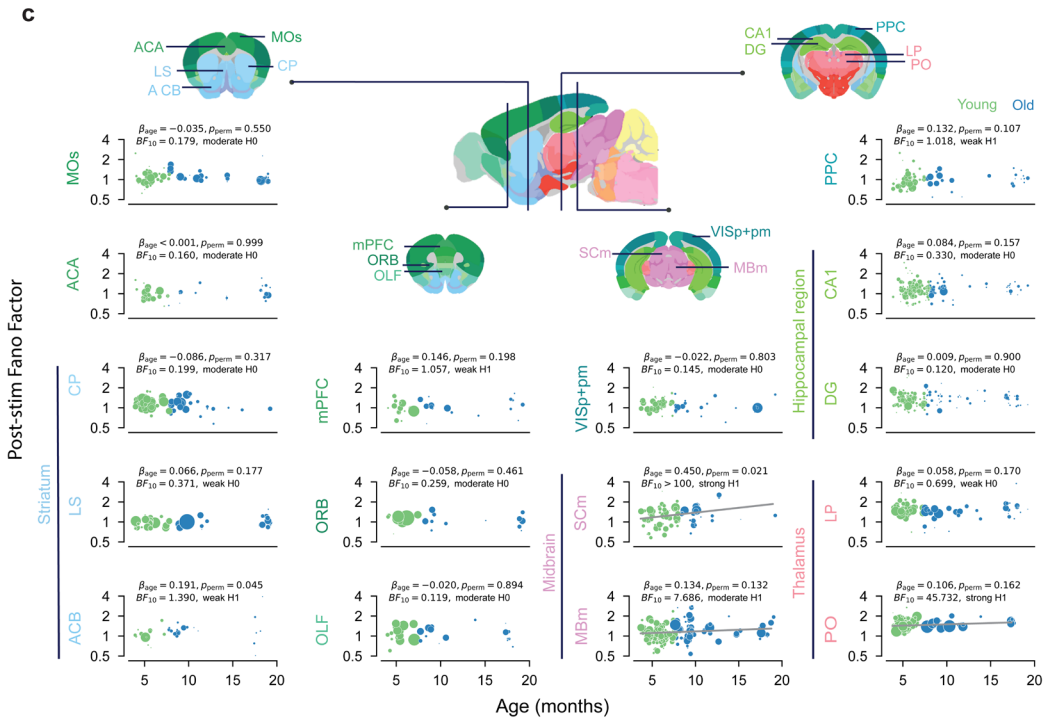


Figure 4 S2c. Regional specificity of post-stimulus mean-subtracted Fano Factors. **(c)**The relationship between post-stimulus mean-subtracted Fano Factor and age. Each dot represents one insertion. The size of the dots shows the number of neurons recorded in the insertion.

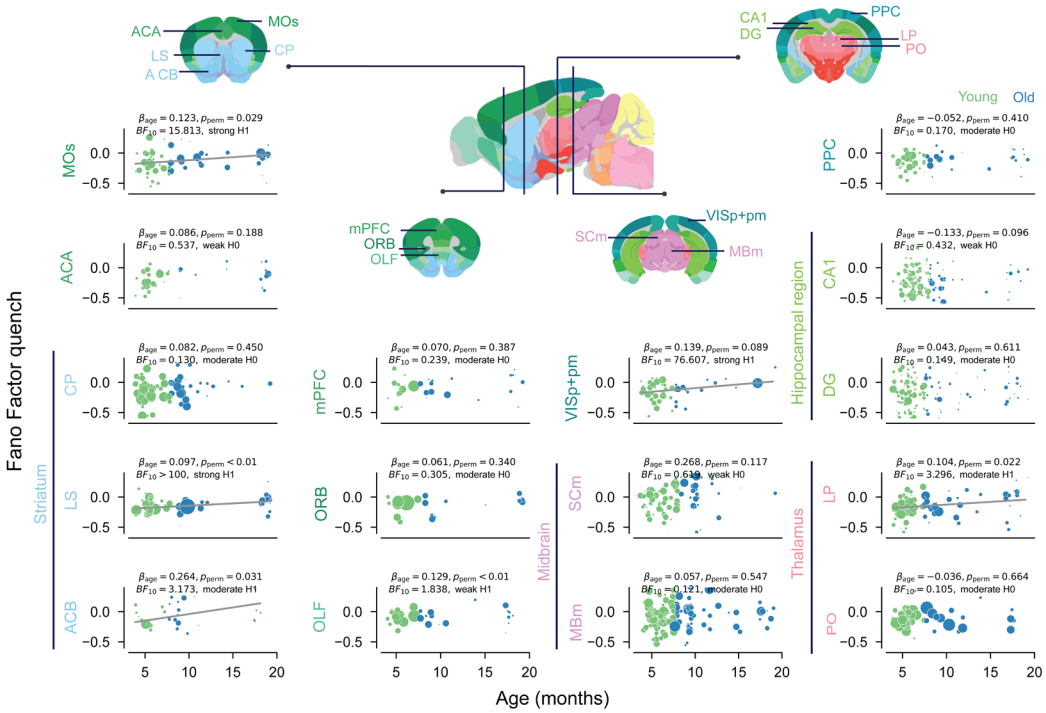


Figure 4 S3. Regional specificity of mean-subtracted Fano Factor quenching. The relationship between the Fano Factor quench index (calculated as post-stimulus Fano Factor minus pre-stimulus Fano Factor for each neuron) and the age of the mice. Each dot represents one insertion. The size of the dots shows the number of neurons recorded in the insertion.

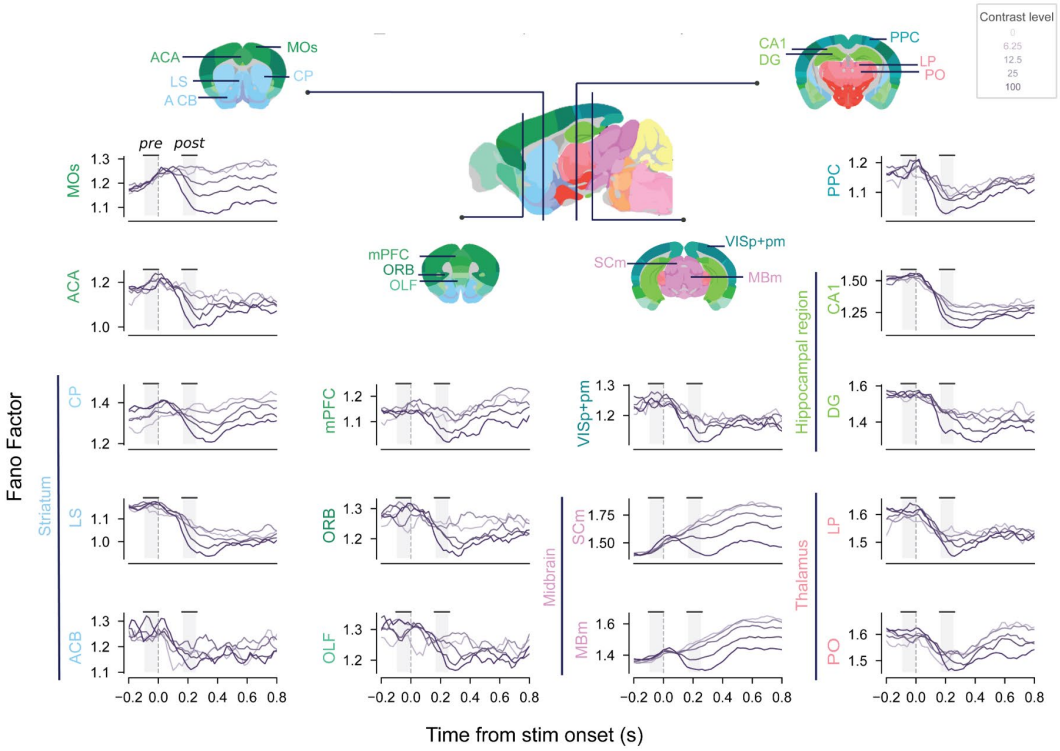


Figure 4 S4. Regional specificity in contrast modulation of the Fano Factor. Time courses of the mean Fano Factor in each ROI across different stimulus contrast levels, combining data from both age groups.

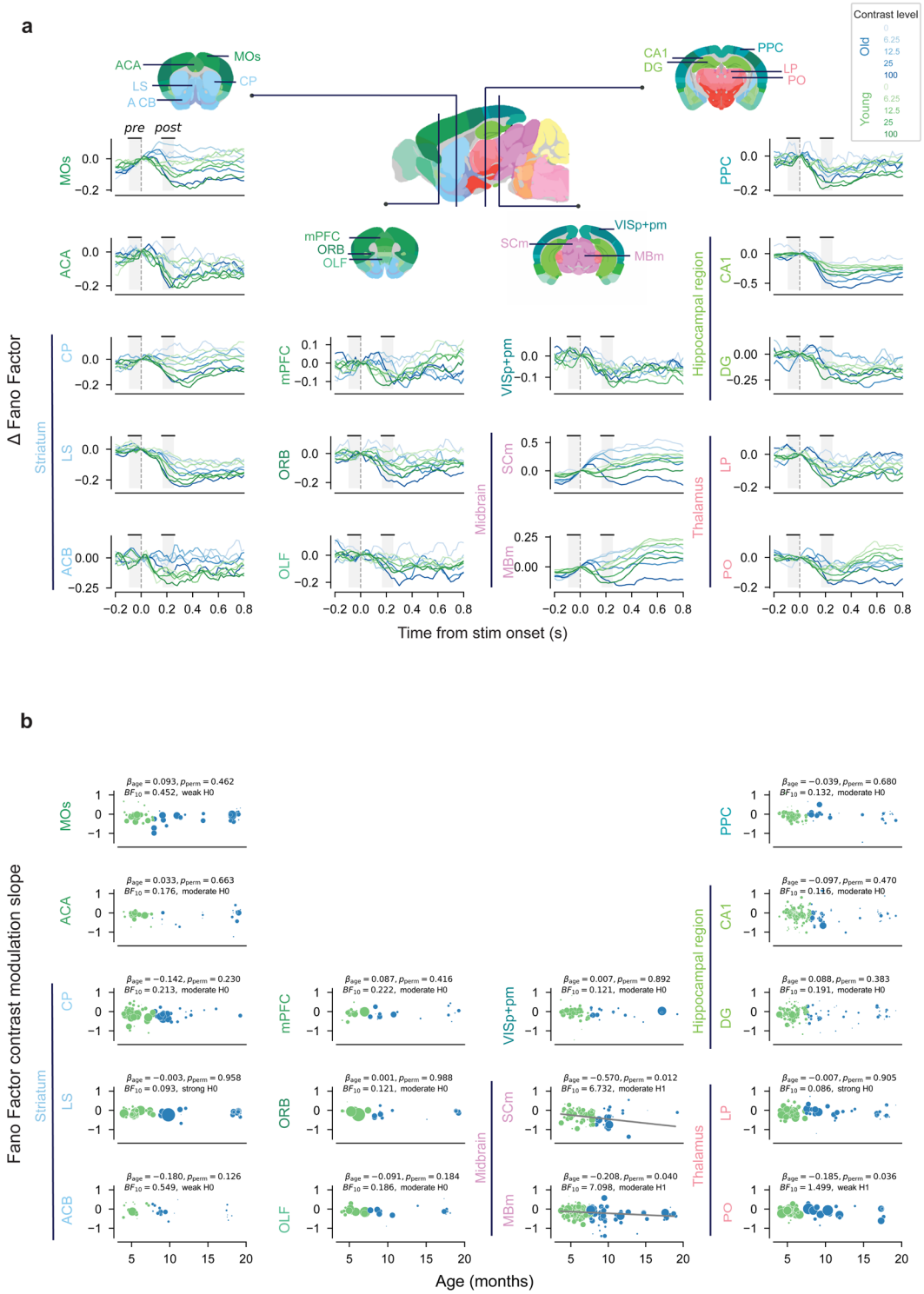


Figure 4 S5 Regional specificity in age-related differences in contrast modulation of the Fano Factor. (a) Baseline corrected mean Fano Factor time courses for different stimulus contrast levels in each ROI, aligned to stimulus onset. Colors represent the age groups (blue = old, green = young). Saturation represents different contrast levels. Sliding window width: 0.1 s; step size: 0.02 s. Grey areas indicate the pre-stimulus (-100ms, 0ms) and post-stimulus (160ms, 260ms) time windows separately. **(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. Each dot represents one insertion.

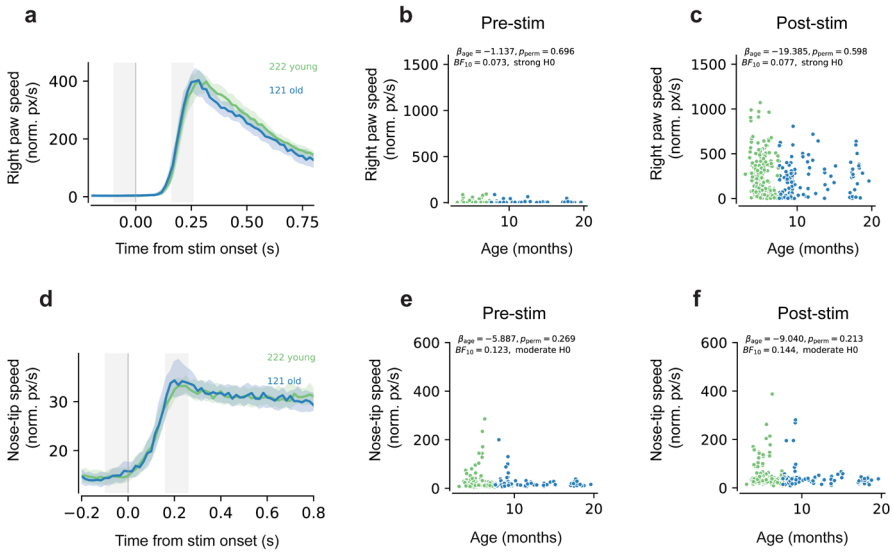


Figure 4 S6. Right-camera DLC movement time courses and session-level summary metrics. (a) Right paw speed time course aligned to stimulus onset (t = 0), summarized across sessions (median ± 95% CI). Gray shaded regions indicate the pre-stimulus (−0.1 to 0 s) and post-stimulus (0.16 to 0.26 s) windows 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% CI), 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 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

Table S1. 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

Table S2. 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

Table S3. 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 session-level 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.