







# Coherent mapping of position and head direction across auditory and visual cortex

Paul E.C. Mertens <sup>1</sup>, Pietro Marchesi <sup>1</sup>, Thijs R. Ruikes <sup>1</sup>, Matthijs Oude Lohuis <sup>1,2</sup>, Quincy Krijger<sup>1</sup>,  
Cyril M.A. Pennartz <sup>1,3,\*</sup>, Carien S. Lansink <sup>1,3</sup>

<sup>1</sup>Center for Neuroscience, Faculty of Science, Swammerdam Institute for Life Sciences, University of Amsterdam, Science Park 904, Amsterdam 1098 XH, The Netherlands,

<sup>2</sup>Champalimaud Neuroscience Programme, Champalimaud Foundation, Lisbon, Portugal,

<sup>3</sup>Research Priority Program Brain and Cognition, University of Amsterdam, Science Park 904, Amsterdam 1098 XH, The Netherlands

\*Corresponding author: Center for Neuroscience, Faculty of Science, Swammerdam Institute for Life Sciences, University of Amsterdam, Science Park 904, Amsterdam 1098 XH, The Netherlands. Email: c.m.a.pennartz@uva.nl

Neurons in primary visual cortex (V1) may not only signal current visual input but also relevant contextual information such as reward expectancy and the subject's spatial position. Such contextual representations need not be restricted to V1 but could participate in a coherent mapping throughout sensory cortices. Here, we show that spiking activity coherently represents a location-specific mapping across auditory cortex (AC) and lateral, secondary visual cortex (V2L) of freely moving rats engaged in a sensory detection task on a figure-8 maze. Single-unit activity of both areas showed extensive similarities in terms of spatial distribution, reliability, and position coding. Importantly, reconstructions of subject position based on spiking activity displayed decoding errors that were correlated between areas. Additionally, we found that head direction, but not locomotor speed or head angular velocity, was an important determinant of activity in AC and V2L. By contrast, variables related to the sensory task cues or to trial correctness and reward were not markedly encoded in AC and V2L. We conclude that sensory cortices participate in coherent, multimodal representations of the subject's sensory-specific location. These may provide a common reference frame for distributed cortical sensory and motor processes and may support crossmodal predictive processing.

**Key words:** ensemble recordings; auditory cortex; secondary lateral visual cortex; spatial mapping; head direction.

## Introduction

Early sensory cortical areas were long viewed to primarily function as collections of unisensory feature detectors (Hubel and Wiesel 1962; Felleman and Van Essen 1991; DiCarlo and Cox 2007; Miller 2016). More recently, single-unit recordings in awake, behaving animals have shown responses in the auditory cortex (AC) and visual cortex (primary visual cortex [V1]) to a wide variety of perceptual and behavioral factors, suggesting these areas have functions beyond unimodal sensory processing. In rodent V1, these include responses to reward and reward timing (Shuler and Bear 2006), reward predictive stimuli (Goltstein et al. 2013), running speed (Niell and Stryker 2010; Ayaz et al. 2013), head-orienting movements (Guitchoyants et al. 2020), orofacial motor activity (Stringer et al. 2019), and responses which are causal to visually cued action timing (Namboodiri et al. 2015). Additionally, a growing number of studies show that many V1 neurons display location-correlated spiking activity (Ji and Wilson 2007), coding the animal's position along real (Haggerty and Ji 2015) and virtual linear tracks (Fiser et al. 2016; Pakan et al. 2018; Saleem et al. 2018; Fournier et al. 2020). Various studies report spatial and temporal correlations in the activity of V1 and hippocampal area CA1, including correlated errors in position decoding (Saleem et al. 2018; Fournier et al. 2020), correlated trial-by-trial shifts in preferred spiking locations (Haggerty and Ji 2015), and significant spike-phase coherence of V1 spiking and hippocampal theta oscillations (Fournier et al. 2020).

A similarly broad variety of single-unit correlates is observed in the primary auditory cortex (A1), including activity selective for visual task cues (Brosch et al. 2005), behavioral demands (Scheich et al. 2007), stimulus expectation (Jaramillo and Zador 2011), reward (Scheich et al. 2007), and instrumental action (Niwa et al. 2012). However, much remains currently unknown about the functional role and origins of such “extra-modal” activity correlates, including whether they primarily contribute to local sensory processes or reflect crossmodal interactions in the service of more general and modality-independent cortical functions. While spiking correlates to stimulus location are present in AC (Town et al. 2017), no activity correlated to the spatial position of the subject has hitherto been reported for AC. This would be expected if the underlying mechanisms reflect general functions of cortical sensory processing, including the maintenance and updating of a coherent representation of space or of current and future sensory states across sensory domains (Rao and Ballard 1999; Friston 2005; Gavornik and Bear 2014; Fiser et al. 2016; Pennartz et al. 2019).

To determine whether location-correlated spiking activity exists in the auditory cortical system and whether such activity provides a representation that is coherent with the visual modality, we analyzed single-unit data recorded simultaneously from 2 anatomically connected, sensory cortical areas of freely moving rats: AC and lateral, secondary visual cortex (V2L). We show spatially modulated firing patterns in large proportions

Received: March 7, 2022. Revised: January 31, 2023. Accepted: February 1, 2023

© The Author(s) 2023. Published by Oxford University Press. All rights reserved. For permissions, please e-mail: journals.permission@oup.com.

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

of V2L and AC single units which are reliable over time. Firing patterns in each area collectively tiled the entire behavioral track so that every location was marked by the activity of a subset of neurons. Reconstructions of the rat's position afforded by the spiking activity of each area showed reconstruction errors that were correlated in magnitude and direction, thereby indicating that the sensory-spatial representations in AC and V2L are coherent. Our freely moving paradigm allowed us to establish the contributions of position, head direction, and their temporal derivatives to location-correlated spiking activity in early sensory cortices dedicated to different modalities. Our results uncover striking similarities as well as quantitative differences in the location-correlated neural activity of AC and V2L, suggesting that such activity supports common functions in coordinated mapping of sensory and contextual representations across different sensory modalities.

## Materials and methods

### Experimental design

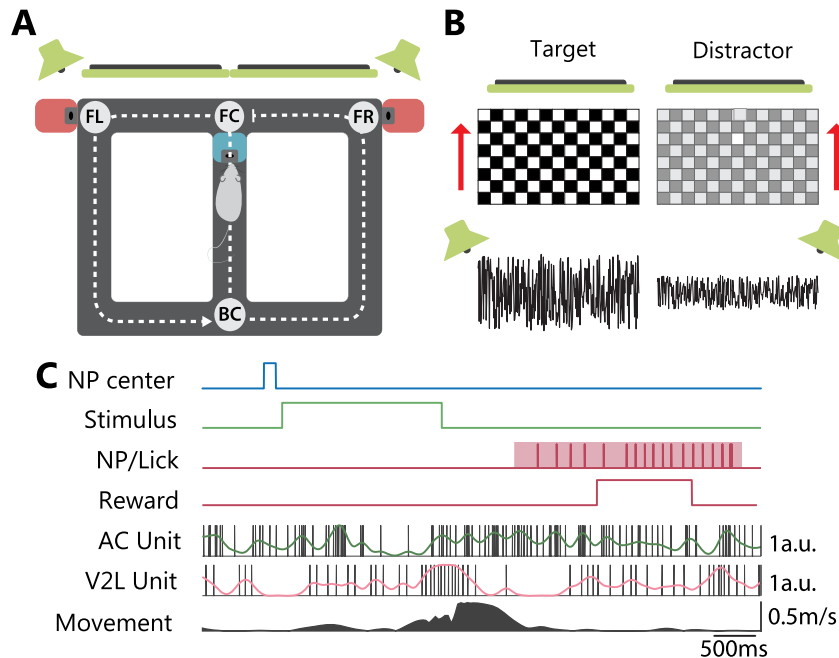
#### Subjects

Experiments were performed on Lister Hooded rats ( $n=3$ , Envigo, The Netherlands) at an age between 9 and 40 weeks. All rats were socially housed during behavioral training, but were individually housed during periods of recordings when rats had an implanted tetrode array, under a normal day/night cycle (lights on: 8:00 am, lights off: 8:00 pm). The rat's food intake was restricted such

that its weight was at least 85% relative to the standard growth curve provided by the breeder (Envigo, The Netherlands), which was corrected for the deviance in weight between the rat and the curve in the week before the start of food restriction. Weights were maintained at a stable level after rats reached healthy, adult body weight (Newby et al. 1990; Clemens et al. 2014). Rats had ad libitum access to water throughout the experiment. All experiments were performed in accordance with the National Guidelines on Animal Experiments and were approved by the Animal Experimentation Committee of the University of Amsterdam.

#### Behavioral setup

Rats were trained to discriminate between auditory and/or visual stimuli on an automated, rectangular figure-8-shaped track (92 × 73 cm) which was raised 55 cm off the ground (Fig. 1A). The track's alleys were made of black-painted aluminum (width = 8.7 cm) and contained raised edges (1.0 cm). At the front of the track, 2 LCD monitors (Iiyama ProLite B2776HDS) and 2 audio speakers (Audaphon Neo CD 3.0) were available for stimulus presentation. The monitors were positioned symmetrically from the center of the track, and the speakers were located above the top-left and top-right corners of the left and right monitors, respectively. During early training stages, transparent polycarbonate walls lined the central alley to prevent the rat from prematurely exiting this alley. Additionally, 2 transparent polycarbonate sliding doors were positioned at the front and back of the central alley. The walls and door at the front of the central alley contained small holes



**Fig. 1.** Behavioral apparatus and task. A) Rats performed a discrimination task on an automatized figure-8 track, which was located in a dark, sound-attenuated laboratory room without other salient visual cues. At the front center T-junction (FC), rats responded to audio-visual stimuli presented from 2 screens and 2 speakers located in front of the track (green) by running to the track's side corresponding to where the most salient stimulus was presented. Rats were rewarded for a correct response with sucrose solution at the ports to the sides of the front alley (pink squares). The dotted line indicates how the track was linearized for analysis. FL, front left; FR, front right; BC, back center. B) Example of a set of stimuli. Stimuli were visual (moving checkerboard), auditory (filtered white noise), or audiovisual. In multisensory trials, target auditory and visual stimuli were always presented at the same side of the track. The rat had to respond to the most salient stimulus (highest contrast and/or volume) and discard the distractor stimulus (which was less salient). C) Trial layout and example AC and V2L spike trains during 7 s of a leftward trial with a correct response. NP center: Blue line indicates the timing of the nose poke (NP) in the central well to initiate stimulus onset. Stimulus: the time interval during which the stimulus was presented (2 s; green line). NP/lick: The time of the rat's nose poke into the left reward well is indicated by the red shaded area and the individual licks by the vertical tick marks. AC/V2L units: Single-unit spikes are indicated by the black vertical tick marks, whereas solid lines indicate Gaussian-smoothed spike trains. Firing rates are observed to fluctuate in relation to stimuli and locations across the maze. Movement: speed of the rat along the linearized trajectory.

to allow for perception of the auditory stimuli. Two reward wells were positioned at the left and right edges of the track's front alley. An additional reward well was positioned in the central alley toward the front-end T-junction. Fluid sucrose solution (15% in tap water) was delivered to the wells by syringe pumps (Razel, VT, USA). All reward wells contained infrared photodetectors to detect nose pokes and licks. The motor activity of the rat on the track was registered with additional photodetectors, which were located near the T-junctions at the front and back ends.

The track was entirely computer-controlled, obviating the need for human interventions during the experiment, and was interfaced with the recording system to ensure synchronized time stamping of behavioral events and neuronal activity patterns. It was positioned in an enclosure of black curtains (2.8 × 2.2 m) within a sound-attenuated room of 3 × 3 m. The room was dimly lit by a small LED light pointed toward the ceiling. The experimenter observed training and experiments from an adjacent room to minimize interference with the rats' behavior.

### Stimuli

On each trial of the behavioral task, rats had to discriminate between a target stimulus, displayed on one of the screens and/or the adjacent speaker, and a distractor stimulus, appearing on the other screen and/or speaker (Fig. 1B). Target and distractor were either unisensory (visual or auditory) or multisensory (audio-visual) and differed in stimulus amplitude. Three types of target/distractor combinations, in equal proportions, were presented to the rat; large-difference unisensory trials (1/3), small-difference unisensory trials (1/3), and small-difference crossmodal trials (1/3). In a separate session for each rat, the threshold amplitude differences (i.e. the amplitude differences at which the rat shows a correct response in 50% of the stimulus presentations) were determined using a staircase procedure. On the basis of these data, stimulus parameters were set for each rat such that the discrimination performance for the large-difference unisensory trials was >70% correct and the discrimination performance for the small-difference unisensory trials was >50% correct. The stimulus settings remained the same for all recording sessions of a rat. The amplitude differences for the stimulus components in the audiovisual trials were identical to the differences for the small-difference unimodal stimuli of the same session. The specific screens (left or right) at which target and distractor were displayed were pseudorandomly selected for each trial such that the target stimulus was never displayed >4 successive trials at the same side and the difference in left and right target presentations across the session was not >4.

Visual stimuli were full screen, upward-moving, and black-and-white checkerboards (0.1 cycles/degree, 4 cycles/s; Fig. 1B). When no visual stimuli were displayed, a gray background was visible. All visual stimuli and the background were gamma corrected and had the same overall luminance as measured with a photometer. Contrast values between light and dark checkers varied between 0 and 1 in which 0 indicates no contrast and 1 indicates the maximum contrast possible with the monitor at its lowest brightness setting.

Auditory stimuli were composed of white noise, which was band-passed between 10 and 25 kHz. When auditory stimuli were absent, background noise was played. Background noise was white noise band-passed between 8 and 12 kHz. For auditory stimuli, the difference between target and distractor was in the relative volume between the speakers. Relative auditory volume ranges between values of 0 and 1, with 0 and 1 indicating that the volume is fully accounted for by 1 speaker or the other speaker.

Background noise was played with a contrast of 0.5, i.e. identical volume through both speakers. Therefore, at every moment during the session, sound was playing at the same volume, which was set to 76 dB at the central reward well.

### Behavioral training

After rats had learned to complete unidirectional laps on the track and to nose poke for a duration of 1 s to earn reward, they were trained to discriminate between target and distractor stimuli and to respond by choosing the side at which the target stimulus was present. When the front door opened at the start of the trial, the target visual stimulus was presented on 1 screen (i.e. on 1 side of the track), while the other screen maintained the gray background. The rat earned a reward if it poked in the well at the side of the track corresponding to where the stimulus was displayed. The stimulus presentation lasted until 3 s following reward delivery, with the aim of strengthening the stimulus–response–outcome association. When the rat made a nose poke at the incorrect side, stimulus presentation stopped immediately. From this stage onward, the length of the nose pokes to start stimulus presentation was increased incrementally from 0 to 0.5–1.5 s (randomized across trials). If the rat performed at least 60 trials within 60 min with 70% correct trials on 3 out of 5 consecutive days, stimuli switched from visual to auditory. If this criterion was met also for auditory stimuli, subsequent sessions included audio, visual stimuli, and audio-visual stimuli in equal proportions and presented according to the pseudorandom schedule (see "Stimuli"). In multisensory trials, the auditory and visual stimulus components were always presented on the same side of the track; i.e. no sensory conflicts were created. Once the rat reached the same criterion with these 3 trial types included in the session, the discrimination problem was made progressively more difficult by introducing distractor stimuli and by lowering the contrast of visual target stimuli in subsequent sessions. The difference in amplitude between target and distractor stimuli was gradually decreased over training sessions but never below the level described above for "small-difference stimuli." In parallel with the increase in difficulty, the display time of the stimuli was progressively shortened until the stimulus duration was 2–3 s at this stage of training (labeled "stage 4"). To prevent rats from developing habitual or stereotyped response preferences, extra sessions were occasionally included in the training. In these sessions, rats were allowed to collect reward from the correct well after sampling the incorrect well. After an incorrect nose poke, trials continued until the correct well was sampled. These sessions did not count toward criterion performance.

In the final task, the trial procedure was similar to training stage 4, and now the small and large difference unimodal stimuli were presented alongside with the crossmodal stimuli (see "Stimuli"). To ensure enough trials were performed in all conditions to allow statistical comparison, each session contained either visual or auditory unimodal trials in addition to multisensory trials. Recordings commenced when rats consistently performed at criterion level: >60 trials in 60 min with >70% correct on large-difference unimodal stimuli and above chance level for small-difference unimodal stimuli and multisensory stimuli.

### Tetrode array and surgical procedure

A custom-made, 128-channel tetrode array was implanted over the right hemisphere of 3 rats (Lansink et al. 2007; Bos et al. 2017). Four bundles of 8 individually moveable tetrodes each were targeted to A1 (AC, −4.4 mm AP, 6.8 mm ML, −4.4 mm DV); lateral, secondary visual cortex (V2L, −5.8 mm AP, 6.0 mm ML,

–2.6 mm DV); hippocampal area CA1 (CA1, –3.6 mm AP, 2.4 mm ML, –2.6 mm DV); and perirhinal cortex area 36 (PRH, –4.4 mm AP, 6.8 mm ML, –7.0 mm DV). Hippocampal and perirhinal data were not used in the current study. Thirty minutes before surgery, rats received the analgesics meloxicam (Metacam, 2 mg/kg) and buprenorphine (Buprecare, 0.04 mg/kg) subcutaneously as well as the antibiotic enrofloxacin (Baytril, 5 mg/kg). Anesthesia was induced using 3% isoflurane in oxygen and maintained with 1–2% isoflurane. Animals were mounted in a stereotaxic device (Kopf; Tujunga, CA, USA) and were placed on a heating pad to maintain their body temperature. A single craniotomy was made such that the points at which the bundles entered the brain were positioned relative to bregma at –5.4 mm AP and 2.8 mm ML (AC), –5.9 mm AP and 4.0 mm ML (V2L), –3.6 mm AP and 1.7 mm ML (CA1), and –3.5 mm AP and 3.8 mm ML (PRH). Six screws were placed into the skull, with the screw positioned over the frontal bone serving as electrical ground for the tetrode array. The hyperdrive was positioned such that the bottom of the bundles touched the cortical surface. The craniotomy was then sealed using silicone adhesive (Kwik-Sil), and dental cement was used to fix the hyperdrive and screws to the skull. Postoperative care consisted of subcutaneous injections of the analgesic meloxicam (2 days) and Baytril antibiotic (1 day) and application of wound healing ointment (Acederm, Ecuphar, Breda, Netherlands). On postoperative days 1–3, rats received 10 g of extra food softened in water to facilitate consumption. Tetrodes were gradually lowered to their target regions across the first week after surgery. During recordings, their depth was estimated from the number of turns to the guiding screws and from the online Local Field Potential (LFP) profiles.

### Data acquisition and preprocessing

Recording sessions took place in the early afternoon during the rats' active period. One recording session was performed per day, with 4–7 sessions per rat. Spikes and LFPs were recorded using tetrodes (Gray et al. 1995) (nichrome, California Fine Wire, 16  $\mu\text{m}$  per lead, gold-plated to an impedance of 500–800 K $\Omega$ ) using a Neuralynx Digitalynx SX recording system (Neuralynx, Bozeman MT). Raw signals were buffered using 4 32-channel unity-gain head stage amplifiers before being passed through an automated commutator (Neuralynx, MN). Each of the 4 tetrode bundles contained an additional electrode that served as a reference channel and was positioned in the white matter near the tetrode bundle. The recorded signals were the raw signals with the reference signal subtracted. For spike recordings, signals were band-pass filtered between 600 and 6,000 Hz. Putative spikes were recorded for 1 ms (16 bit, 32 kHz) from all leads of a tetrode whenever the signal on any lead of that tetrode crossed a predefined threshold. LFPs were low-pass filtered below 300 Hz and were recorded continuously (16 bit, 3.2 kHz). The behavior of the rat was tracked by a ceiling-mounted camera (at 720 by 576 pixels, 25 fps) and were timestamped by the Digitalynx SX.

Spikes were attributed off-line to putative single units (clusters) using the KlustaKwik automatic clustering algorithm (Kadir et al. 2014), which was followed by manual refinement (MClust 3.5). Waveform features used for clustering were energy, the first derivative of the energy, the overall peak height, and the peak height during samples 6–11 where the action potential peak is expected. Clusters were included for further analysis based on a combination of quality metrics (Schmitzer-Torbert et al. 2005): L-ratio (<0.2–0.8), isolation distance (>15–24), and interspike interval (1.2 ms) violations (<0.1%–0.5%).

### Histology

Following the last recording session, animals were anesthetized using isoflurane and electrolytic lesions were made at each tetrode tip by passing current (18  $\mu\text{A}$  for 2 s) through 2 leads of each tetrode. At least 24 h later, the animal was deeply anesthetized with an intraperitoneal injection of nembutal (1.0 mL, sodium pentobarbital, 60 mg/mL, Ceva Sante Animale, Maassluis, Netherlands) and was transcardially perfused with 0.9% NaCl solution, which was followed by perfusion of a 4% paraformaldehyde solution (pH 7.4 phosphate buffered). The brain was extracted and was placed in 4% paraformaldehyde solution for at least 24 h postfixation after which 40- $\mu\text{m}$  transverse sections were made using a vibratome. These were stained with Cresyl Violet, which allowed the reconstruction of tetrode tracks and their endpoints, which were marked by electrolytic lesions (Paxinos and Watson 2007) (Supplementary Fig. S1).

The endpoints of 20 tetrodes were located in A1, while an additional 2 endpoints were located in the adjacent dorsal secondary auditory cortex. Therefore, we will refer to these recordings as made from AC even though the large majority of cells was recorded from A1. Endpoints for 19 tetrodes targeted at V2L were located in that area, while 1 was located in the adjacent dorsal posterior parietal cortex. Because it was not possible to determine with certainty which recording channel this tetrode was associated with, we removed each tetrode of this rat from the dataset in turn and repeated our analyses. This did not lead to qualitatively different results.

### Statistical analysis

#### Statistical procedures

Unless specified otherwise, all statistics were performed using linear mixed models (LMMs) or generalized linear mixed models (GLMMs) in MatLab (MathWorks, Natick, MA), depending on the distribution of the data. All reported statistical quantities (group means, regression slopes, confidence intervals [CIs], etc.) were derived from the (G)LMMs. All reported CIs are 95% prediction intervals. To estimate the degrees of freedom of the denominator for the *F*-tests, the Satterthwaite approximation was used for LMMs and the residual degrees of freedom was used for GLMMs. To correct for multiple comparisons, *P*-values are adjusted using the Holm-Bonferroni method where applicable (Holm 1979) (Fig. 2H).

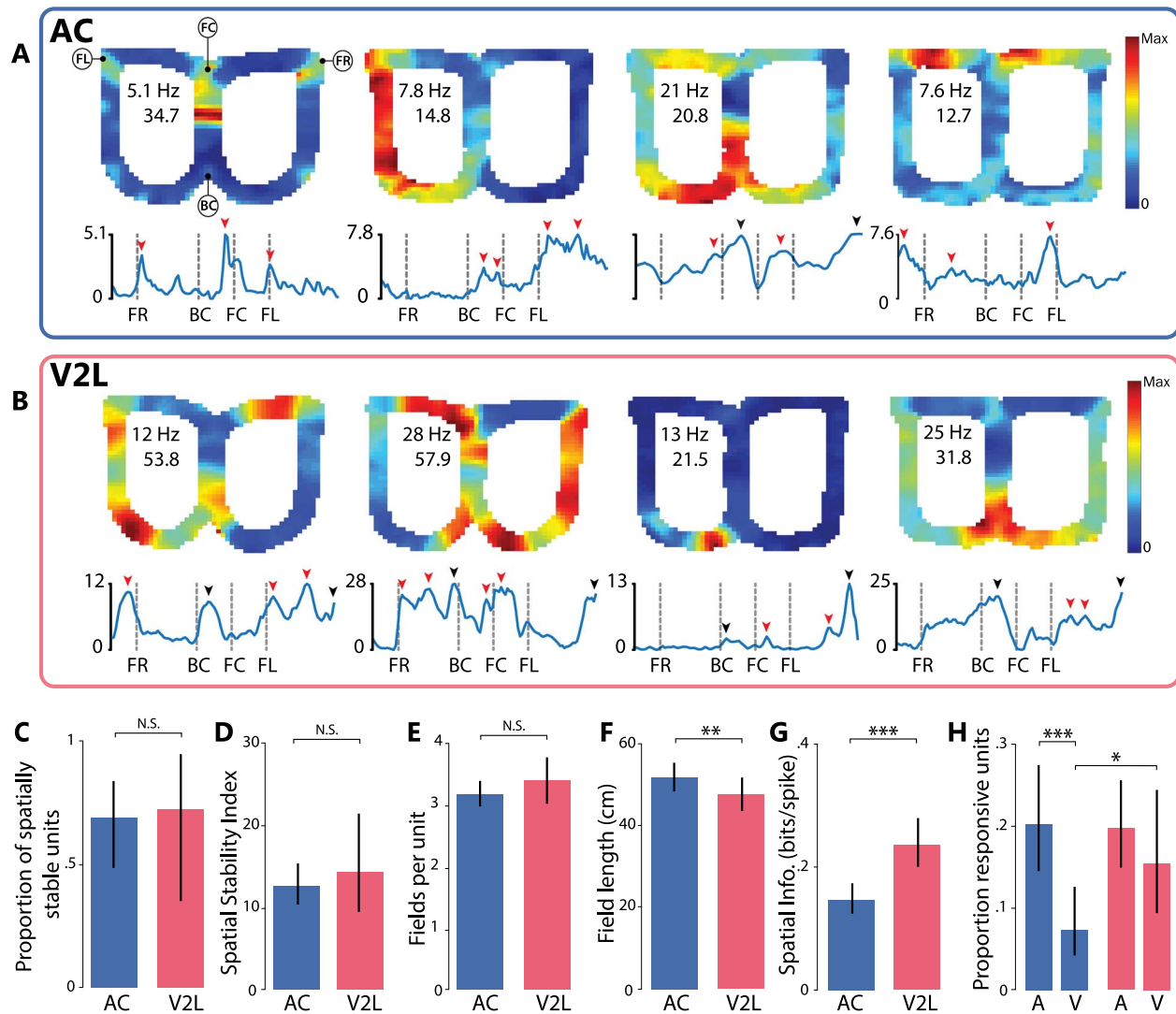
#### Inclusion criteria

Behavioral performance was measured as the percentage of correct responses to stimulus presentation in a session. Sessions were included for analysis only if the performance was above chance level, which we defined as the number of observed correct responses being >95% of a binomial distribution with  $P=0.5$  (chance level) and  $n$  equal to the number of trials in the session. Additionally, to exclude sessions where animals had a preference for a particular response direction, sessions were included only when responses to both left and right stimuli were performed above chance using the same procedure.

#### Z-scored spiking activity and firing rates

Spike trains of individual units were binned into 1-ms temporal bins and were smoothed with an exponential window with a time constant of 150 ms. Z-scored spiking activity was calculated by subtracting the mean from the smoothed spike train and dividing the result by the standard deviation.





**Fig. 2.** Spatially modulated firing patterns of AC and V2L units. A) Top panels show rate maps indicating the spatial firing-rate distribution of four example AC units. Numbers in the left wing of each map represent peak firing rate (top) and spatial stability index (SSI, bottom). Bottom panels show firing rate distributions along the linearized version of the track corresponding to the rate maps shown in the top panels. Red arrow heads indicate individual firing-field peaks. Black arrow heads indicate a single firing-field peak that spans across an edge of the linearized track. Abbreviations below linear rate maps refer to landmark locations (Fig. 1A). B) As (A) but for 4 example V2L units. C) Estimate of the proportion of spatially stable units in the AC and V2L populations. Black vertical lines throughout (C–H) indicate 95% CIs (N.S.,  $P > 0.05$ , F-test). D) SSI of spatially stable units. E) Number of firing fields per spatially stable unit (\*\*,  $P < 0.01$ ). F) Field length per unit (\*\*,  $P < 0.01$ ). G) SI of spatially stable units (\*\*\*,  $P < 0.001$ ). H) Proportions of stimulus responsive neurons in AC (blue bars) and V2L (red) to auditory and visual stimuli (\*,  $P < 0.05$ ).

### Position tracking and 2D rate maps

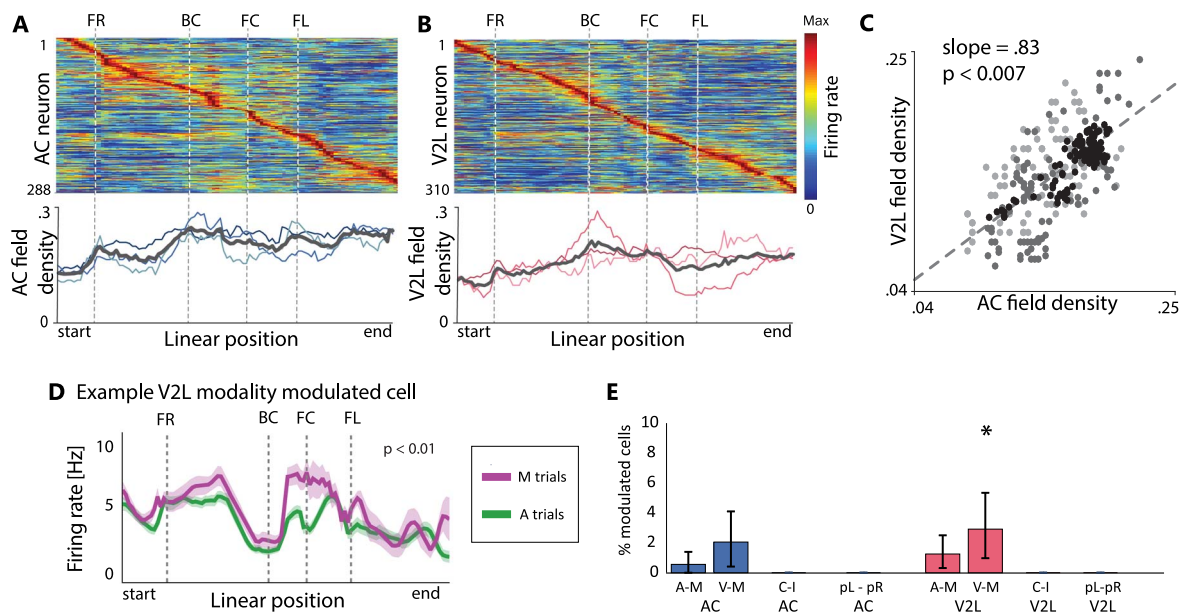
The location of the rat's body and head were determined separately for each frame of recorded video using custom scripts made with Bonsai Editor (Lopes et al. 2015). Frames with erroneously assigned positions were manually corrected. The raw tracking data were smoothed using the “smooth()” function in Matlab using the “rlowess” method and a span of 5 pixels. Head direction was determined for each video frame as the angle of the body-head axis of the rat with respect to the left-to-right axis of the setup.

Rate maps were constructed in 2D by spatially binning (bin size = 10 by 10 pixels or 4.2 cm<sup>2</sup>) the smoothed spike trains of single units into the spatial bin occupied by the rat in each videoframe, summing the firing rates per bin and then averaging over the time spent in each bin (occupancy). Binned spike trains and occupancy maps were independently smoothed by convolution with a 2D Gaussian (std = 1 bin) before averaging.

Time segments in which the running velocity of the rat was  $< 6$  cm/s were excluded.

### Linearization of position data and 1D rate maps

The goal of linearizing position data is to allow more powerful analyses relating spatially modulated spiking activity to behavior. The lateral range of body motion on the track was limited, and the firing fields observed on the 2D rate maps generally spanned the full width of the track. Linearization, therefore, allows to focus on the spatial dimension containing the majority of the rate map structure. Linearization of the rat's position was achieved by first determining the average path of the rat across the setup for each recording session. This was achieved by manually tracing the locations with the highest occupancy of each session's occupancy map. The starting point of the linearized track was chosen as the starting point of the right alley segment at the 3-way junction at the front of the setup, front center (FC; Fig. 1A). The linearized



**Fig. 3.** Spatial distribution of firing fields and lack of field sensitivity to task parameters. A) Top panel: Joint linear rate map including all individual, spatially stable units recorded from AC. Firing rate is color-coded and the individual rate maps are sorted by peak location. Bottom panel: The firing-field density across the spatial bins on the track are shown for the individual rats (thin, colored lines) and the mean across rats (thick, gray lines). The firing-field density expresses for each spatial bin the proportion of individual firing fields that include that bin. B) As (A) but for V2L. C) Correlation between firing-field densities of AC and V2L across spatial bins. Dots correspond to individual spatial bins; different shades of gray correspond to individual rats. The dashed line indicates the linear regression of mean field densities across rats according to the linear mixed effects model. The value of  $P$  indicates the significance of the regression line. D) One of the few AC cells whose rate map differed between unisensory auditory (A, green) and multisensory (M, magenta) trials ( $P < 0.01$ , permutation test). E) Rate maps of AC and V2L cells were assessed for dependency on stimulus modality; A-M, modality-dependent contrast: unisensory A versus multisensory M trials; V-M, unisensory V versus multisensory M; C-I, outcome-dependent contrast, rewarded (correct, C) versus unrewarded (incorrect, I) trials, prev. trial: choice in previous trial contrast (previous trial choice left [pL] vs. previous trial choice right [pR]). Error bars denote 95% bootstrapped CIs, asterisk denotes significant fractions of cells (significant if the CIs do not overlap with 0%). The small fractions of modality-modulated cells are likely stimulus-responsive cells such as in (D). Stimuli were presented when the animal occupied locations near the front center (FC).

track then consisted, in this order, of the right side of the track, central alley, and left side of the track. The end position was the center of the back-alley segment at the 3-way junction. The linear position was then determined as the point along this linearized track closest to each observed 2D position. Raw linearized position was smoothed by convolution with a Gaussian with a standard deviation of 2 pixels (0.4 cm). Linear speed was then calculated from the linear location using the “gradient()” function in Matlab, and linear acceleration was calculated similarly from linear speed. Finally, because the linearized track length varied slightly between rats, linear location was normalized to the mean length across animals (334 cm). Linear rate maps were made similarly to the 2D rate maps but using the linear position data by using the same inclusion criteria and smoothing parameters and with a bin size of 16 pixels (3.3 cm). Joint rate maps (Fig. 3A and B) were constructed from the linear rate maps of each neuron by normalizing each rate map between 0 and 1 and then sorting all rate maps of all sessions and rats by the location of the peak firing rate (dark red in Fig. 3A). These rate maps were crossvalidated (Supplementary Fig. S3B and D) using the following procedure. First, the rate maps were recomputed separately on the first and second halves of the session and then neurons in both rate maps were sorted using their peak location in the ratemap from the first half of the session. This method reveals whether the reported structures (as in Fig. 3A and B) could be an artifact of the normalization and sorting procedures.

### Firing fields

The procedure for determining the location of firing fields was similar to the method described by Haggerty and Ji (2015) using the unit’s linear rate map as a basis. First, the rate map was smoothed using the “smooth()” function in Matlab (MathWorks) with the “rlowess” method and a span of 5 bins. The baseline firing rate was determined as the 40th percentile of this smoothed rate map. The baseline was subtracted from the rate map, and local maxima were determined for the baseline-corrected rate map. Local maxima were kept for further processing if the rate was  $>1$  Hz and  $>0.2 \times$  the baseline rate (we found quantitatively and qualitatively very similar results with more restrictive inclusion criteria). Field boundaries were determined as the bins nearest to a peak, where the firing rate was  $<10\%$  of the peak firing rate. This procedure sometimes produced very small fields near or on the slope of larger fields (“shoulders”). Such shoulders were discarded if the border of the field was  $\leq 2$  bins apart ( $\leq 6.6$  cm) from the border of a taller field unless the peak was  $\geq 2$  bins ( $\geq 6.6$  cm) away from its own border and the peak was  $>1.5$  times the value at this border. After the removal of spurious small fields, the borders of remaining fields were extended to the bins where the activity fell below the peak cutoff of 1 Hz and 0.2 times the baseline rate, or until they reached the border of another peak. This procedure was followed for each single unit from the tallest to the lowest peak. Single fields that spanned the T-junctions at the front and back of the experimental track were prevented from being detected as

2 separate fields by considering the bins of the linear rate maps that border the T-junctions as adjacent.

### Spatial stability

The spatial stability of the firing of individual units was determined using a permutation analysis. First, separate linear rate maps of each individual trial were constructed, similarly to the procedure described above, except that the data were split into leftward and rightward trials. The correlation between a single-trial rate map and the rate maps of all other trials on the same side of the track was computed (Pearson correlation) for each trial and then the average correlation coefficient was computed. These steps were repeated for 1,000 shuffled versions of the spiking data. For each shuffled iteration, the spiking data were temporally rotated within a single trial by a random number of samples (Louie and Wilson 2001). This method of shuffling leaves the temporal structure of spiking patterns largely intact. Positions and spikes emitted during periods of immobility (<0.06 cm/s) were excluded from analysis. A unit was considered as spatially stable if its average single-trial rate map correlation, for both leftward and rightward laps, was >95% of the distribution of average, shuffled single-trial rate map correlations of the corresponding side. To calculate the spatial stability index (SSI), the mean of the distribution of average, shuffled single-trial rate map correlations was subtracted from the average observed single-trial rate map correlation; the result was divided by the standard deviation of the shuffled distribution. This procedure was done for each unit for both leftward and rightward trials and the reported SSI was the mean of those 2.

The probability of the spatial stability of a unit was modeled using a GLMM with link function  $g(P_{\text{stable}}) = \ln(P_{\text{stable}})$  and the following equation:

$$P_{\text{stable}} \sim \text{Area} + (1 + \text{Area}|\text{Subject}),$$

where **Area** is a categorical variable indicating the cortical area where the unit was recorded, and **Subject** is a categorical variable indicating from which experimental subject the data originated.

### Spatial information

Spatial information (SI) for each unit was calculated from the linearized location data and firing rate following (Skaggs et al. 1992):

$$\text{SI} = \sum_{i=1}^N P_i \frac{X_i}{r} \log_2 \frac{X_i}{r},$$

where  $P_i$  is the probability of finding the animal in bin  $i$ ,  $X_i$  is the sum of the firing rates observed when the animal was found in bin  $i$ ,  $r$  is the mean spiking activity of the neuron, and  $N$  is the number of bins of the linearized trajectory (104). The SI for each unit was modeled with a GLMM using a gamma distribution with link function  $g(\text{SI}) = \text{SI}^{-0.01}$  and the following equation:

$$\text{SI} \sim \text{Area} + (1 + \text{Area}|\text{Subject}).$$

### Correlating firing-field densities

Firing-field densities were calculated per rat by first determining for each linear spatial bin the number of firing fields, across all units, in which this bin took part. Then, this number was divided by the subject's total number of fields. The average firing-field density is reported as the mean of the field densities for

the individual rats. Firing-field densities of AC and V2L were correlated with an LMM using the following equation:

$$\text{Density}_{\text{AC}} \sim \text{Density}_{\text{V2L}} + (1 + \text{Density}_{\text{V2L}}|\text{Subject}).$$

### Linear regression of spiking activity on behavioral covariates

Linear regression was used to determine whether firing-field correlations between AC and V2L could be explained by similar linear dependencies between single-unit firing rates and behavioral covariates. The instantaneous, z-scored spiking activity of each unit was regressed on linear running speed, acceleration, head direction, and change in head direction (angular velocity) using the general linear model:

$$Z \sim \text{Spd} + \text{Acc} + \text{Sin}_{\text{HD}} + \text{Cos}_{\text{HD}} + \omega_{\text{HD}},$$

where  $Z$  is the instantaneous, z-scored spiking activity, **Spd** is linear speed, **Acc** is linear acceleration, **Sin<sub>HD</sub>** and **Cos<sub>HD</sub>** are the sine and cosine of head direction, and  $\omega_{\text{HD}}$  is the angular velocity in head direction. The model residuals of each unit were used to produce linear rate maps and to subsequently determine firing fields without linear dependencies on the behavioral covariates.

### Stimulus responsiveness

The responsiveness of a unit to the sensory stimuli was assessed by statistically comparing the mean, z-scored firing rates of the time intervals of [350–50]-ms prestimulus onset and [0–300]-ms poststimulus onset using Wilcoxon's signed-rank test ( $\alpha = 0.05$ ). A subset of units gradually increased or decreased firing rates before stimulus onset ("ramping activity") without showing a change in spiking activity at stimulus onset. To preclude that such activity would erroneously be considered as significantly stimulus responsive, only units with a stable prestimulus onset firing rate were considered; i.e. the unit's firing rates between [1,000–700]- and [350–50]-ms prestimulus onset were required to be similar (Wilcoxon's signed-rank test,  $P > 0.05$ ). The probability of a unit's responsiveness to stimuli was modeled using a GLMM with link function  $g(P_{\text{resp}}) = \ln(P_{\text{resp}})$  and following equation:

$$P_{\text{resp}} \sim \text{Area} * \text{Modality} + (1 + \text{Area} + \text{Modality} | \text{Subject}),$$

where **Modality** is a categorical variable indicating the modality (auditory or visual) of the stimulus being considered.

### Information theoretic analysis

For the information theory-based analyses, linear spatial bins of ~16.4 cm were used such that the linearized track was divided into 21 bins. As predictors, we used running speed, head direction ( $\theta_{\text{head}}$ ), and head direction change ( $\Delta\theta_{\text{head}}$ ). Running speed and  $\theta_{\text{head}}$  were calculated as described above, and  $\Delta\theta_{\text{head}}$  was calculated as the degrees per second change in  $\theta_{\text{head}}$ . The predictors were binned in 21 equipopulated bins to match the number of location bins. Spike counts were binned in 300-ms time bins, and instances with running speed <0.1 m/s were excluded. Discrete mutual information (MI) and discrete conditional mutual information (cMI) were computed using the Java Information Dynamics Toolkit (Lizier 2014). Bias due to finite sample sizes was corrected by generating, for every computation of MI and cMI, a population of 500 surrogates. When creating the surrogates, only spike count vectors were shuffled such that the relationship between the target and conditional predictor (in the case of cMI)

was preserved. The MI and cMI values reported are the differences between the observed values and the average of the surrogate populations. Differences in MI between AC and V2L were statistically assessed using Mann–Whitney’s U test. The *P*-value of cMI values of individual units was computed as the fraction of the surrogate dataset which had higher information values than the observed one. To correct for multiple comparisons, a Bonferroni correction was applied across all neurons and was tested at the 0.05 significance level. Reported confidence bounds correspond to 95% bootstrap CIs, which were computed with the [Waskom \(2021\)](#).

### Encoding: predicting spike trains from behavioral variables

First, the position of the rat was taken as the 2D variable with *x*- and *y*-coordinates with a resolution of 0.205 cm/pixel recorded by the system without further binning. Encoding was performed with a random forest encoder using 100 trees and 5-fold cross-validation with randomized folds ([Benjamin et al. 2018](#)). Encoding over time was performed using continuous folds to preserve the order in time. Encoding quality was measured with the Poisson pseudo- $R^2$  score and was averaged over folds. Statistical comparisons of encoding quality for individual predictors, and comparisons of improvement in encoding quality above all other predictors, were made using Wilcoxon’s signed-rank test and used the average of the AC and V2L (improvement in) encoding quality for each predictor. Reported confidence bounds correspond to 95% bootstrap CIs, which were computed with the [Waskom \(2021\)](#).

### Decoding of animal position

The position of the rat was decoded from the neuronal data recorded from AC or V2L if a session included at least 16 neurons from that area showing a rate map peak > 2 Hz; only those neurons were included for each session. This number was determined to provide a balance between decoding quality and number of included sessions. Spikes were binned in 400-ms bins, and the true position (i.e. the actual position of the rat on the track) at every timeframe was assigned to a spatial bin on the linearized track (total of 35 bins, bin size ~9.8 cm). When the linear position changed spatial bins within a temporal bin, the position was assigned to the spatial bin which occurred most often within the temporal bin. Running speed was linearly interpolated at the centers of the temporal bins. Samples with speed <0.1 m/s and with spike count <5 were excluded. A Bayesian classifier was employed to predict the spatial bin occupied by the rat on the basis of the temporally binned neural data ([Davidson et al. 2009](#)). A 5-fold crossvalidation routine with shuffling was used, with identical shuffling (i.e. similar-sized training set for each fold) across the 2 areas for a given session.

Decoding errors were used as a main metric for decoding performance and were computed as the Euclidian distance between the centers of the true and the decoded spatial bin in 2D space. Pearson correlations of instantaneous decoding errors between the 2 cortical areas in time were calculated to assess whether AC and V2L encoded the same position. This was performed separately for the error in the *x*-direction and that in the *y*-direction in 2D space to preserve the directionality of the error in addition to its magnitude.

Instantaneous error correlations resulting from these computations were compared with error correlations that were computed following shuffling of the errors within the same spatial bin and running speed range ([Saleem et al. 2018](#)). Running speed bins were defined per session by taking the full range of speeds and subdividing it into 5 equipopulated bins. Significance of differences in error correlations before and after shuffling were tested using Wilcoxon’s signed-rank tests. Joint error

density maps for AC and V2L were computed for the error correlations of recorded and shuffled data and for the *x*- and *y*-directions separately. Joint error density maps were averaged across all included sessions and were smoothed with a Gaussian filter with a standard deviation of 4 spatial bins. The relative probability of observing an error of a particular size and direction in the recorded versus shuffled data was calculated by taking the difference between the actual and shuffled joint density maps.

A bootstrapping procedure was performed for testing how the decoding performance depended on the size of the population. First, for each included session, 50 unique, random groups of units were selected for each ensemble size (ranging from 5 to the maximum number of units in each session minus one). Then, the decoding analysis was performed for each group before averaging the decoding performance across the groups. For the largest ensembles, with 1 fewer unit than the session total, it was not possible to create 50 unique groups. For these ensembles, some groups were included twice. Crossvalidation was performed for every group by splitting the data into a training set (80%) and test set (20%), with identical shuffling across all groups of a session.

To exclude the possibility that the observed correlations in instantaneous decoding errors are a result of decoding artifacts in sessions with poor decoding, Pearson’s correlation was computed for the average decoding errors and the correlation in instantaneous decoding errors across sessions for AC and V2L average errors separately and for instantaneous errors in the *x*- and *y*-directions separately.

### Head direction tuning curves

Observed spike counts as regarding head direction were determined per video frame and were counted for 60 overlapping bins along the unitary circle. Each bin was  $\pi/6$  radians wide and the bin overlap was 80%. The firing rate at each of the 60 directions was assumed to follow a Poisson distribution. Maximum likelihood estimates and 95% confidence bounds of the Poisson rate were calculated from a Gamma distribution using MatLab’s “`makedist()`” and “`icdf()`” functions, with  $\alpha$  equal to the number of spikes in each bin plus one, and  $\beta$  equal to the inverse of the number of observed head directions in each bin plus one. This resulted in rate estimates per video frame, which were multiplied by the framerate to produce firing rates in Hertz.

## Results

We investigated the responsiveness of neurons in the lateral, secondary visual cortex (V2L) and AC to spatial location when rats were running on a rectangular, figure-8-shaped track ([Fig. 1A](#)). On the track, rats performed an audio-visual discrimination task in which they earned reward by responding to the most salient stimulus out of 2 by running from the stimulus presentation site to the reward well on the track side corresponding to the location of that stimulus ([Fig. 1A and B](#)). Our analyses were primarily based on the spatial components of the rats’ behavior regardless of task performance (a summary of which is provided in [Supplementary Fig. S2](#)). If the neuron’s spiking activity was reliably modulated by the rat’s body location over the course of a recording session, we refer to this as “location-correlated activity.” We emphasize that this definition includes not only responsiveness to allocentric position but also to specific conjunctions of locally available sensory cues and task-related information. With this caveat in mind, we will sometimes use the term “location-selective” as well.



We made simultaneous recordings from both V2L and AC, with a total of 526 single units from V2L and 603 units from AC across 17 recording sessions from 3 rats. Units in both AC and V2L showed firing patterns with  $\geq 1$  peaks in firing rate on various locations of the track (2A-B). For all further analyses, we selected the units with sufficient firing on the track, i.e. peak firing rate  $> 2$  Hz in the spatial map, which amounted to 400 units from V2L and 413 from AC. The locations of the tetrode endpoints were verified with histology (Supplementary Fig. S1).

### Spiking activity of most auditory and visual cortical neurons is reliably correlated to location

First, we assessed whether neurons in AC and V2L displayed spatially modulated firing patterns. Visual inspection of rate maps of the spatial firing distribution of individual neurons indicated that some units in AC and V2L showed increased spiking activity in a single, concentrated location on the track, whereas other units displayed modulations of spiking activity at multiple areas or across a larger area of the track (Fig. 2A and B, top panels). An important constraint for a unit to reliably code location is that the firing activity at that location is consistently, rather than incidentally, present across individual traversals through that location. Therefore, we assessed the reliability of each unit by computing pairwise correlations between all single-trial rate maps of that unit and comparing the observed mean pairwise correlation with a shuffled distribution. A unit was considered spatially stable if its observed mean correlation was  $> 95\%$  of the shuffled distribution. The proportions of spatially stable units were very similar for AC and V2L and comprised on average about 70% of the total population; Fig. 2C; AC: 0.69 (95% CI: 0.49, 0.84), V2L: 0.72 (CI: 0.31, 0.94),  $F(1, 32) = 0.072$ ,  $P = 0.79$ , ANOVA. We quantified the spatial stability of a unit using a spatial stability index (SSI), which is defined as the number of standard deviations by which the observed mean correlation was removed from the shuffled distribution (Fig. 2D). The SSI of spatially stable units was similar between AC and V2L; Fig. 2D; AC: 12.7 a.u. (CI: 8.39, 15.44), V2L: 14.3 a.u. (CI: 10.4, 24.4),  $F(1, 1194) = 0.44$ ,  $P = 0.50$ , ANOVA. Supplementary Figure S3 shows firing rate maps for the same, spatially stable example units as in Fig. 2A and B, split between the first and last 30% of the recording session, illustrating spatial stability. Examples of spatially unstable units are shown in Supplementary Fig. S4. In summary, large fractions of AC and V2L units display spatially stable activity patterns with high reliability.

Individual firing fields were then identified for spatially stable units as local increases in the mean spiking activity of a unit on the linearized representation of the track (Fig. 2A and B). Neurons in AC and V2L exhibited a similar number of around 3 firing fields per unit; Fig. 2E; AC: 3.17 (CI: 2.97, 3.39), V2L: 3.40 (CI: 3.20, 3.61),  $F(1, 588) = 2.22$ ,  $P = 0.14$ , ANOVA; the average length of firing fields was significantly smaller in V2L than in AC—Fig. 2F; AC: 50.5 cm (CI: 47.2, 54.1), V2L: 46.3 cm (CI: 43.5, 49.4),  $F(1, 1939) = 9.0$ ,  $P = 0.003$ , ANOVA—indicating that the spatial granularity of spiking activity is modestly finer in V2L than AC.

The extent to which units show spatially modulated firing can be expressed as the information about the rat's position that is conveyed by a single spike, which is quantified as spatial information (SI; Skaggs et al. 1992). In line with the smaller firing fields in V2L, the SI was significantly higher for V2L units; Fig. 2G; AC: 0.15 bits/spike (CI: 0.12, 0.17), V2L: 0.24 bits/spike (CI: 0.20, 0.28),  $F(1, 556) = 16.5$ ,  $P = 5 \times 10^{-5}$ , ANOVA.

### Both auditory and visual cortical neurons respond to discrete auditory and visual stimuli

Besides showing location-selective firing, subsets of units in both AC and V2L responded with significant firing-rate changes to the auditory and visual stimuli, presented as individual (unisensory) task cues; Fig. 2H; stimulus conditions were pooled for each modality; see Materials and methods. As expected, a larger proportion of AC neurons responded to auditory than to visual stimuli in all rats; AC auditory: 0.20 (CI: 0.15, 0.28), AC visual: 0.07 (CI: 0.04, 0.13),  $F(1, 30) = 10.34$ ,  $P = 0.003$ , ANOVA. Additionally, responsiveness to visual stimuli was more common in V2L than AC; V2L visual: 0.15 (CI: 0.09, 0.24),  $F(1, 30) = 4.60$ ,  $P = 0.0402$ , ANOVA. Surprisingly, however, a comparable proportion of V2L neurons responded to both visual and auditory stimuli; V2L auditory: 0.20, (CI: 0.15, 0.26),  $F(1, 30) = 0.93$ ,  $P = 0.34$ , ANOVA; and auditory responses were equally common in both areas;  $F(1, 30) = 0.013$ ,  $P = 0.91$ , ANOVA. Although AC and V2L responsiveness to discrete stimuli is not the focus of our current analyses, these results not only indicate that responses to stimuli in  $> 1$  modality are common in both AC and especially V2L but also underscore the existence of substantial heterogeneity in cortical sensory selectivity.

Stable, spatially modulated activity patterns were much more abundant in both areas than responses to sensory stimuli (AC:  $P < 10^{-15}$ ; V2L:  $P < 10^{-13}$ , binomial tests on pooled data). We then asked whether the spatial tuning of cells depends on behavioral parameters such as stimulus modality and reward (see, example cell, Fig. 3D). To assess this, we constructed rate maps for different splits of the trials; Fig. 3E; modality: unisensory (i.e. auditory or visual) versus multisensory trials, reward outcome: rewarded versus unrewarded trials, and trial history: left-side choice versus right-side choice in the previous trial; and we tested whether these rate maps differed significantly (permutation testing,  $n = 1,000$  shuffles). In both areas, small proportions of cells had their rate maps modulated by modality; AC, unisensory auditory versus multimodal trials: 0.6 (0.0–1.4) % of cells (mean  $\pm$  95% bootstrapped CI), AC, unisensory visual versus multimodal trials: 2.1 (0.4–3.7) % of cells, V2L, unisensory auditory versus multimodal trials: 1.25 (0.3–2.5) % of cells, V2L, unisensory visual versus multimodal trials: 2.9 (0.8–5.3) % of cells; while no modulation of rate maps as a result of trial outcome (rewarded vs. unrewarded) or choice in the previous trial was found in either area. The small fractions of cells having rate maps that were dependent on modality were not significantly larger than chance in AC and were borderline significant in V2L (the fraction of cells is significant if the bootstrapped CIs do not overlap or border with 0%).

### Spatial distributions of firing fields are highly correlated across auditory and visual cortex

The reliable coding of specific locations by individual units is necessary but is not sufficient for building a representation of an environment or the sequence of sensory states an animal experiences when navigating across the track. Another prerequisite for either type of representation would be that the spatial distribution of firing fields in each area covers the entire series of locations traversed by the animal. Figure 3A and B shows the linear rate maps for all spatially stable units ordered by peak location for AC and V2L and reveals that firing-field peaks occur at every location along the track. We further analyzed the density of firing fields tiling all locations by computing the proportion of firing fields which include a specific spatial bin on the linearized track. The distributions of firing-field densities confirm that the

entire track was covered by AC and V2L firing fields (Fig. 3A and B, bottom; for crossvalidation of population rate maps, see Supplemental Fig. S3).

The joint rate maps and field density plots indicate similar firing-field distributions in AC and V2L. Indeed, the field densities across spatial bins of the track were highly correlated between the areas; Fig. 3C; regression slope: 0.83 (CI: 0.57, 1.09),  $F(1, 39.5) = 3.2$ ,  $P = 0.007$ , ANOVA. This strong correlation, however, may also be due to similar responses of AC and V2L neurons to behavioral covariates such as locomotion speed, acceleration, and head direction ( $\theta_{\text{head}}$ ). To correct for the potential influence of these covariates on field densities, we performed linear regression of spiking activity on the running speed, acceleration, head direction, and the angular change in head direction ( $\Delta\theta_{\text{head}}$ ) and repeated the detection of firing fields and field density correlations on the residuals. When using the model's residuals, field densities between AC and V2L were still highly correlated; regression slope: 0.71 (CI: 0.36, 1.06),  $F(1, 3.0) = 16.1$ ,  $P = 0.028$ , ANOVA. It is, therefore, unlikely that locomotion and head direction can fully explain the firing-field densities of AC and V2L neurons on the track or their shared spatial distribution. These analyses, however, did not take into account possible nonlinear relationships between location and behavioral covariates.

### Spiking activity carries information about position and head direction

We next used information-theoretic measures to quantify the influence of navigation parameters, i.e. position, running speed, head direction ( $\theta_{\text{head}}$ ), and changes in head direction ( $\Delta\theta_{\text{head}}$ ), on the spiking activity of recorded units. MI quantifies the reduction in uncertainty obtained about spiking activity after observing these factors and captures both linear and nonlinear relationships (Lizier 2014; Olcese et al. 2016). Of these individual factors, position carried the most information about spiking activity of both AC and V2L units, which was followed by head direction (Fig. 4A; MI of position vs.  $\theta_{\text{head}}$ : AC:  $P = 1.1 \times 10^{-15}$ ; V2L:  $P = 3.0 \times 10^{-15}$ , Wilcoxon's signed-rank test). To give an impression of head direction-associated spiking activity at the level of single neurons, Supplemental Figure S5 shows the average firing rates for all head directions of example neurons as they were recorded during the experiment. Running speed and changes in head direction carried relatively little information about the spiking activity in both areas. For all individual factors, the mean MI was higher in V2L than in AC ( $P = 2.0 \times 10^{-6}$ ; running speed:  $P = 4.7 \times 10^{-4}$ ,  $\theta_{\text{head}}$ :  $P = 1.4 \times 10^{-7}$ ,  $\Delta\theta_{\text{head}}$ :  $P = 1.6 \times 10^{-5}$ , Mann-Whitney's U test).

When a particular  $\theta_{\text{head}}$  is predominantly encountered at a given position, MI cannot distinguish between the possibilities of a neuron encoding either that position, head direction, or a combination of the 2. To determine the amount of information carried by spikes about speed,  $\theta_{\text{head}}$  and  $\Delta\theta_{\text{head}}$ , that cannot be explained by (nonlinear) correlations with position, we computed the debiased cMI between each of the 3 factors conditional on position (Fig. 4B) (Lizier 2014; Bos et al. 2019). Averaged across all spatially stable units, the amount of information that spike trains carried about running speed and changes in head direction beyond position, i.e. mean cMI (spikes, speed | position) and mean cMI (spikes,  $\Delta\theta_{\text{head}}$  | position), was negative for both AC and V2L, confirming that running speed and changes in head direction contributed little information about the spiking activity on top of position (Fig. 4B). By contrast, the average cMI for head direction, cMI (spikes,  $\theta_{\text{head}}$  | position), was significantly larger than 0 for both AC and V2L (Fig. 4B, AC:  $P < 10^{-15}$ ; V2L:  $P < 10^{-15}$ ; Wilcoxon's signed-rank test), indicating that the spiking activity of both AC

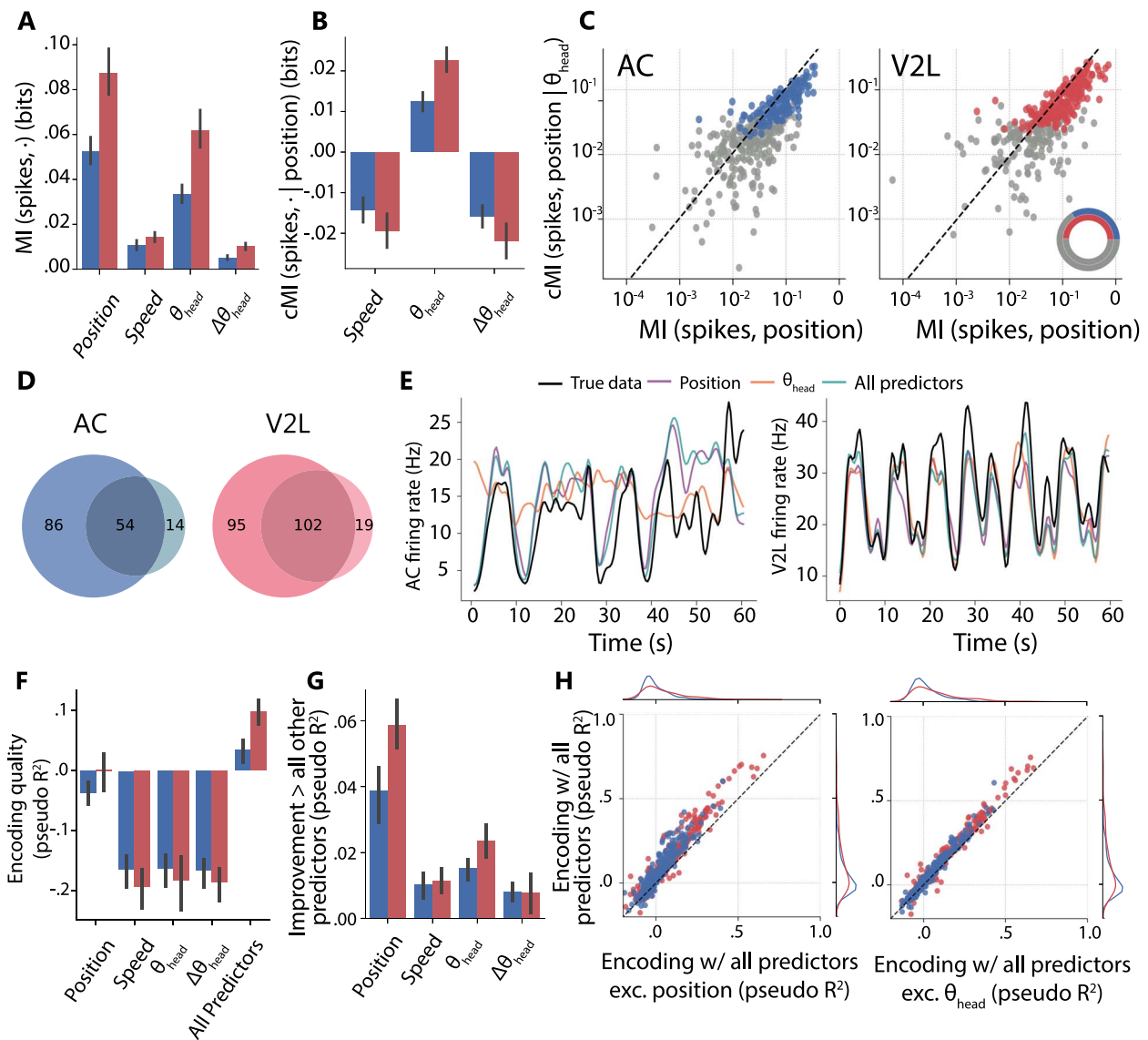
and V2L units coded information about head direction in addition to information about position.

To study the prevalence of position coding in populations of single units, we calculated cMI between spikes and position conditional on  $\theta_{\text{head}}$ , thereby excluding any contributions from  $\theta_{\text{head}}$  (Fig. 4C). Significant cMI (spikes, position |  $\theta_{\text{head}}$ ) was found in 140 out of 413 AC units (33%) and in 197 out of 400 V2L units (49%), indicating that spiking activity of large proportions of single units in both areas contains information on spatial position on top of any information on  $\theta_{\text{head}}$ , with V2L showing a significantly larger proportion ( $z = 4.64$ ,  $P = 1.7 \times 10^{-6}$ , binomial test; Fig. 4D). Smaller, but substantial subsets of neurons in each area showed significant cMI (spikes,  $\theta_{\text{head}}$  | position), indicating that the spike trains of these neurons carry information about head direction even after correcting for position (Fig. 4D, AC: 16.5% and V2L: 31.2%), with V2L here also showing a significantly larger proportion of neurons ( $z = 4.9$ ,  $P = 4.2 \times 10^{-7}$ , binomial test). Among all units encoding position and/or head direction, V2L units were also significantly more likely to provide information on both factors (Fig. 4D, AC: 13.1%, V2L: 25.5%,  $z = -4.50$ ,  $P = 3.4 \times 10^{-6}$ , binomial test). In summary, this analysis confirmed that substantial fractions of AC and particularly V2L neurons continued to show location- and head direction-associated selectivity after correcting for the influence of the other behavioral covariate.

### Position is the strongest predictor of auditory and visual cortex single-unit activity, followed by head direction

Next, we used a random forest encoding algorithm (Benjamin et al. 2018) to determine how well spiking activity could be predicted by each of the behavioral factors. This provided the opportunity to extend the analysis of dependencies between behavioral factors and spiking activity beyond 2 behavioral factors. Encoding of firing rates of 2 units from AC and V2L is exemplified in Fig. 4E. For both areas, encoding quality was best for a model incorporating all predictors (Fig. 4F; all predictors vs. only position for AC and V2L combined,  $P < 10^{-15}$ , Wilcoxon's signed-rank test), whereas of the individual factors, position provided the best performance (position vs. each other individual predictor for AC and V2L, all  $P_s < 10^{-15}$ , Wilcoxon's signed-rank tests). The large effect of position on encoding performance could not be explained by confounds arising from the other behavioral factors since adding position to a model already utilizing all other factors led to substantial improvements in encoding performance for both AC and V2L (Fig. 4G; position vs. each other individual factor, all  $P_s < 10^{-15}$ , AC and V2L combined, Wilcoxon's signed-rank tests). Comparing the contributions of position and  $\theta_{\text{head}}$  separately for individual units confirmed that, among the behavioral factors considered, position was dominant in explaining spiking activity in both areas, while  $\theta_{\text{head}}$  provided the second strongest contribution (Fig. 4G). This is also illustrated in Fig. 4H, which shows, for each neuron, the improvement in encoding when adding  $\theta_{\text{head}}$  or position to models containing all other behavioral factors. For the vast majority of units in both areas, adding either factor improved encoding quality, but this difference was larger for position than for  $\theta_{\text{head}}$ . When analyzing the mean improvement in decoding performance achieved beyond all other predictors, we found that this value was larger for position (0.048) than for  $\theta_{\text{head}}$  (0.019; all  $P_s < 10^{-15}$ , Wilcoxon's signed-rank tests) or the other 2 covariates. Furthermore, the improvement due to  $\theta_{\text{head}}$  was larger than for these other factors (speed: 0.011;  $P < 10^{-15}$ ;  $\Delta\theta_{\text{head}}$ : 0.008;  $P < 10^{-15}$ , Wilcoxon's signed-rank tests).

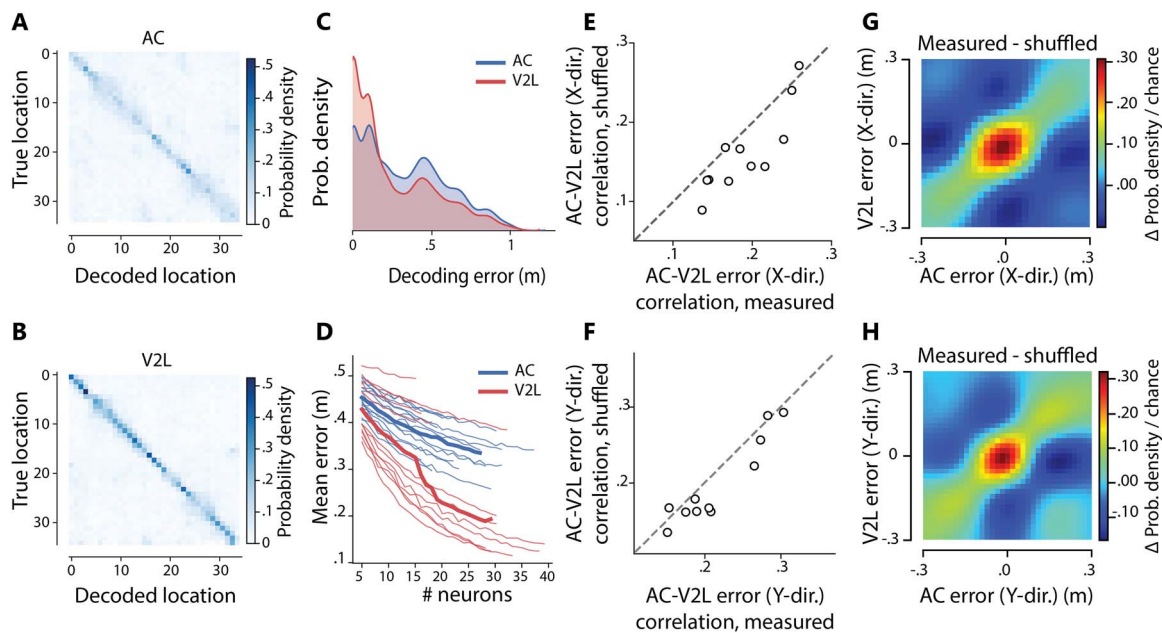
Both the cMI and encoding analysis show that neurons can carry significant information on both factors. However, other



**Fig. 4.** Encoding: Predicting single-unit spiking activity from behavioral factors. A) Average MI between spiking activity of AC (blue) and V2L (red) single neurons and the behavioral factors position, running speed, head direction ( $\theta_{\text{head}}$ ), and head direction change ( $\Delta\theta_{\text{head}}$ ). Error bars represent 95% bootstrapped confidence bounds. B) Debiased cMI between spiking activity and behavioral factors speed, head direction and changes in head direction, conditional on position. Error bars as in (A). C) Relationship between the MI of spikes and position and the cMI between spikes and position conditional on head direction for all spatially stable single units for AC (left) and V2L (right). Blue/red points mark units with significant cMI about position conditional on  $\theta_{\text{head}}$ , indicating that these units carry significant information on position that cannot be explained by  $\theta_{\text{head}}$ . Inset shows in color the number of units per area having significant cMI about position conditional on  $\theta_{\text{head}}$  as fraction of the total number of units (gray: non-significant units). D) Venn diagrams showing for each area the number of neurons transmitting significant cMI (spikes, position |  $\theta_{\text{head}}$ ) (dark color) and cMI (spikes,  $\theta_{\text{head}}$  | position) (light color). Overlapping region indicates neurons which transmit significant information on both position and  $\theta_{\text{head}}$ , neither of which can be explained entirely by the other factor. E) A random forest encoder was used to predict spiking behavior on the basis of position, running speed,  $\theta_{\text{head}}$  and  $\Delta\theta_{\text{head}}$ . Figure shows 60 s of firing rate of an example AC unit (black line; left panel) and V2L unit (black line; right panel) and the predicted firing rate based on the models, including the different behavioral parameters (colored lines). F) Mean encoding quality across all AC (blue) and V2L (red) units using single behavioral factors as predictor and using all predictors. Error bars as in (A). G) Mean improvement in encoding quality across all single units following the addition of the indicated behavioral factor to a model already containing all other factors. Error bars are 95% confidence bounds. H) The relationship between encoding quality of individual single units when all predictors are considered and the encoding quality when all predictors except linear position (left) or head direction (right) are considered. Points above the diagonal belong to units with improved encoding due to the inclusion of linear position/head direction which cannot be attributed to any other included factor. Blue: AC units, red: V2L units. Diagrams to the right and top of the main scatterplot show the empirical distributions of the data depicted in the scatterplots projected onto a single dimension.

factors, such as stimulus availability, choice, and reward, may have influenced these results because, during track running, they are closely associated with both position and head direction in space and/or time (Zaidel et al. 2017). Re-analysis of conditional information, now on data in which the period of stimulus presentation  $\pm 1$  s was eliminated, showed a decrease in the amount of information on  $\theta_{\text{head}}$  conditional on position, which was

nevertheless still significantly larger than 0 for both areas (cMI difference from 0, AC:  $P = 1.7 \cdot 10^{-4}$ , V2L:  $P < 10^{-15}$ , Wilcoxon's signed-rank tests; Supplementary Fig. S6A, cf. Fig. 4B). Similarly, re-analysis of encoding improvement with data from which the stimulus presentation period  $\pm 1$  s was eliminated yielded similar results compared with the original finding (Supplementary Fig. S6B, cf. Fig. 4G). Together, these results



**Fig. 5.** Reconstruction of spatial position from auditory and visual cortical populations and between-area correlation of errors. A) Confusion matrices indicating the performance of Bayesian decoding of linearized position from AC neuronal populations averaged across sessions. The probability that a sample of spiking activity is assigned by the decoder to the true location of the rat is coded in blue shades. B) As (A) but for V2L populations. C) Distribution of decoding errors (i.e. distances between the true and decoded position) across AC (blue) and V2L (red) sessions. D) Decoding error as a function of population size, obtained by randomly selecting units from the neuronal populations. Thin lines indicate means for individual sessions for AC (blue) and V2L (red). Thick lines indicate means over sessions. E) Correlations of measured and shuffled (ordinate) instantaneous decoding errors in the X-dimension of the maze. Shuffling was performed inside the same spatial bin and within the same running speed range. F) Same as (E) but for the Y-dimension. G) The size and direction of instantaneous errors were correlated between AC and V2L. We computed the joint probability distribution of decoding errors along the X-direction in AC and V2L and subtracted from it the joint distribution computed on shuffled data, to obtain a relative probability ( $\Delta P$ ; a value of +0.3 indicates that a particular error is 30% more likely to be observed in actual data than in shuffled data). H) Same as (G) but for the Y-dimension.

suggest that position and, to a lesser degree, head direction are important drivers of firing activity of neurons in AC and V2L. Along with position and head direction, additional subject-induced sensations may have occurred, potentially contributing to spike-coded information on these 2 behavioral parameters.

### Coherent mapping of location in auditory and visual cortical populations

If the neuronal populations in AC and V2L code the animal's track position, it should be possible to infer the position of the rat from the collective neuronal activity. We used a Bayesian decoder to test whether spatial location can indeed be predicted from population activity (sessions with >16 neurons/area were included). Decoding of position was successful for both AC and V2L, because the chance that a decoded position matched the true position of the rat was larger than the chance that it matched any other location on the track (i.e. clear diagonal structure in Fig. 5A and B; AC:  $n = 15$  sessions; V2L:  $n = 12$  sessions). The distributions of decoding errors, which are the Euclidean distances between the true and decoded positions, clustered near 0 m for both areas (Fig. 5C). V2L showed, on average, higher proportions of smaller errors than AC, indicating that decoding spatial position from V2L was more accurate than from AC (V2L error: 0.11 m, 0.10–0.42; AC error: 0.29 m, 0.26–0.37; median and interquartile range, in meters).

Increasing the number of neurons included in the decoding analysis decreased the mean decoding error for both areas. We considered several population sizes, which were similar for both areas (AC: 16–38 neurons, V2L: 16–40 neurons), but bootstrap analysis showed no signs of plateau performance with progressive

population sizes (Fig. 5D). Although decoding from both areas likely could be improved had more neurons been recorded, at identical population size, the decoding performance for V2L was better than for AC both for the mean across sessions and for most individual sessions. Furthermore, the average performance of the decoder was virtually independent from stimulus modality, reward, and trial history (Supplementary Fig. S7), which is in agreement with Fig. 3E.

To investigate whether locations are encoded coherently across both areas, we calculated the correlation in instantaneous decoding error between AC and V2L. Instead of only taking the magnitude of the error into account, errors in the x and y dimensions of the maze were considered separately and their directionality was preserved. When decoding errors were computed for sessions which contained >16 simultaneously recorded neurons in each area, we found significant correlations ( $P < 0.001$  for all 11 sessions). This was also the case when including all sessions where at least 1 of the 2 areas provided 16 neurons, again indicating that AC and V2L encode locations coherently. However, these correlations may be confounded if certain locations are represented more accurately than others in both areas, or if firing patterns are subjected to a common modulation by locomotion speed. To control for these possibilities, we shuffled the data across time points at which the animal was in the same spatial bin and running within the same range of speeds. The correlations between the shuffled decoding errors were significantly lower than the observed correlations (x-direction,  $P = 0.01$ , y-direction  $P = 0.02$ , Wilcoxon's signed-rank test). This difference remained significant when we included all sessions where at least 1 area provided 16 neurons (x-direction,  $P = 0.001$ , y-direction,  $P = 0.01$ ).



The residual decoding errors, obtained by subtracting the shuffled distribution from the observed joint distributions, displayed a diagonal structure (Fig. 5G and H), indicating that representations of location in AC and V2L remain coherent even when the position cannot be decoded accurately from the population of either area. This coherency exceeds what would be expected from a common influence of speed. In addition to the mean across sessions (Fig. 5G and H), these results also held for individual rats (Supplementary Fig. S8). It is unlikely that error correlations can be accounted for by decoding artifacts in sessions with poor decoding performance because there was no evidence of a positive relationship between instantaneous decoding error correlations and average decoding error; rather, their linear correlations were not significant (Supplementary Fig. S9).

## Discussion

We showed location- and head direction-correlated firing patterns in large proportions of neurons in AC and V2L of freely moving rats navigating a figure-8 maze. This activity takes part in coherent representations across areas, as indicated by highly correlated firing-field densities (Fig. 3A–C) and correlated errors in reconstructed position (Fig. 5). A first implication of these results is that neural representations bound to subject location exist in the AC, thus outside the visual system. Such activity may be expected in V2L because of similar observations in V1 (Ji and Wilson 2007; Haggerty and Ji 2015; Fiser et al. 2016; Saleem et al. 2018; ) and egocentric trajectory correlates in the spiking activity of posterior parietal cortex (McNaughton et al. 1994; Nitz 2006; Whitlock et al. 2012; Krumin et al. 2018). These areas share anatomical borders with V2L and bidirectional, monosynaptic connectivity (Miller and Vogt 1984; McNaughton et al. 1994; Haggerty and Ji 2015; Krumin et al. 2018). Previously, spatial modulation was also found in higher visual areas of the mouse cortex, but in this study, animals were head-restrained, thus lacking changes in vestibular inputs (Diamanti et al. 2021). That AC neurons also show location-tuning is more surprising because AC activity has hitherto not been associated with animal location. However, in many mammals, this area is required for sound localization (Heffner 1978; Thompson and Cortez 1983; Jenkins and Merzenich 1984; Kavanagh and Kelly 1987) and contains neurons tuned to the location of sound sources (Middlebrooks and Pettigrew 1981; Town et al. 2017; Wang et al. 2019). In addition to visual cortex and AC, indications for spatial mapping in primary somatosensory cortex were presented by Long and Zhang (2021).

### Nature and function of location-correlated firing in sensory cortices

In addition to single or multiple peaks in single-cell firing rates tessellating the entire maze, we found that AC and V2L spike patterns were spatially stable across trials and coded significant amounts of information on animal position (Figs. 2 and 4). Our maze harbored repetitive elements requiring the same local behavior, such as directional body turns or running along straight maze stretches, whereas many cells showed single firing-rate peaks and thus did not reflect these repetitions (Fig. 2). Furthermore, the encoding analysis revealed animal position as the strongest predictor of AC and V2L activity even after correcting for head direction (Fig. 4). Moreover, subject position could be inferred from AC and V2L ensemble activity using Bayesian decoding (Fig. 5). Despite these indications, we argue that neither our current findings nor previous results on V1

(Ji and Wilson 2007; Haggerty and Ji 2015; Fiser et al. 2016; Saleem et al. 2018) necessarily imply coding of (allocentric) position *per se*, because proving this would require additional experimental manipulations to establish that location-selective activity is independent of the locomotion direction through a location and tolerates the manipulation of local sensory cues (Speakman and O'Keefe 1990; Knierim 2002; Leutgeb et al. 2005; Lansink et al. 2012; Wilber et al. 2014).

We propose the more general alternative that sensory cortical areas integrate modality-specific evidence with information from other sensory, motor, and association areas to generate sequential representations, which are not merely sensitive to local sensorimotor cues but also to contextual elements (which may comprise spatial and also other task-relevant elements, such as reward proximity, task rule execution, etc.). For instance, activation of visual cortex neurons depends on the subject's field of view which in turn depends on animal position and head direction (cf. Haggerty and Ji 2015). This combination of view, position, and head direction may give rise to a predictive visual representation in the cortex, which will be compared to further visual input to compute error signals, as posited by predictive processing models (Rao and Ballard 1999; Friston 2005; Fiser et al. 2016; Keller and Mrsic-Flogel 2018; Pennartz et al. 2019). Although this proposal needs further testing, it is generally supported by the literature documenting extensive corticocortical connections (Felleman and Van Essen 1991; Groenewegen and Uylings 2000; Bizley et al. 2007; Budinger and Scheich 2009; Campi et al. 2010; Laramée et al. 2011; D'Souza et al. 2016; Leinweber et al. 2017; Gămănuț et al. 2018; Harris et al. 2019), contributions to neural coding in sensory cortices by nonsensory parameters (Shuler and Bear 2006; Goltstein et al. 2013; Namboodiri et al. 2015; Pakan et al. 2018), and auditory-visual cortical interactions (Iurilli et al. 2012; Ibrahim et al. 2016; Meijer et al. 2017; Morrill and Hasenstaub 2018; Knöpfel et al. 2019; Meijer et al. 2020).

This view does not conflict with a potential role for sensory cortices in updating spatial (e.g. hippocampal) representations in a bottom-up fashion, or with the navigational system contributing to top-down sensory predictions (Fournier et al. 2020). However, the hypothesis of the hippocampus causally driving spatial coding in V1 (Saleem et al. 2018) faces the issue that the hippocampus proper does not directly project to the sensory cortices, and its output is transformed by the synaptic matrices of intermediate parahippocampal regions on which the nonhippocampal structures also converge (Witter et al. 2000; Furtak et al. 2007; Rusu and Pennartz 2020). Instead of guiding spatial navigation, the bidirectional, cortico-hippocampal circuitry may subservise declarative memory consolidation (O'Keefe and Nadel 1978; Squire 1986; Eichenbaum 2000; McGaugh 2000; Rusu and Pennartz 2020). Although the medial entorhinal cortex seems to be better situated to issue spatial-mapping signals, as it is known to project to the entire neocortical mantle (Swanson and Kohler 1986; Insausti et al. 1997; cf. Sargolini et al. 2006), the hippocampus may nonetheless act as a driver for spatial coding in neocortex, as hippocampal lesions were shown to impair the development of location-specificity of neocortical cells (Esteves et al. 2021).

### Sensitivity to head direction in auditory and secondary visual cortex

For 33% of neurons in AC and 49% in V2L, the information conveyed on head direction was significant after correcting for position (Fig. 4B and D). This finding is consistent with results of

Guitchounts et al. (2020) in rat V1, albeit that they did not correct for changes in self-position within the home cage. Recently, Mimica et al. (2022) also reported coding of head-orienting features in auditory and visual cortices of freely foraging rats. While their results are consistent with our findings as concerns head orientation, Mimica et al. focused on neural correlates of elementary behaviors, not on maze running or self-location. A further reason why they did not report prominent location-selective firing may lie in differences in behavioral settings, as they studied rats foraging for randomly scattered food crumbs, whereas self-location in our task is deemed relevant for mapping of, e.g., proximity to reward sites and stimulus displays.

As is the case for location, these overall findings relate to how sensory-specific states coded by AC and V2L depend on head direction and how these sensory cortices may use this sensitivity to emit head direction signals to target areas. Basic head direction signaling is regulated by the vestibular nuclei carrying information about the head's motion relative to external space (Cullen 2014) and, in rodents, about static neck position (Barresi et al. 2013; Medrea and Cullen 2013). Vestibular input is a key contributor to the brain's head direction system (Taube 2007), including anterior thalamus (Taube 1995), postsubiculum (Taube et al. 1990), and medial entorhinal cortex (Sargolini et al. 2006). Vestibular information was shown to reach the visual system directly and indirectly, viz. via the retrosplenial cortex (Vélez-Fort et al. 2018), which was proposed to mediate between the sensory cortices and the head direction system of the temporal lobe (Page and Jeffery 2018). In addition to head direction signaling based on vestibular inputs, it should be noted that, in the current task, head direction sensitivity may also depend on other sensory inputs associated with maze running (e.g. visual and proprioceptive).

Whether head direction signaling in AC and V2L has a causal role in distributing head direction information to target areas is up for further research. There is evidence to support the hypothesis that areas along the cortical hierarchy may use both allocentric and egocentric representations, with a gradient of egocentric-to-allocentric processing from sensory to temporal cortices, as also proposed for parietal-retrosplenial circuitry, where V2L is sometimes included as part of parietal cortex (Wilber et al. 2014; Chen et al. 2018; Clark et al. 2018). Recent studies, however, question the validity of such a gradient, pointing to egocentric coding in several temporal lobe areas in addition to allocentric bearings (Wang et al. 2018; LaChance et al. 2019; Kunz et al. 2021).

### Comparison between auditory and secondary visual cortex

We were particularly struck by the broad, qualitative similarities between AC and V2L spiking patterns. Both areas showed comparable levels of spatial stability, amounts of firing fields per unit (Fig. 2) and spatial distributions of firing fields (Fig. 3). Moreover, the contributions to predictions of firing-rate patterns from position, head direction, and other factors were highly similar for AC and V2L (Fig. 4). These observations lend support to the hypothesis that coordinated representations are a general feature of sensory cortical areas.

Neural coding in AC and V2L also showed interesting quantitative differences, which consistently point to a higher SI content (Fig. 2), stronger correlations with location and head direction (Fig. 4), and better position reconstruction in V2L than AC. How this greater accuracy arises in V2L is unknown, but it may relate to a larger amount of spatially informative visual cues in our maze compared to auditory cues and to more consistent changes in visual inputs due to self-motion than to (self-induced) auditory inputs.

### Significance of crossareal coordination in cortical mapping of location and head direction

Arguably, our most novel result is the highly coherent nature of location-selective representations across sensory domains of the cortex (Figs. 3 and 5), suggesting that location- and head direction-sensitive mappings in auditory and visual cortical systems are not computed independently but are coordinated. Such coordination of context-dependent sensory mappings offers the computational advantage that evidence from multiple sensory modalities can be combined to improve estimates of the subject's task-relevant state, although this comes at the expense of error sharing.

Previous studies on auditory-visual interactions were predominantly guided by the theoretical framework of multisensory cue integration, whereby evidence for the detection of a stimulus in 1 modality is augmented by another modality (Fetsch et al. 2013; Stein et al. 2014; Meijer et al. 2019). Our findings go beyond integration of discrete sensory cues and instantaneous sensory states as they indicate a crossmodal coordination of context-sensitive representations. In our proposal, this common mapping subserves the construction of a multimodal survey of the subject's current situation, thereby enabling efficient goal-directed action planning and execution (Pennartz 2018). As argued by Hawkins et al. (2017), self-parameters, including head direction and position within a task sequence, are key priors in determining how and when to undertake goal-directed actions. In line with predictive processing, knowledge of these self-parameters is required to interpret novel sensory inputs and to anticipate sensory outcomes of actions (cf. Schürmann et al. 2019).

### Conclusion

In conclusion, both the auditory cortex and secondary visual cortex map the subject's position and head direction, and both areas coherently generate this neural code during spatial navigation. Together, these sensory cortical areas are proposed to integrate modality-specific evidence with information from other neocortical areas to sustain context-sensitive representations.

### Acknowledgments

We thank Kenneth D. Harris and A. David Redish for the availability of unit isolation software KlustaKwik and MClust, respectively. We also gratefully acknowledge the use of a Seaborn Python library in our study (v0.9.0, DOI 10.5281/zenodo.1313201, 2018). The work of the Technology Center at the University of Amsterdam for building the recording setup and the tetrode microdrives is highly appreciated.

### Supplementary material

Supplementary material is available at *Cerebral Cortex* online.

### Data availability

Data are available on request.

### Authors' contributions

Paul E.C. Mertens (Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Software, Validation, Visualization, Writing—original draft), Pietro Marchesi (Formal analysis, Software, Validation, Visualization, Writing—original draft, Writing—review & editing), Thijs R. Ruikes (Data curation, Formal analysis, Software, Validation, Visualization, Writing—review & editing), Quincy Krijger (Investigation), Matthijs oude Lohuis (Investigation), Cyriel M.A. Pennartz

(Conceptualization, Funding acquisition, Project administration, Supervision, Writing—original draft, Writing—review & editing, and Carien S. Lansink (Conceptualization, Supervision, Writing—original draft)

## Funding

This work was supported by the Netherlands Organization for Scientific Research VENI Grant (863.11.010 to CSL) and by the European Union's Horizon 2020 Framework Program for Research and Innovation under the Human Brain Project SGA3 (Specific Grant Agreement number 945539 to CMAP).

*Conflict of interest statement:* None declared.

## References

- Ayaz A, Saleem AB, Schölvinck ML, Carandini M. Locomotion controls spatial integration in mouse visual cortex. *Curr Biol*. 2013;23:890–894.
- Barresi M, Grasso C, Li Volsi G, Manzoni D. Effects of body to head rotation on the labyrinthine responses of rat vestibular neurons. *Neuroscience*. 2013;244:134–146.
- Benjamin AS, Fernandes HL, Tomlinson T, Ramkumar P, Versteeg C, Chowdhury RH, Miller LE, Kording KP. Modern machine learning as a benchmark for fitting neural responses. *Front Comput Neurosci*. 2018;12.
- Bizley JK, Nodal FR, Bajo VM, Nelken I, King AJ. Physiological and anatomical evidence for multisensory interactions in auditory cortex. *Cereb Cortex*. 2007;17:2172–2189.
- Bos JJ, Vinck M, van Mourik-Donga LA, Jackson JC, Witter MP, Pennartz CMA. Perirhinal firing patterns are sustained across large spatial segments of the task environment. *Nat Commun*. 2017;8:15602.
- Bos JJ, Vinck M, Marchesi P, Keestra A, van Mourik-Donga LA, Jackson JC, Verschure PFMJ, Pennartz CMA. Multiplexing of information about self and others in hippocampal ensembles. *Cell Rep*. 2019;29:3859–3871.e3856.
- Brosch M, Selezneva E, Scheich H. Nonauditory events of a behavioral procedure activate auditory cortex of highly trained monkeys. *J Neurosci*. 2005;25:6797–6806.
- Budinger E, Scheich H. Anatomical connections suitable for the direct processing of neuronal information of different modalities via the rodent primary auditory cortex. *Hear Res*. 2009;258:16–27.
- Campi KL, Bales KL, Grunewald R, Krubitzer L. Connections of auditory and visual cortex in the prairie vole (*Microtus ochrogaster*): evidence for multisensory processing in primary sensory areas. *Cereb Cortex*. 2010;20:89–108.
- Chen X, DeAngelis GC, Angelaki DE. Flexible egocentric and allocentric representations of heading signals in parietal cortex. *Proc Natl Acad Sci U S A*. 2018;115:E3305–E3312.
- Clark BJ, Simmons CM, Berkowitz LE, Wilber AA. The retrosplenial-parietal network and reference frame coordination for spatial navigation. *Behav Neurosci*. 2018;132:416–429.
- Clemens LE, Jansson EKH, Portal E, Riess O, Nguyen HP. A behavioral comparison of the common laboratory rat strains Lister hooded, Lewis, Fischer 344 and Wistar in an automated homecage system. *Genes Brain Behav*. 2014;13:305–321.
- Cullen KE. The neural encoding of self-generated and externally applied movement: implications for the perception of self-motion and spatial memory. *Front Integr Neurosci*. 2014;7:108.
- Davidson TJ, Kloosterman F, Wilson MA. Hippocampal replay of extended experience. *Neuron*. 2009;63(4):497–507.
- Diamanti EM, Reddy CB, Schröder S, Muzzu T, Harris KD, Saleem AB, Carandini M. Spatial modulation of visual responses arises in cortex with active navigation. *elife*. 2021;10:e63705.
- DiCarlo JJ, Cox DD. Untangling invariant object recognition. *Trends Cogn Sci*. 2007;11(8):333–341.
- D'Souza RD, Meier AM, Bista P, Wang Q, Burkhalter A. Recruitment of inhibition and excitation across mouse visual cortex depends on the hierarchy of interconnecting areas. *elife*. 2016;5:e19332.
- Eichenbaum H. A cortical-hippocampal system for declarative memory. *Nat Rev Neurosci*. 2000;1:41–50.
- Esteves IM, Chang H, Neumann AR, Sun J, Mohajerani MH, McNaughton BL. Spatial information encoding across multiple neocortical regions depends on an intact hippocampus. *J Neurosci*. 2021;41(2):307–319.
- Felleman DJ, Van Essen DC. Distributed hierarchical processing in the primate cerebral cortex. *Cereb Cortex*. 1991;1:1–47.
- Fetsch CR, DeAngelis GC, Angelaki DE. Bridging the gap between theories of sensory cue integration and the physiology of multisensory neurons. *Nat Rev Neurosci*. 2013;14:429–442.
- Fiser A, Mahringer D, Oyibo HK, Petersen AV, Leinweber M, Keller GB. Experience-dependent spatial expectations in mouse visual cortex. *Nat Neurosci*. 2016;19:1658–1664.
- Fournier J, Saleem AB, Diamanti EM, Wells MJ, Harris KD, Carandini M. Mouse visual cortex is modulated by distance traveled and by theta oscillations. *Curr Biol*. 2020;30:3811–3817.e3816.
- Friston K. A theory of cortical responses. *Philos Trans R Soc Lond B Biol Sci*. 2005;360:815–836.
- Furtak SC, Wei SM, Agster KL, Burwell RD. Functional neuroanatomy of the parahippocampal region in the rat: the perirhinal and postrhinal cortices. *Hippocampus*. 2007;17:709–722.
- Gămănuț R, Kennedy H, Toroczka Z, Ercsey-Ravasz M, Van Essen DC, Knoblauch K, Burkhalter A. The mouse cortical connectome, characterized by an ultra-dense cortical graph, maintains specificity by distinct connectivity profiles. *Neuron*. 2018;97:698–715.e610.
- Gavornik JP, Bear MF. Learned spatiotemporal sequence recognition and prediction in primary visual cortex. *Nat Neurosci*. 2014;17:732–737.
- Goltstein PM, Coffey EBJ, Roelfsema PR, Pennartz CMA. In vivo two-photon Ca2+ imaging reveals selective reward effects on stimulus-specific assemblies in mouse visual cortex. *J Neurosci*. 2013;33:11540–11555.
- Gray CM, Maldonado PE, Wilson M, McNaughton B. Tetrodes markedly improve the reliability and yield of multiple single-unit isolation from multi-unit recordings in cat striate cortex. *J Neurosci Methods*. 1995;63:43–54.
- Groenewegen HJ, Uylings HB. The prefrontal cortex and the integration of sensory, limbic and autonomic information. *Prog Brain Res*. 2000;126:3–28.
- Guitchoants G, Masís J, Wolff SBE, Cox D. Encoding of 3D head orienting movements in the primary visual cortex. *Neuron*. 2020;108:512–525.e514.
- Haggerty DC, Ji D. Activities of visual cortical and hippocampal neurons co-fluctuate in freely moving rats during spatial behavior. *elife*. 2015;4:e08902.
- Harris JA, Mihalas S, Hirokawa KE, Whitesell JD, Choi H, Bernard A, Bohn P, Caldejon S, Casal L, Cho A, et al. Hierarchical organization of cortical and thalamic connectivity. *Nature*. 2019;575:195–202.
- Hawkins J, Ahmad S, Cui Y. A theory of how columns in the neocortex enable learning the structure of the world. *Front Neural Circuits*. 2017;11:81.
- Heffner H. Effect of auditory cortex ablation on localization and discrimination of brief sounds. *J Neurophysiol*. 1978;41:963–976.

- Holm S. A simple sequentially rejective multiple test procedure. *Scand J Stat.* 1979;6:65–70.
- Hubel DH, Wiesel TN. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J Physiol.* 1962;160:106–154.
- Ibrahim LA, Mesik L, Ji X-y, Fang Q, Li H-F, Li Y-T, Zingg B, Zhang LI, Tao HW. Cross-modality sharpening of visual cortical processing through layer-1-mediated inhibition and disinhibition. *Neuron.* 2016;89:1031–1045.
- Insausti R, Herrero MT, Witter MP. Entorhinal cortex of the rat: cytoarchitectonic subdivisions and the origin and distribution of cortical efferents. *Hippocampus.* 1997;7:146–183.
- Iurilli G, Ghezzi D, Olcese U, Lassi G, Nazzaro C, Tonini R, Tucci V, Benfenati F, Medini P. Sound-driven synaptic inhibition in primary visual cortex. *Neuron.* 2012;73:814–828.
- Jaramillo S, Zador AM. The auditory cortex mediates the perceptual effects of acoustic temporal expectation. *Nat Neurosci.* 2011;14:246–251.
- Jenkins WM, Merzenich MM. Role of cat primary auditory cortex for sound-localization behavior. *J Neurophysiol.* 1984;52:819–847.
- Ji D, Wilson MA. Coordinated memory replay in the visual cortex and hippocampus during sleep. *Nat Neurosci.* 2007;10:100–107.
- Kadir SN, Goodman DFM, Harris KD. High-dimensional cluster analysis with the masked EM algorithm. *Neural Comput.* 2014;26:2379–2394.
- Kavanagh GL, Kelly JB. Contribution of auditory cortex to sound localization by the ferret (*Mustela putorius*). *J Neurophysiol.* 1987;57:1746–1766.
- Keller GB, Mrsic-Flogel TD. Predictive processing: a canonical cortical computation. *Neuron.* 2018;100:424–435.
- Knierim JJ. Dynamic interactions between local surface cues, distal landmarks, and intrinsic circuitry in hippocampal place cells. *J Neurosci.* 2002;22:6254–6264.
- Knöpfel T, Sweeney Y, Radulescu CI, Zabouri N, Doostdar N, Clopath C, Barnes SJ. Audio-visual experience strengthens multisensory assemblies in adult mouse visual cortex. *Nat Commun.* 2019;10:5684.
- Krumin M, Lee JJ, Harris KD, Carandini M. Decision and navigation in mouse parietal cortex. *elife.* 2018;7:e42583.
- Kunz L, Brandt A, Reinacher PC, Staresina BP, Reifensstein ET, Weidemann CT, Herweg NA, Patel A, Tsitsiklis M, Kempter R, et al. A neural code for egocentric spatial maps in the human medial temporal lobe. *Neuron.* 2021;109(17):2781–2796.
- LaChance PA, Todd TP, Taube JS. A sense of space in postrhinal cortex. *Science.* 2019;365(6449):eaax4192.
- Lansink CS, Bakker M, Buster W, Lankelma J, van der Blom R, Westdorp R, Joosten RNJMA, McNaughton BL, Pennartz CMA. A split microdrive for simultaneous multi-electrode recordings from two brain areas in awake small animals. *J Neurosci Methods.* 2007;162:129–138.
- Lansink CS, Jackson JC, Lankelma JV, Ito R, Robbins TW, Everitt BJ, Pennartz CM. Reward cues in space: commonalities and differences in neural coding by hippocampal and ventral striatal ensembles. *J Neurosci.* 2012;32(36):12444–12459.
- Laramée ME, Kurotani T, Rockland KS, Bronchti G, Boire D. Indirect pathway between the primary auditory and visual cortices through layer V pyramidal neurons in V2L in mouse and the effects of bilateral enucleation. *Eur J Neurosci.* 2011;34:65–78.
- Leinweber M, Ward DR, Sobczak JM, Attinger A, Keller GB. A sensorimotor circuit in mouse cortex for visual flow predictions. *Neuron.* 2017;95:1420–1432.e1425.
- Leutgeb S, Leutgeb JK, Barnes CA, Moser EI, McNaughton BL, Moser MB. Independent codes for spatial and episodic memory in hippocampal neuronal ensembles. *Science.* 2005;309:619–623.
- Lizier JT. JIDT: an information-theoretic toolkit for studying the dynamics of complex systems. *Front Robot AI.* 2014:1.
- Long X, Zhang SJ. A novel somatosensory spatial navigation system outside the hippocampal formation. *Cell Res.* 2021;31(6):649–663.
- Lopes G, Bonacchi N, Frazão J, Neto JP, Atallah BV, Soares S, Moreira L, Matias S, Itskov PM, Correia PA, et al. Bonsai: an event-based framework for processing and controlling data streams. *Front Neuroinform.* 2015:9.
- Louie K, Wilson MA. Temporally structured replay of awake hippocampal ensemble activity during rapid eye movement sleep. *Neuron.* 2001;29(1):145–156.
- McGaugh JL. Memory—a century of consolidation. *Science.* 2000;287:248–251.
- McNaughton BL, Mizumori SJ, Barnes CA, Leonard BJ, Marquis M, Green EJ. Cortical representation of motion during unrestrained spatial navigation in the rat. *Cereb Cortex.* 1994;4:27–39.
- Medrea I, Cullen KE. Multisensory integration in early vestibular processing in mice: the encoding of passive vs. active motion. *J Neurophysiol.* 2013;110:2704–2717.
- Meijer GT, Montijn JS, Pennartz CMA, Lansink CS. Audiovisual modulation in mouse primary visual cortex depends on cross-modal stimulus configuration and congruency. *J Neurosci.* 2017;37:8783–8796.
- Meijer GT, Mertens PEC, Pennartz CMA, Olcese U, Lansink CS. The circuit architecture of cortical multisensory processing: distinct functions jointly operating within a common anatomical network. *Prog Neurobiol.* 2019;174:1–15.
- Meijer GT, Marchesi P, Mejias JF, Montijn JS, Lansink CS, Pennartz CM. Neural correlates of multisensory detection behavior: comparison of primary and higher-order visual cortex. *Cell Rep.* 2020;31:107636.
- Middlebrooks JC, Pettigrew JD. Functional classes of neurons in primary auditory cortex of the cat distinguished by sensitivity to sound location. *J Neurosci.* 1981;1:107–120.
- Miller KD. Canonical computations of cerebral cortex. *Curr Opin Neurobiol.* 2016;37:75–84.
- Miller MW, Vogt BA. Direct connections of rat visual cortex with sensory, motor, and association cortices. *J Comp Neurol.* 1984;226(2):184–202.
- Mimica B, Tombaz T, Battistin C, Fuglstad JG, Dunn BA, Whitlock JR. Behavioral decomposition reveals rich encoding structure employed across neocortex. 2022. bioRxiv.
- Morrill RJ, Hasenstaub AR. Visual information present in infragranular layers of mouse auditory cortex. *J Neurosci.* 2018;38:2854–2862.
- Namoodiri VMK, Huertas MA, Monk KJ, Shouval HZ, Shuler H, Marshall G. Visually cued action timing in the primary visual cortex. *Neuron.* 2015;86:319–330.
- Newby FD, DiGirolamo M, Cotsonis GA, Kutner MH. Model of spontaneous obesity in aging male Wistar rats. *Am J Phys.* 1990:259:R1117–R1125.
- Niell CM, Stryker MP. Modulation of visual responses by behavioral state in mouse visual cortex. *Neuron.* 2010;65:472–479.
- Nitz DA. Tracking route progression in the posterior parietal cortex. *Neuron.* 2006;49:747–756.
- Niwa M, Johnson JS, O'Connor KN, Sutter ML. Activity related to perceptual judgment and action in primary auditory cortex. *J Neurosci.* 2012;32:3193–3210.
- O'Keefe J, Nadel L. *The hippocampus as a cognitive map*. Clarendon Press, Oxford; 1978.



- Olcese U, Bos JJ, Vinck M, Lankelma JV, Mourik-Donga LB, Schlumm F, Pennartz CMA. Spike-based functional connectivity in cerebral cortex and hippocampus: loss of global connectivity is coupled to preservation of local connectivity during non-REM sleep. *J Neurosci*. 2016;36:7676–7692.
- Page HJL, Jeffery KJ. Landmark-based updating of the head direction system by retrosplenial cortex: a computational model. *Front Cell Neurosci*. 2018;12:191.
- Pakan JMP, Currie SP, Fischer L, Rochefort NL. The impact of visual cues, reward, and motor feedback on the representation of behaviorally relevant spatial locations in primary visual cortex. *Cell Rep*. 2018;24:2521–2528.
- Paxinos G, Watson C. *The rat brain in stereotaxic coordinates*. 6th ed. Academic Press, New York, USA; 2007.
- Pennartz CMA. Consciousness, representation, action: the importance of being goal-directed. *Trends Cogn Sci*. 2018;22:137–153.
- Pennartz CMA, Dora S, Muckli L, Lorteije JAM. Towards a unified view on pathways and functions of neural recurrent processing. *Trends Neurosci*. 2019;42:589–603.
- Rao RPN, Ballard DH. Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nat Neurosci*. 1999;2:79.
- Rusu SI, Pennartz CMA. Learning, memory and consolidation mechanisms for behavioral control in hierarchically organized cortico-basal ganglia systems. *Hippocampus*. 2020;30:73–98.
- Saleem AB, Diamanti EM, Fournier J, Harris KD, Carandini M. Coherent encoding of subjective spatial position in visual cortex and hippocampus. *Nature*. 2018;562:124.
- Sargolini F, Fyhn M, Hafting T, McNaughton BL, Witter MP, Moser M-B, Moser EI. Conjunctive representation of position, direction, and velocity in entorhinal cortex. *Science*. 2006;312:758–762.
- Scheich H, Brechmann A, Brosch M, Budinger E, Ohl FW. The cognitive auditory cortex: task-specificity of stimulus representations. *Hear Res*. 2007;229:213–224.
- Schmitzer-Torbert N, Jackson J, Henze D, Harris K, Redish AD. Quantitative measures of cluster quality for use in extracellular recordings. *Neuroscience*. 2005;131:1–11.
- Schürmann T, Vogt J, Christ O, Beckerle P. The Bayesian causal inference model benefits from an informed prior to predict proprioceptive drift in the rubber foot illusion. *Cogn Process*. 2019;20(4):447–457.
- Shuler MG, Bear MF. Reward timing in the primary visual cortex. *Science*. 2006;311:1606–1609.
- Skaggs WE, McNaughton BL, Gothard KM, and Markus EJ. 1992. An information-theoretic approach to deciphering the hippocampal code. In: *Proceedings of the 5th International Conference on Neural Information Processing Systems*. Morgan Kaufmann Publishers Inc., San Francisco; 1992. p. 1030–1037.
- Speakman A, O'Keefe J. Hippocampal complex spike cells do not change their place fields if the goal is moved within a cue controlled environment. *Eur J Neurosci*. 1990;2:544–555, 6.
- Squire LR. Mechanisms of memory. *Science*. 1986;232:1612–1619.
- Stein BE, Stanford TR, Rowland BA. Development of multisensory integration from the perspective of the individual neuron. *Nat Rev Neurosci*. 2014;15:520–535.
- Stringer C, Pachitariu M, Steinmetz N, Reddy CB, Carandini M, Harris KD. Spontaneous behaviors drive multidimensional, brainwide activity. *Science*. 2019;364:255.
- Swanson L, Kohler C. Anatomical evidence for direct projections from the entorhinal area to the entire cortical mantle in the rat. *J Neurosci*. 1986;6:3010–3023.
- Taube JS. Head direction cells recorded in the anterior thalamic nuclei of freely moving rats. *J Neurosci*. 1995;15:70–86.
- Taube JS. The head direction signal: origins and sensory-motor integration. *Annu Rev Neurosci*. 2007;30:181–207.
- Taube JS, Muller RU, Ranck JB. Head-direction cells recorded from the postsubiculum in freely moving rats. II. Effects of environmental manipulations. *J Neurosci*. 1990;10:436–447.
- Thompson GC, Cortez AM. The inability of squirrel monkeys to localize sound after unilateral ablation of auditory cortex. *Behav Brain Res*. 1983;8:211–216.
- Town SM, Brimijoin WO, Bizley JK. Egocentric and allocentric representations in auditory cortex. *PLoS Biol*. 2017;15:e2001878.
- Vélez-Fort M, Bracey EF, Keshavarzi S, Rousseau CV, Cossell L, Lenzi SC, Strom M, Margrie TW. A circuit for integration of head- and visual-motion signals in layer 6 of mouse primary visual cortex. *Neuron*. 2018;98:179–191.e176.
- Wang C, Chen X, Lee H, Deshmukh SS, Yoganarasimha D, Savelli F, Knierim JJ. Egocentric coding of external items in the lateral entorhinal cortex. *Science*. 2018;362(6417):945–949.
- Wang X, Liu J, Zhang J. Chronic unilateral hearing loss disrupts neural tuning to sound-source azimuth in the rat primary auditory cortex. *Front Neurosci*. 2019;13.
- Waskom 2021. v.0.9.0. 2018. <https://doi.org/10.5281/zenodo.1313201>.
- Waskom M. Seaborn: statistical data visualization. *J Open Source Softw*. 2021;6(60):3021. <https://doi.org/10.21105/joss.03021>.
- Whitlock JR, Pfuhl G, Dagslott N, Moser M-B, Moser EI. Functional split between parietal and entorhinal cortices in the rat. *Neuron*. 2012;73:789–802.
- Wilber AA, Clark BJ, Forster TC, Tatsuno M, McNaughton BL. Interaction of egocentric and world-centered reference frames in the rat posterior parietal cortex. *J Neurosci*. 2014;34:5431–5446.
- Witter MP, Naber PA, van Haeften T, Machielsen WC, Rombouts SA, Barkhof F, Scheltens P, Lopes da Silva FH. Cortico-hippocampal communication by way of parallel parahippocampal-subicular pathways. *Hippocampus*. 2000;10:398–410.
- Zaidel A, DeAngelis GC, Angelaki DE. Decoupled choice-driven and stimulus-related activity in parietal neurons may be misrepresented by choice probabilities. *Nat Commun*. 2017;8:715.