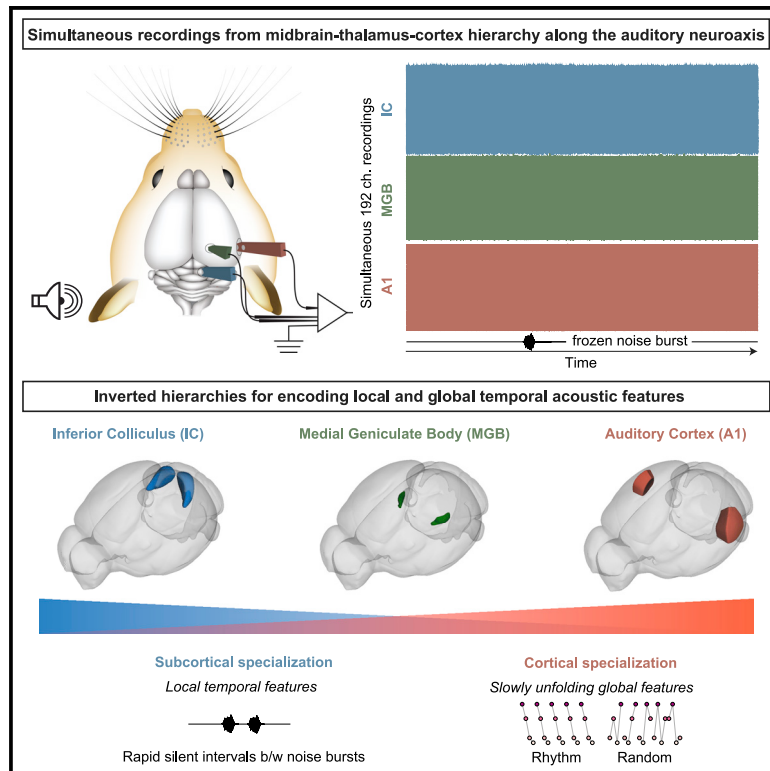


Current Biology

Inverted central auditory hierarchies for encoding local intervals and global temporal patterns

Graphical Abstract



Authors

Meenakshi M. Asokan,
Ross S. Williamson,
Kenneth E. Hancock, Daniel B. Polley

Correspondence

daniel_polley@meei.harvard.edu

In Brief

Asokan et al. show an emergent sensitivity for slow temporal rhythms at the level of the auditory cortex. Whereas single units in the mouse midbrain and thalamus encode local temporal intervals with high fidelity, cortical units integrate over extended periods to distinguish between random or regular interval sequences by adjusting spike timing.

Highlights

- Simultaneous unit recordings from the mouse auditory midbrain, thalamus, and cortex
- Subcortical stations decode short temporal intervals, not longer rhythmic patterns
- Cortical units decode slower changes in stimulus patterns, not short intervals
- Changes in stimulus regularity are reflected in cortical spike timing, not rate



Report

Inverted central auditory hierarchies for encoding local intervals and global temporal patterns

Meenakshi M. Asokan,^{1,2} Ross S. Williamson,^{1,3,4} Kenneth E. Hancock,^{1,3} and Daniel B. Polley^{1,3,5,*}¹Eaton-Peabody Laboratories, Massachusetts Eye and Ear Infirmary, Boston MA 02114 USA²Division of Medical Sciences, Harvard Medical School, Boston MA 02114 USA³Department of Otolaryngology - Head and Neck Surgery, Harvard Medical School, Boston MA 02114 USA⁴Present address: Department of Otolaryngology, University of Pittsburgh, Pittsburgh PA 15260⁵Lead contact*Correspondence: daniel_polley@meei.harvard.edu<https://doi.org/10.1016/j.cub.2021.01.076>

SUMMARY

In sensory systems, representational features of increasing complexity emerge at successive stages of processing. In the mammalian auditory pathway, the clearest change from brainstem to cortex is defined by what is lost, not by what is gained, in that high-fidelity temporal coding becomes increasingly restricted to slower acoustic modulation rates.^{1,2} Here, we explore the idea that sluggish temporal processing is more than just an inability for fast processing, but instead reflects an emergent specialization for encoding sound features that unfold on very slow timescales.^{3,4} We performed simultaneous single unit ensemble recordings from three hierarchical stages of auditory processing in awake mice – the inferior colliculus (IC), medial geniculate body of the thalamus (MGB) and primary auditory cortex (A1). As expected, temporal coding of brief local intervals (0.001 – 0.1 s) separating consecutive noise bursts was robust in the IC and declined across MGB and A1. By contrast, slowly developing (~1 s period) global rhythmic patterns of inter-burst interval sequences strongly modulated A1 spiking, were weakly captured by MGB neurons, and not at all by IC neurons. Shifts in stimulus regularity were not represented by changes in A1 spike rates, but rather in how the spikes were arranged in time. These findings show that low-level auditory neurons with fast timescales encode isolated sound features but not the longer gestalt, while the extended timescales in higher-level areas can facilitate sensitivity to slower contextual changes in the sensory environment.

RESULTS AND DISCUSSION

Our perception of music, speech and the surrounding soundscape all originate from patterns of vibrations in the cochlea. To encode these high-speed vibrations, evolutionary pressures designed an exquisite hydro/electro/mechanical force transducer, signal amplifier and frequency analyzer – the Organ of Corti – with sensory cells that can individually encode a 1×10^{12} range of signal amplitudes at rates up to several kilohertz. Neurons in the auditory brainstem synchronize action potential timing to frequency modulations in excess of 1 kHz and compute sub-millisecond discrepancies in the timing of sound waves that reach each ear. Each of these remarkable feats of high-speed processing are enabled by cellular morphologies, intrinsic electrical properties and synaptic architectures that are unique to the first stages of auditory processing and are indispensable for encoding the rapid acoustic building blocks of sound perception.⁵

On the other hand, higher-order structure in music, vocal communication, and auditory scenes are organized on inherently slow timescales (Figure 1A, top). For example, recognizing that a

scratched record is repeating the same short section of a song, or entraining the tapping of your foot to a musical rhythm, or integrating across the slower structures of vocal communication that organize the faster components (e.g., sentences or bouts of bird-song), all require a specialized selectivity for *slow* timescales. Here, we addressed whether low-level auditory neurons with fast timescales encode brief local features of sounds but not the longer gestalt, while the slower neural timescales in higher-level auditory areas may support the encoding of global sound patterns organized over longer timescales (Figure 1A, bottom).

Subcortical specialization for encoding rapid, local intervals

To test this idea, we used three high-density probes to make simultaneous recordings from hundreds of single units along the IC-MGB-A1 hierarchy in awake, head-fixed mice (Figure 1B). In response to a single 20 ms burst of frozen white noise, IC units fire short-latency spikes in rapid succession, whereas A1 units produced protracted spiking that persisted for tens of milliseconds after sound offset (Figure 1C). Neural timescales also increase across hierarchies of visual cortical areas, where it is

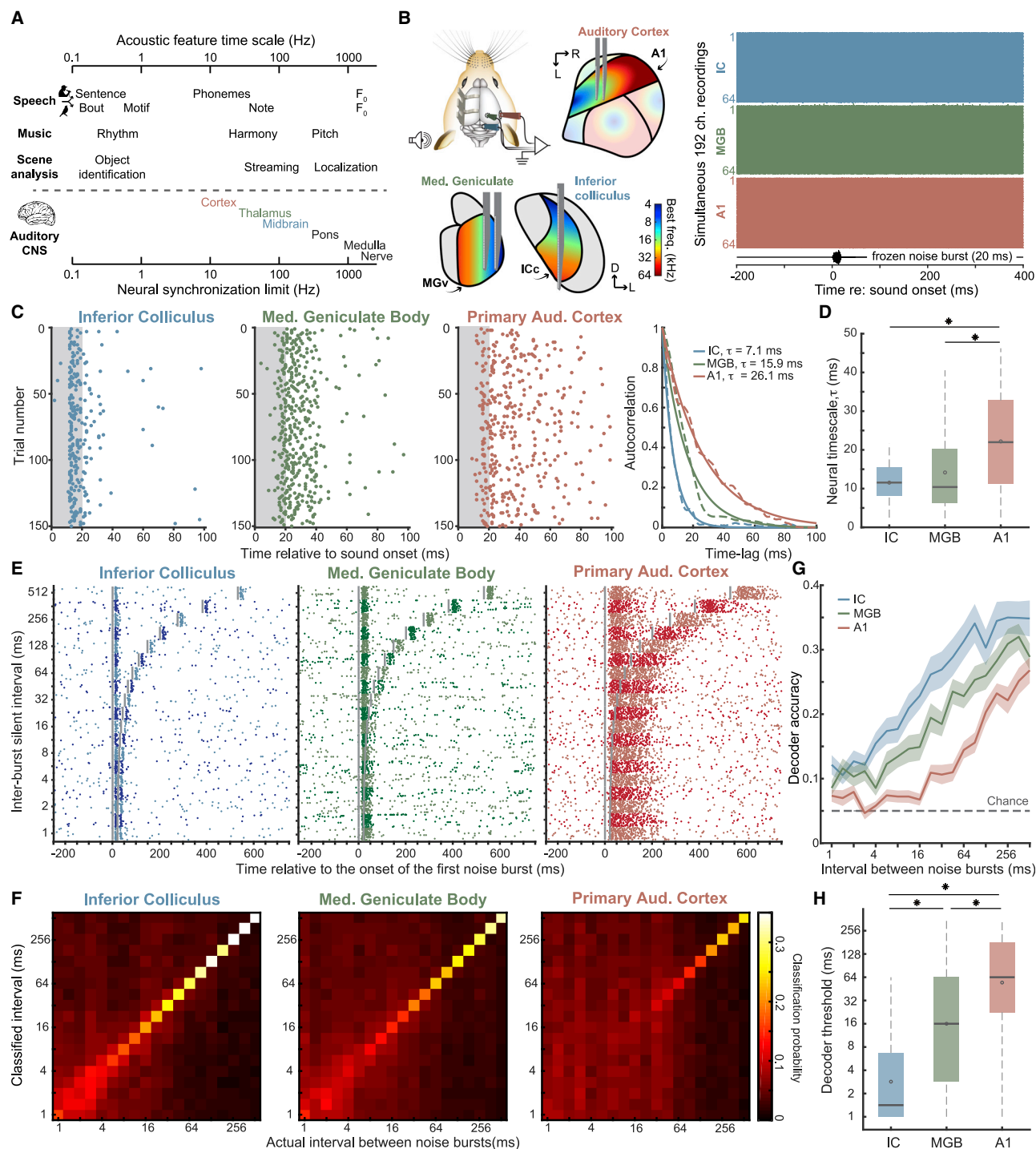


Figure 1. Across the midbrain to cortex hierarchy, neural timescales expand as temporal interval decoding accuracy declines

(A) Top: acoustic features supporting the perception of speech, music and auditory scene analysis are inherently organized on a wide range of timescales. Bottom: the typical synchronization limit for neurons at each stage of auditory processing. Each word is roughly centered on the typical synchronization limit to amplitude modulated sound, as reviewed in¹.

(B) Left: schematic of simultaneous multi-regional extracellular recordings from the IC, MGB and A1 of awake, head-fixed mice. One- or two-shanked 64-channel probe positioning relative to a schematic of the best frequency tonotopic gradients in each structure. Right: a single sweep of 192-channel multiunit activity across the IC, MGB, and A1 before and after presentation of a 20 ms white noise burst to the contralateral ear at 70 dB SPL.

(C) Left: spike rasters from representative single units in the IC, MGB, and A1. Gray shaded area denotes the timing of the 20 ms white noise burst. Right: autocorrelation function for each unit with exponential fit and computed decay constant.

(legend continued on next page)

quantified as the decay time constant (τ) of an exponential fit to the temporal spike count autocorrelation.^{6,7} Adapting this approach to our data, we confirmed a progression of neural timescales in the auditory system that approximately doubled from midbrain to cortex (Figure 1D). We reasoned that protracted spiking responses at higher hierarchical stages was likely associated with reduced temporal resolution from IC to MGB to A1. To demonstrate this, we presented pairs of noise bursts separated by a variable silent interval and used a spike timing-based template matching classifier to decode the inter-burst interval (Figures 1E and 1F). From midbrain to cortex, decoding accuracy was reduced overall and restricted to a limited range of longer intervals (Figure 1G), such that the median threshold of interval decoding rose from 1.4 ms for IC units, to 16 ms for MGB units, and 64 ms in A1 (Figure 1H), reflecting a progressive loss of temporal fidelity along the ascending auditory pathway that has been documented in many previous studies (for reviews see^{1,2,4}).

Cortical specialization for encoding slower, global rhythms

While slow neural timescales in the cortex may interfere with precise temporal coding, we reasoned that sluggish temporal processing could suggest a cortical advantage over subcortical stations for encoding slowly evolving features. To test this idea, we arranged the intervals separating consecutive noise bursts randomly or in a pattern (i.e., a rhythm) (Figure 2A, top). With this approach, the same set of four intervals comprised one cycle, but the intervals are arranged in a regular sequence (rhythmic context) or else occur in a random sequence. We created several libraries composed of different temporal interval arrangements, each set to a mean inter-burst interval of 200 ms (with no interval shorter than 100 ms) to avoid forward suppression, as supported by the analysis above (Figure 1C, Figures 1D–1H). Because the stimuli, the silent interval durations and the total cycle duration for random and rhythmic periods were matched, the sole feature distinguishing random and rhythmic contexts was regularity, which can only be encoded by integrating over at least one complete cycle (0.8 s).

When noise burst tokens switched between rhythmic and random sequences, we noted a clear change in the spiking patterns of A1 units that were not observed in simultaneously recorded MGB and IC ensembles (Figure 2A, bottom). To determine whether changes in ensemble spiking patterns could be used to decode whether individual cycles occurred in a rhythmic or random context, we first reduced the dimensionality of the

ensemble data matrix with principal components analysis. We then used a binary support vector machine (SVM) on the principal components projections to classify whether ensemble activity for each cycle came from the rhythmic or random stimulus context (Figure 2B). We observed that the changes in ensemble spiking were sufficient to classify the change from rhythmic to random sequences in A1, but not in MGB or IC (Figures 2C and 2D).

To determine the features of spike patterning that most clearly captured the difference between regular and random arrangements, we quantified the responses of single units across all 100 bursts (25 cycles) within each context. Intriguingly, it was the neural timescale itself that conveyed whether noise bursts were arranged randomly or in a rhythm; A1 spiking decayed more rapidly when the time intervals fell into a predictable rhythm than when they were arranged in a random sequence. To quantify neural timescale dynamics, we calculated the autocorrelation decay time constants for single units in random and rhythmic contexts using the same approach described above (Figures 3A and 3B). We found that neural timescales were unaffected by random or rhythmic sequences in the IC, were slightly and inconsistently dampened in the MGB, but were strongly and reliably dampened in A1 during rhythmic patterns (Figures 3C and 3D). Presenting the intervals in rhythmic patterns had no consistent effect on mean firing rate within the 100 ms response window (Figures 3E and 3F), but instead only on how the spikes were arranged within the 100 ms response window, both in terms of shorter decay constants and less variation in first spike latency (Figure S1). Neural timescales were equivalently compressed for regular spiking putative pyramidal neurons across A1 layers (Figure S2), yet no effect of context was observed in fast spiking putative parvalbumin-expressing interneurons (Figure S3). To control for the possibility that the ordinal arrangement of random following rhythmic could account for differences in spike patterning, we created a random-random stimulus but found that classification was near chance throughout and no systematic changes in neural timescales were observed (Figure S4).

When individual cycles were composed of four intervals, sensitivity to stimulus regularity required integration over at least 0.8 s, the duration of a single cycle. To determine if A1 neurons could integrate over even longer stimulus timescales, we tested the same units with random or rhythmic noise burst sequences composed of eight and twelve intervals per cycle (which repeated every 1.6 s and 2.4 s, respectively) (Figure 4a). Spiking

(D) Neural timescale measurements from the IC ($N/n = 5/48$, mice/single units), MGB (5/131), and A1 (4/140). Neural timescales significantly increased across the hierarchy (Kruskal-Wallis, 1.7×10^{-10} ; post hoc pairwise comparisons with Bonferroni correction for multiple comparisons, IC versus MGB, $1/0.05$; IC versus A1, $1 \times 10^{-6}/0.51$; MGB versus A1, $1 \times 10^{-8}/0.41$ for p value/Cliff's delta). Box-and-whisker plots show median values in solid gray line and 25th and 75th percentiles. Whiskers, range of non-outlier values. Circles, mean.

(E) Spike rasters for three representative single units for paired noise bursts. Alternating colors are presented for ease of visualization. Vertical gray lines = timing of 20 ms noise bursts.

(F) Mean confusion matrices for single-trial PSTH-based classification of inter-burst interval for the same IC, MGB, and A1 single units in (E).

(G) Mean \pm SEM probability of veridical interval classification (the upward diagonal from the confusion matrices in F) for all IC, MGB and A1 units. Interval classification accuracy is significantly reduced at successive stages of the IC-MGB-A1 hierarchy (mixed design ANOVA, main effect for structure, $F = 20.95$, $p = 6 \times 10^{-9}$).

(H) Decoder accuracy threshold in IC ($n = 5/48$, mice/single units), MGB (5/111), and A1 (4/109). Threshold increases across the hierarchy (Kruskal-Wallis, 8×10^{-17} ; post hoc pairwise comparisons with Bonferroni correction for multiple comparisons, IC versus MGB, $3.7 \times 10^{-6}/0.50$; IC versus A1, $6 \times 10^{-17}/0.83$; MGB versus A1, $8 \times 10^{-6}/0.37$ for p value/Cliff's delta). Box-and-whisker plots show median values in solid gray line and 25th and 75th percentiles. Whiskers, range of non-outlier values. Circles, mean.

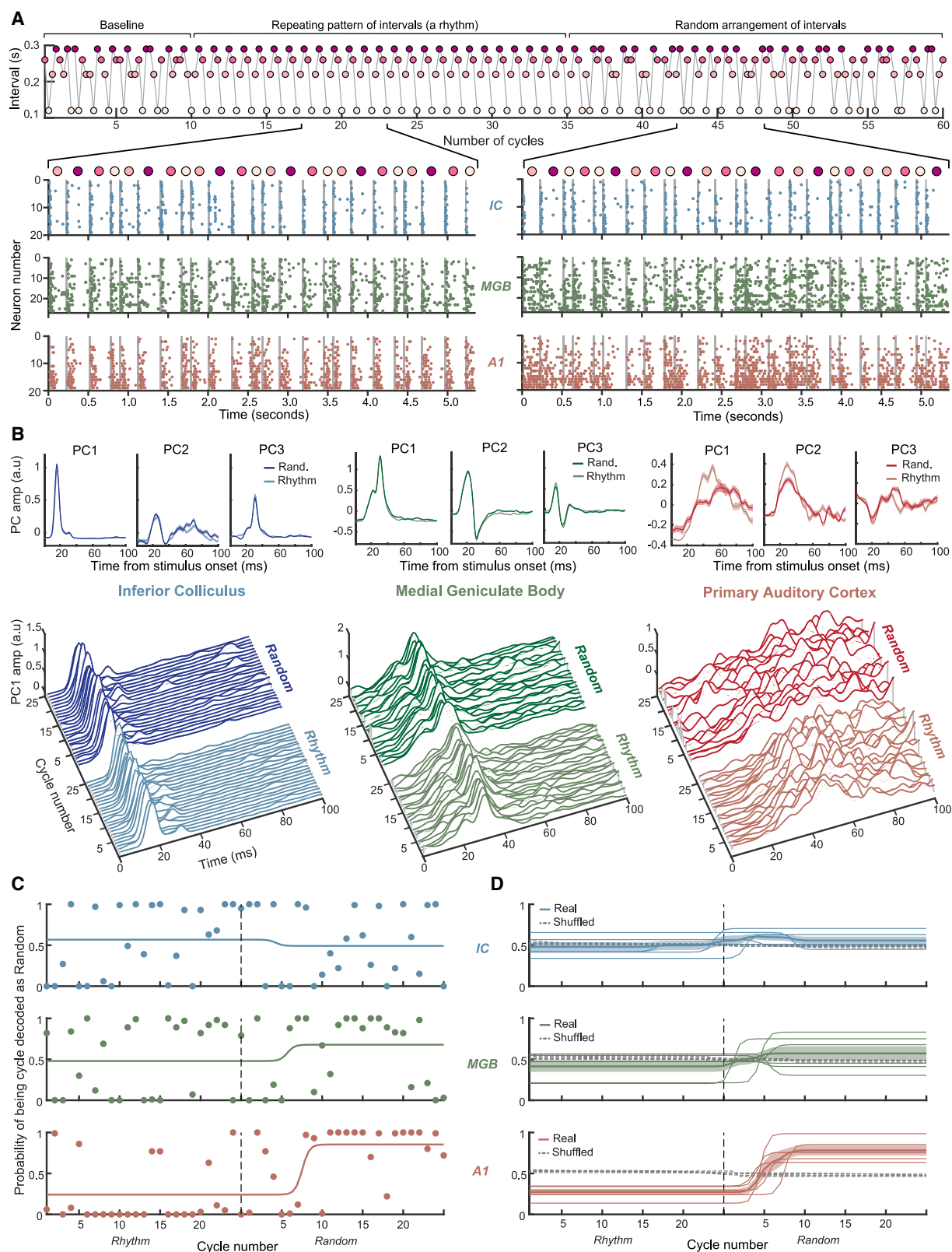


Figure 2. Accurate decoding of random or regular temporal interval arrangements observed in A1, not subcortically

(A) Top: schematic of rhythmic and random noise burst sequences. A single cycle is composed from four intervals which are joined by the gray line. The four intervals are presented in a stereotyped order to form a rhythm, or in a random order during the baseline and random periods. Bottom: six cycles from the rhythmic

(legend continued on next page)

decay constants in A1 were slightly reduced during rhythms composed of eight repeating intervals and were equivalent between random and rhythmic epochs with twelve intervals per cycle (Figures 4B–4E). Therefore, at least for A1 in passively listening mice, single unit sensitivity to regularity is limited to stimulus timescales on the order of 4–8 repeating elements or 0.8–1.6 s.

Spike rate adaptation versus spike timescale dynamics

We observed that the decay constants for sound-evoked spiking increased from the midbrain to cortex, and that the cortical timescale was itself shaped by slower changes in stimulus context, such that rhythmic, predictable sequences featured more orderly first spike latencies and more strongly dampened spiking decay. There is a long history of studying cortical adaptation to stimulus history in the auditory system, most notably with paradigms related to stimulus-specific adaptation (SSA), in which sound-evoked firing rates are modulated according to the probability of stimulus presentation.^{8–10} Like the protocol we used here, SSA is characterized by contrasting the neural response to an identical stimulus that differs only according to global context. While SSA has historically been studied with tones of varying frequency, recent work has also relied on silent gaps in broadband stimuli, demonstrating that it cannot be explained solely as adaptation in narrowly tuned frequency channels.¹¹ As the name implies, SSA is fundamentally a spike rate adaptation process dominated by synaptic fatigue for standard inputs, whereas synaptic inputs elicited by improbable deviant inputs remain unadapted, or even potentiated through as yet undefined recurrent or descending circuit mechanisms.^{12–14}

The protocol we introduce here was inspired by changes in cortical activity and perceptual awareness of complex repeating acoustic patterns in human subjects.^{15,16} While our approach is conceptually similar to prior work on SSA or contrast gain control in animal models, the methodological approach and neurophysiological results differ in several key respects.¹⁷ First, with the protocol described here, there are no low- or high-probability stimuli. All stimuli were identical frozen noise bursts separated by a single set of intervals used in both random and regular contexts. The only explicit feature that distinguished random from regular contexts was whether the sequence of inter-burst intervals repeated across cycles or were randomly arranged from one cycle to the next. Second, sensitivity to regularity in our paradigm required integration over periods lasting at least one cycle length (0.8 s), whereas high- and low-probability stimulus statistics that drive SSA or contrast gain control can generally be encoded on shorter timescales (but see^{9,18,19}). Third, while SSA,^{20–22} contrast gain

control,^{23–25} or other neural signatures of stimulus context changes²⁶ are robustly expressed in the A1, they are also clearly represented the IC and MGB. The context sensitivity we report here is completely absent in our sample of IC units and is weakly expressed in MGB neurons. Fourth, and perhaps most importantly, SSA, contrast gain control and other forms of context-dependent adaptation are based on changes in spike rate, whereas our findings demonstrate that random or rhythmic arrangements of temporal intervals can be decoded from A1 spiking (Figure 2D) without any change in mean spike rate (Figures 3E and 3F). Instead, stimulus context was represented solely by changes in how sound-evoked spikes were arranged in time.

Neural circuit mechanisms that may modulate spiking timescales, but not overall rate

When not presented in a regular pattern, sound-evoked A1 spiking persisted for approximately 20 ms (Figures 1D and 3C), with first spike latencies that vary on the order of 10 ms (Figure S1). It is interesting that cortical spiking timescales are sensitive to a stimulus feature that is approximately 40 × slower than the timescale itself (800 ms versus 20 ms), which suggests that A1 regular-spiking unit timescale dynamics are likely the product of other local or long-range circuits that respond more directly to the change in predictable stimulus timing.²⁷ Prior work has shown that the temporal decay of sensory-evoked spikes can be regulated by recurrent local circuits and inter-regional feedback loops.^{6,7,28–31} Cortical GABAergic neurons that express somatostatin or neuron-derived neurotrophic factor are clear candidates to dampen principal neuron spiking timing because they target the distal dendrites of A1 pyramidal neurons to regulate network excitability and recurrent excitation, but are also sensitive to slowly changing internal state variables.^{32–37} On the other hand, a subset of thalamic units also showed strong modulation of neural timescales and reduced firing rates during rhythmic contexts. These effects hint at an additional (or alternative) circuit mediated through corticothalamic feedback, which could dampen feedforward thalamocortical synaptic inputs via inhibition from the thalamic reticular nucleus.^{28,38}

Relevance to predictive sensory processing and perception

Whether studied in humans, birds, non-human primates or other species – including mice³⁹ – fundamental elements of vocal communication are captured in the time domain. In humans, primates and songbirds, emergent processing of slowly developing temporal patterns have been identified in the auditory cortex and other higher-order brain regions.^{16,40–44} Within the auditory

and random epochs are presented with the single unit spike rasters from 68 simultaneously recorded units in the IC, MGB, and A1. Vertical gray bars denote the timing of individual noise bursts. See also supplemental audio files presenting examples of regular and rhythmic stimulus sequences. Refer to Audio S1 and S2 for audio excerpts of random and rhythmic interval sequences.

(B) Top: evoked response over a 100 ms period beginning at stimulus onset for the top three principal components in each region averaged across all 100 noise bursts in each stimulus context (4 bursts per cycle, 25 cycles per epoch). Bottom: amplitude of the first principal component's response for each of 25 cycles within the rhythmic and random contexts.

(C) An SVM was used to classify the ensemble spiking for each cycle to random or rhythmic contexts, as shown from an example mouse with simultaneous IC, MGB, and A1 recordings. Solid line represents sigmoidal fit of the cycle-by-cycle decoding.

(D) Decoder output for ensembles in IC (N = 5, mice), MGB (N = 6), and A1 (N = 5). Gray lines indicate the mean random cycle shuffled control from each recording. Thick line with shading, mean ± SEM. Classification accuracy increases along the IC-MGB-A1 hierarchy (two-way repeated-measures ANOVAs, main effect for classification over cycles: IC [F = 0.5, p = 0.99], MGB [F = 0.69, p = 0.93], A1 [F = 18.81, p = 0]; interaction term for classification × condition, IC [F = 1.23, p = 0.15], MGB [F = 1.55, p = 0.01], A1, [F = 21.28, p = 0]).

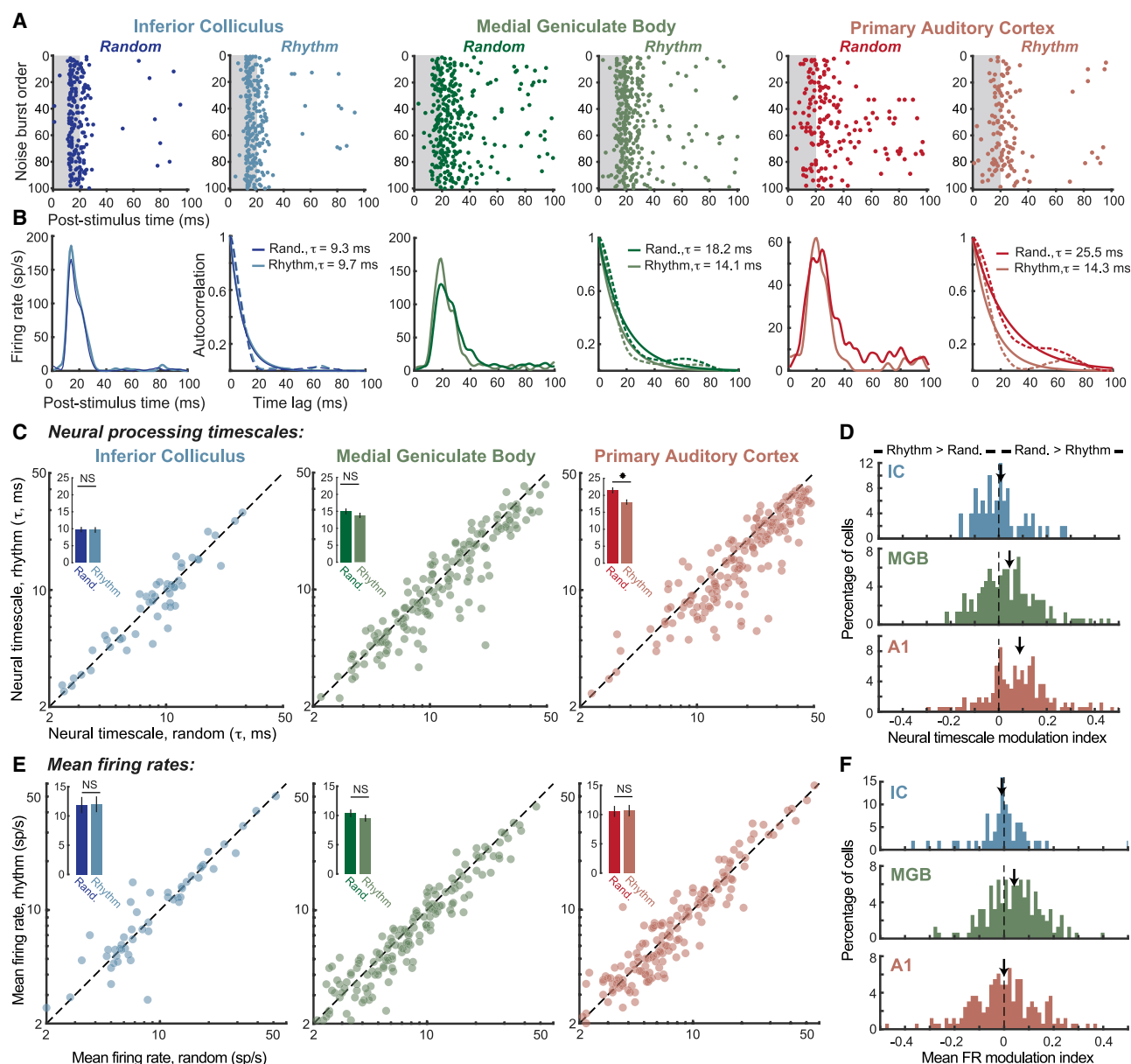


Figure 3. A top-down representation of global temporal patterns via spike timing dynamics, not spike rate

(A) Representative spike rasters from IC, MGB, and A1 units for 100 noise bursts presented in random and rhythmic sequences. Gray shading = 20 ms noise burst. (B) Post-stimulus firing rate histograms and autocorrelation functions for the example units in (A). See also Figure S1 for quantification of first spike latency variability during random and rhythmic epochs.

(C) Neural timescale measurements from sound-evoked spiking during random and rhythmic epochs in IC ($N/n = 5/50$, mice/single units), MGB (6/154), and A1 (5/164). Inset: mean \pm SEM neural timescale from each region. Asterisk denotes statistical significance with a paired t test ($p < 0.05$ and Cohen's $d > 0.5$); (IC, 0.7/0.05; MGB, 0.00009/0.33; A1, 3×10^{-14} /0.65 for p value/Cohen's d). See also Figure S2 for spike timescale changes across layers and Figure S3 for spike timescale changes between regular- and fast-spiking units.

(D) Histogram of neural timescale asymmetry index ($[(\text{random} - \text{rhythm}) / (\text{random} + \text{rhythm})]$), where values < 0 indicate more dampened responses during random intervals and values > 0 indicate more dampened responses during rhythmic intervals. Arrows denote sample means. Neural timescales are significantly reduced during rhythmic epochs in A1, slightly reduced in the MGB, but not affected in the IC (one-sample t test against a population mean of 0; IC, 0.27/0.08; MGB, 0.00002/0.33; A1, 8×10^{-13} /0.59 for p value/Cohen's d). See also Figure S4 for a characterization of a control experiments in which the stimulus remained random for all 50 post-baseline cycles.

(E) Sound-evoked firing measured from random and rhythmic epochs. Inset: mean \pm SEM spike rate from each region. NS, not significant. Asterisk denotes statistical significance with a paired t test ($p < 0.05$ and Cohen's $d > 0.5$); (IC, 0.68/−0.06; MGB, 1×10^{-5} /0.36; A1, 0.47/−0.06 for p value/Cohen's d).

(F) Firing rate asymmetry during random and rhythmic stimulus epochs. Plotting conventions match D. Mean firing rates are only weakly modulated by stimulus context in all brain areas (one-sample t test against a population mean of 0; IC, 0.6/0.07; MGB, 1×10^{-5} /0.36; A1, 0.97/0.00, for p value/Cohen's d).

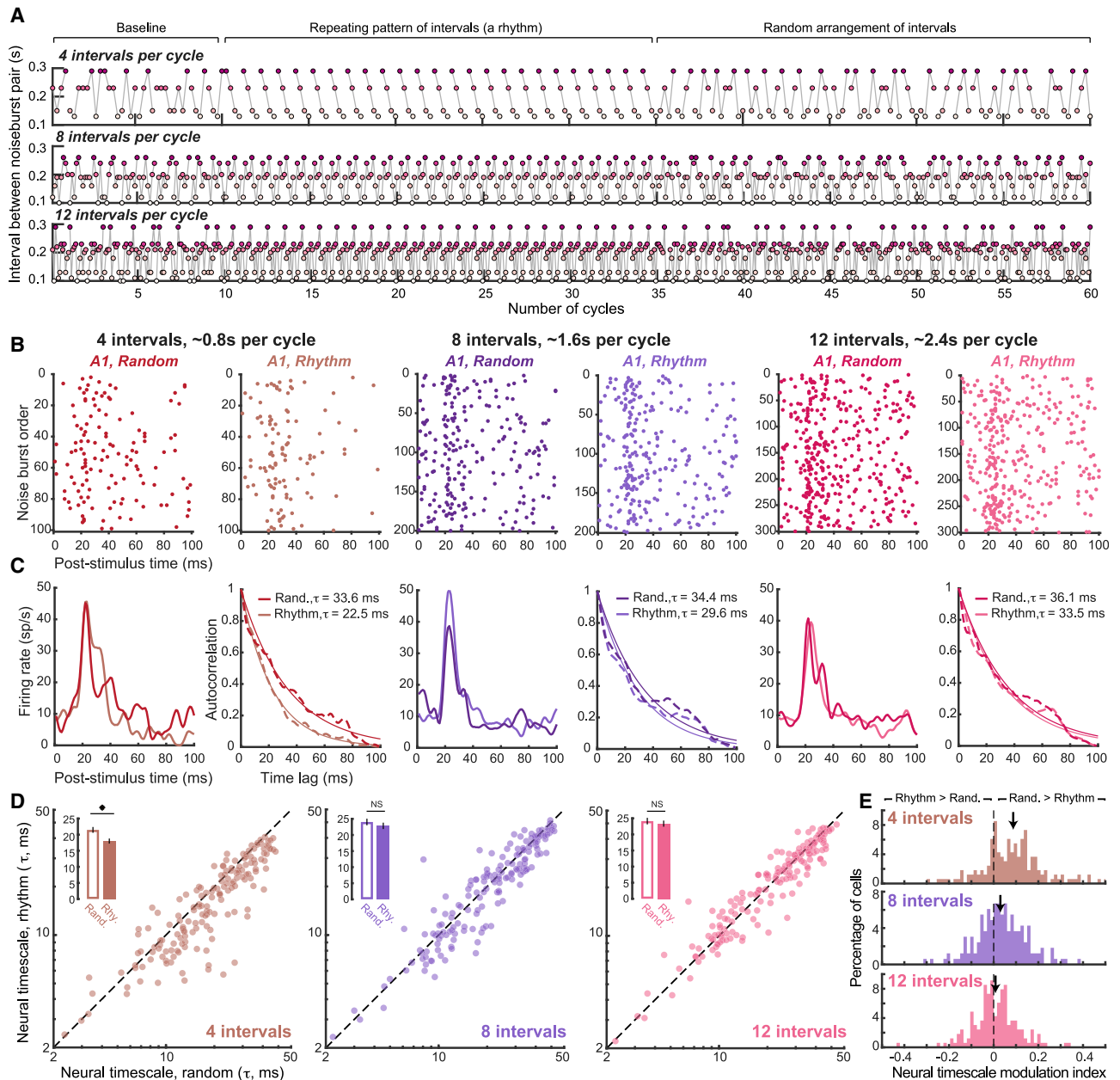


Figure 4. Dampened A1 neural timescales occur with short repeating rhythms, not with longer rhythms

(A) Sequences of baseline-rhythm-random interval arrangements where individual cycles are composed of 4, 8, or 12 noise burst intervals. Refer to [Audio S3](#) and [S4](#) for audio excerpts of rhythmic interval sequences comprised of eight and twelve intervals per cycle.

(B) Spike rasters from representative A1 single units recorded with sequences of varying cycle size. Random and rhythmic epochs present 100, 200, or 300 epochs (25 cycles per epoch \times 4/8/12 noise burst intervals per cycle).

(C) Post-stimulus firing rate histograms and autocorrelation functions for the example units in (B).

(D) Neural timescale measurements from the random and rhythmic epochs are presented for 164 single units, each recorded with a cycle size of 4, 8, and 12 noise burst intervals. Inset: mean \pm SEM neural timescale for each cycle length. NS, not significant. Asterisk denotes statistical significance with a paired t test ($p < 0.05$ and Cohen's $d > 0.5$); (4 intervals, $3 \times 10^{-14}/0.65$; 8 intervals, $0.003/0.23$; 12 intervals, $0.03/0.17$ for p value/Cohen's d).

(E) Histogram of neural timescale asymmetry index ($[(\text{random} - \text{rhythm}) / (\text{random} + \text{rhythm})]$), where values < 0 indicate more dampened responses during random intervals and values > 0 indicate more dampened responses during rhythmic intervals. Arrows denote sample means. Neural timescales are significantly reduced when noise bursts form rhythms than during random arrangements with cycle sizes of 4, but not 8 or 12 (one-sample t test against a population mean of 0, 4 intervals, $2 \times 10^{-12}/0.59$; 8 intervals, $0.002/0.24$; 12 intervals, $0.31/0.08$, for p value/Cohen's d).

cortex, the time constants that regulate temporal integration of speech and auditory regularity increases across core to higher-order fields.^{45–47} In our study, neural timescales in A1 did not integrate over longer, more complex rhythmic sequences that repeated less frequently (Figure 4). In future work, it would be important to test whether even more protracted spiking timescales and more integrative processing of slower developing regularities are encoded in candidate higher-order areas of the mouse auditory cortex, such as the dorsal-posterior and ventral-posterior fields, particularly when mice are behaviorally engaged in active listening tasks.⁴⁸ Our findings show that – at least from the perspective of first spike latency jitter and spike decay time constants – the cortex becomes more like the thalamus when sound features are regular and predictable. This may reflect a physiological transition in the cortex from a “deliberative” mode when sound features are unpredictable, typified by strong inter- and intra-columnar recurrent activity, to a “pass through” mode when environmental stimuli match internal predictive models.^{15,49,50}

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
 - Lead Contact
 - Materials Availability
 - Data and Code Availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
 - Stimulus design
 - Surgical preparation
 - Simultaneous multi-regional recordings
- QUANTIFICATION AND STATISTICAL ANALYSIS
 - Data acquisition and online analysis
 - Single units sorting
 - Spike timing-based template matching classifier
 - Ensemble decoding analysis using dimensionality reduction and a support vector machine classifier
 - Neural timescales
 - Modulation indices
 - Statistical Analysis

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at <https://doi.org/10.1016/j.cub.2021.01.076>.

ACKNOWLEDGMENTS

Financial support for this work was provided by NIH grant DC017178 (D.B.P.), NIH fellowship DC015376 (R.S.W.), and a Herchel Smith Harvard Scholarship (M.M.A.). We thank Aravind Parthasarathy for helpful discussions about stimulus design.

AUTHOR CONTRIBUTIONS

M.M.A., K.E.H., and D.B.P. designed the experiments. M.M.A. performed all experiments and data analysis, with supervisory input from D.B.P. R.S.W.

trained M.M.A. to perform single unit recordings and wrote code for single unit data pre-processing. K.E.H. developed software for stimulus control and neural data acquisition. D.B.P. wrote the manuscript with input from all the authors.

DECLARATION OF INTERESTS

The authors have no competing interests to declare.

Received: August 20, 2020

Revised: December 1, 2020

Accepted: January 21, 2021

Published: February 19, 2021

REFERENCES

1. Joris, P.X., Schreiner, C.E., and Rees, A. (2004). Neural processing of amplitude-modulated sounds. *Physiol. Rev.* 84, 541–577.
2. Wang, X., Lu, T., Bendor, D., and Bartlett, E. (2008). Neural coding of temporal information in auditory thalamus and cortex. *Neuroscience* 157, 484–494.
3. King, A.J., and Nelken, I. (2009). Unraveling the principles of auditory cortical processing: can we learn from the visual system? *Nat. Neurosci.* 12, 698–701.
4. Cassey, J.H., and Covey, E. (1996). A neuroethological theory of the operation of the inferior colliculus. *Brain Behav. Evol.* 47, 311–336.
5. Yin, T.C.T., Smith, P.H., and Joris, P.X. (2019). Neural mechanisms of binaural processing in the auditory brainstem. *Compr. Physiol.* 9, 1503–1575.
6. Chaudhuri, R., Knoblauch, K., Gariel, M.A., Kennedy, H., and Wang, X.J. (2015). A large-scale circuit mechanism for hierarchical dynamical processing in the primate cortex. *Neuron* 88, 419–431.
7. Murray, J.D., Bernacchia, A., Freedman, D.J., Romo, R., Wallis, J.D., Cai, X., Padoa-Schioppa, C., Pasternak, T., Seo, H., Lee, D., and Wang, X.J. (2014). A hierarchy of intrinsic timescales across primate cortex. *Nat. Neurosci.* 17, 1661–1663.
8. Latimer, K.W., Barbera, D., Sokoletsky, M., Awwad, B., Katz, Y., Nelken, I., Lampl, I., Fairhall, A.L., and Priebe, N.J. (2019). Multiple timescales account for adaptive responses across sensory cortices. *J. Neurosci.* 39, 10019–10033.
9. Ulanovsky, N., Las, L., Farkas, D., and Nelken, I. (2004). Multiple time scales of adaptation in auditory cortex neurons. *J. Neurosci.* 24, 10440–10453.
10. Pérez-González, D., and Malmierca, M.S. (2014). Adaptation in the auditory system: an overview. *Front. Integr. Neurosci.* 8.
11. Awwad, B., Jankowski, M.M., and Nelken, I. (2020). Synaptic recruitment enhances gap termination responses in auditory cortex. *Cereb. Cortex* 30, 4465–4480.
12. Solomon, S.G., and Kohn, A. (2014). Moving sensory adaptation beyond suppressive effects in single neurons. *Curr. Biol.* 24, R1012–R1022.
13. Whitmire, C.J., and Stanley, G.B. (2016). Rapid sensory adaptation redux: a circuit perspective. *Neuron* 92, 298–315.
14. Polterovich, A., Jankowski, M.M., and Nelken, I. (2018). Deviance sensitivity in the auditory cortex of freely moving rats. *PLoS one* 13, e0197678.
15. Heilbron, M., and Chait, M. (2018). Great expectations: is there evidence for predictive coding in auditory cortex? *Neuroscience* 389, 54–73.
16. Barascud, N., Pearce, M.T., Griffiths, T.D., Friston, K.J., and Chait, M. (2016). Brain responses in humans reveal ideal observer-like sensitivity to complex acoustic patterns. *Proc. Natl. Acad. Sci. USA* 113, E616–E625.
17. Robinson, B.L., and McAlpine, D. (2009). Gain control mechanisms in the auditory pathway. *Curr. Opin. Neurobiol.* 19, 402–407.
18. Yaron, A., Hershenhoren, I., and Nelken, I. (2012). Sensitivity to complex statistical regularities in rat auditory cortex. *Neuron* 76, 603–615.

19. Rubin, J., Ulanovsky, N., Nelken, I., and Tishby, N. (2016). The representation of prediction error in auditory cortex. *PLoS Comput. Biol.* **12**, e1005058.
20. Parras, G.G., Nieto-Diego, J., Carbajal, G.V., Valdés-Baizabal, C., Escera, C., and Malmierca, M.S. (2017). Neurons along the auditory pathway exhibit a hierarchical organization of prediction error. *Nat. Commun.* **8**, 2148.
21. Ayala, Y.A., and Malmierca, M.S. (2013). Stimulus-specific adaptation and deviance detection in the inferior colliculus. *Front. Neural Circuits* **6**, 89.
22. Anderson, L.A., Christianson, G.B., and Linden, J.F. (2009). Stimulus-specific adaptation occurs in the auditory thalamus. *J. Neurosci.* **29**, 7359–7363.
23. Dean, I., Harper, N.S., and McAlpine, D. (2005). Neural population coding of sound level adapts to stimulus statistics. *Nat. Neurosci.* **8**, 1684–1689.
24. Lohse, M., Bajo, V.M., King, A.J., and Willmore, B.D.B. (2020). Neural circuits underlying auditory contrast gain control and their perceptual implications. *Nat. Commun.* **11**, 324.
25. Kvale, M.N., and Schreiner, C.E. (2004). Short-term adaptation of auditory receptive fields to dynamic stimuli. *J. Neurophysiol.* **91**, 604–612.
26. Cai, R., Richardson, B.D., and Caspary, D.M. (2016). Responses to predictable versus random temporally complex stimuli from single units in auditory thalamus: impact of aging and anesthesia. *J. Neurosci.* **36**, 10696–10706.
27. Cannon, J.J., and Patel, A.D. (2021). How beat perception co-opts motor neurophysiology. *Trends Cogn. Sci.* **25**, 137–150.
28. Reinhold, K., Lien, A.D., and Scanziani, M. (2015). Distinct recurrent versus afferent dynamics in cortical visual processing. *Nat. Neurosci.* **18**, 1789–1797.
29. Runyan, C.A., Piasini, E., Panzeri, S., and Harvey, C.D. (2017). Distinct timescales of population coding across cortex. *Nature* **548**, 92–96.
30. Hart, E., and Huk, A.C. (2020). Recurrent circuit dynamics underlie persistent activity in the macaque frontoparietal network. *eLife* **9**, 1–22.
31. Gao, R., van den Brink, R., Pfeffer, T., and Voytek, B. (2020). Neuronal timescales are functionally dynamic and shaped by cortical microarchitecture. *Elife* **9**, p.e61277.
32. Kato, H.K., Gillet, S.N., and Isaacson, J.S. (2015). Flexible sensory representations in auditory cortex driven by behavioral relevance. *Neuron* **88**, 1027–1039.
33. Abs, E., Poorthuis, R.B., Apelblat, D., Muhammad, K., Pardi, M.B., Enke, L., Kushinsky, D., Pu, D.L., Eizinger, M.F., Conzelmann, K.K., et al. (2018). Learning-related plasticity in dendrite-targeting layer 1 interneurons. *Neuron* **100**, 684–699.e6.
34. Kato, H.K., Asinof, S.K., and Isaacson, J.S. (2017). Network-level control of frequency tuning in auditory cortex. *Neuron* **95**, 412–423.e4.
35. Park, Y., and Geffen, M.N. (2020). A circuit model of auditory cortex. *PLoS Comput. Biol.* **16**, e1008016.
36. Natan, R.G., Briguglio, J.J., Mwilambwe-Tshilobo, L., Jones, S.I., Aizenberg, M., Goldberg, E.M., and Geffen, M.N. (2015). Complementary control of sensory adaptation by two types of cortical interneurons. *eLife* **4**, 4.
37. Fan, L.Z., Kheifets, S., Böhm, U.L., Wu, H., Piatkevich, K.D., Xie, M.E., Parot, V., Ha, Y., Evans, K.E., Boyden, E.S., et al. (2020). All-optical electrophysiology reveals the role of lateral inhibition in sensory processing in cortical layer 1. *Cell* **180**, 521–535.e18.
38. Guo, W., Clause, A.R., Barth-Maron, A., and Polley, D.B. (2017). A cortico-thalamic circuit for dynamic switching between feature detection and discrimination. *Neuron* **95**, 180–194.e5.
39. Castellucci, G.A., Calbick, D., and McCormick, D. (2018). The temporal organization of mouse ultrasonic vocalizations. *PLoS ONE* **13**, e0199929.
40. Lim, Y., Lagoy, R., Shinn-Cunningham, B.G., and Gardner, T.J. (2016). Transformation of temporal sequences in the zebra finch auditory system. *eLife* **5**.
41. Cazala, A., Giret, N., Edeline, J.M., and Del Negro, C. (2019). Neuronal encoding in a high-level auditory area: from sequential order of elements to grammatical structure. *J. Neurosci.* **39**, 6150–6161.
42. Barczak, A., O'Connell, M.N., McGinnis, T., Ross, D., Mowery, T., Falchier, A., and Lakatos, P. (2018). Top-down, contextual entrainment of neuronal oscillations in the auditory thalamocortical circuit. *Proc. Natl. Acad. Sci. USA* **115**, E7605–E7614.
43. Herrmann, B., and Johnsrude, I.S. (2018). Neural signatures of the processing of temporal patterns in sound. *J. Neurosci.* **38**, 5466–5477.
44. Bekinschtein, T.A., Dehaene, S., Rohaut, B., Tadel, F., Cohen, L., and Naccache, L. (2009). Neural signature of the conscious processing of auditory regularities. *Proc. Natl. Acad. Sci. USA* **106**, 1672–1677.
45. Overath, T., McDermott, J.H., Zarate, J.M., and Poeppel, D. (2015). The cortical analysis of speech-specific temporal structure revealed by responses to sound quilts. *Nat. Neurosci.* **18**, 903–911.
46. Lerner, Y., Honey, C.J., Silbert, L.J., and Hasson, U. (2011). Topographic mapping of a hierarchy of temporal receptive windows using a narrated story. *J. Neurosci.* **31**, 2906–2915.
47. Keceli, S., Okamoto, H., and Kakigi, R. (2015). Hierarchical neural encoding of temporal regularity in the human auditory cortex. *Brain Topogr.* **28**, 459–470.
48. Romero, S., Hight, A.E., Clayton, K.K., Resnik, J., Williamson, R.S., Hancock, K.E., and Polley, D.B. (2020). Cellular and widefield imaging of sound frequency organization in primary and higher order fields of the mouse auditory cortex. *Cereb. Cortex* **30**, 1603–1622.
49. Keller, G.B., and Mrsic-Flogel, T.D. (2018). Predictive processing: a canonical cortical computation. *Neuron* **100**, 424–435.
50. Nobre, A.C., and van Ede, F. (2018). Anticipated moments: temporal structure in attention. *Nat. Rev. Neurosci.* **19**, 34–48.
51. Pachitariu, M., Steinmetz, N., Kadir, S., Carandini, M., and Kennedy, D.H. (2016). Kilosort: realtime spike-sorting for extracellular electrophysiology with hundreds of channels. *bioRxiv*, 061481.
52. Foffani, G., and Moxon, K.A. (2004). PSTH-based classification of sensory stimuli using ensembles of single neurons. *J. Neurosci. Methods* **135**, 107–120.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited Data		
Sorted spikes	Current study	Mendeley Data: https://doi.org/10.17632/v7tpm6gsvp.1
Experimental Models: Organisms/Strains		
Mouse; C57BL/6J	The Jackson Laboratory	Cat# JAX: 000664; RRID: IMSR_JAX:000664
Software and Algorithms		
LabVIEW	National Instruments	https://www.ni.com/en-us/shop/labview.html
Kilosort	GitHub	https://github.com/MouseLand/Kilosort
MATLAB	MathWorks	https://www.mathworks.com/products/matlab.html
Synapse	Tucker-Davis Technologies	http://www.tdt.com/component/synapse-software/
Other		
PXI	National Instruments	PXI-4461
Free-field speaker	Parts Express	275-010
Silicon recording probes	Cambridge NeuroTech	H2, H3
Neurodigitizer and preamplifier	Tucker-Davis Technologies	PZ5
Data processor and real-time controller	Tucker-Davis Technologies	RZ2
Data streamer	Tucker-Davis Technologies	RS4
ZIF-Clip headstage components	Tucker-Davis Technologies	ZC64, Z-ROD64, ZCA-NN64

RESOURCE AVAILABILITY

Lead Contact

Further information and requests should be directed to and will be fulfilled by the lead contact, Daniel Polley (daniel_polley@meei.harvard.edu).

Materials Availability

This study did not generate new reagents.

Data and Code Availability

Original data have been deposited to Mendeley Data: <https://doi.org/10.17632/v7tpm6gsvp.1>.

The analysis codes used in the study are available on GitHub: https://github.com/Meenakshi-Asokan/Asokan_et_al_2021_CurrentBiology.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

We used adult male and female C57BL6 mice aged 6-8 weeks at the time of recording. All procedures were approved by the Massachusetts Eye and Ear Infirmary Animal Care and Use Committee and followed the guidelines established by the National Institute of Health for the care and use of laboratory animals.

METHOD DETAILS

Stimulus design

Auditory stimuli were frozen broadband noise bursts (20 ms duration, 5 ms raised cosine onset/offset ramps, 70 dB SPL, 1-64 kHz bandwidth). Stimuli were generated with a 24-bit digital-to-analog converter (National Instruments model PXI-4461) and presented

via a free-field speaker (Parts Express 275-010) placed approximately 10 cm from the left (contralateral) ear canal. Free-field stimuli were calibrated using a wide-band free-field microphone (PCB Electronics, 378C01).

Noise bursts were presented in pairs (1-512 ms silent inter burst intervals, incremented in octave steps, 25 repetitions each, 1 s inter-trial interval) or in longer sequences with durations set to 48, 96 or 144 s for 4, 8 and 12 intervals per cycle, respectively. All sequences used the same range of inter-pulse intervals (100 – 300 ms in 20 ms increments). We then pseudo randomly selected a set of intervals for a cycle to have a mean value of 200 ms. For the random sequence, we concatenated 25 cycles, where the same set of intervals were randomly ordered for each cycle. The rhythmic pattern was generated similarly, except that the order of intervals selected for the first cycle was repeated for the next 24 cycles. A complete sequence began with a baseline period consisting of at least 10 cycles of random interval arrangements followed by 25 cycles of the repeating rhythm, which was followed by 25 cycles of randomly ordered intervals. We generated multiple sequences of the baseline-rhythmic-random arrangement for each cycle duration, each based on a different pseudorandom selection of intervals. We also tested a control stimulus in which random temporal interval arrangement was used throughout (Figure S4). To reduce the influence of memory or habituation, we only used the first sequence per mouse for each of the three cycle durations. This paradigm ensures that the explicit auditory inputs (i.e., the noise burst itself and the particular set of silent intervals separating the noise bursts) are matched throughout the sequence, and that repetition is the sole feature distinguishing the rhythmic epoch from the baseline and random epochs. Refer to [Audio S1](#), [Audio S2](#), [Audio S3](#), and [Audio S4](#) for excerpts of random and rhythmic interval arrangements of varying cycle lengths.

Surgical preparation

We first anesthetized the mice with isoflurane in oxygen (5% for induction, 1.5% for maintenance), and used a homeothermic blanket system (FHC) to maintain their body temperature at 36.5°C. After administering Lidocaine subcutaneously to numb the surgical area, we removed the scalp and the periosteum overlying the dorsal surface of the skull. To create a better adhesive surface, we prepared the skull with etchant (C&B metabond) and 70% ethanol before affixing a custom titanium head plate (iMaterialise) using dental cement (C&B metabond). We also created a small burr hole over the left occipital cortex to implant a ground wire (AgCl). We administered Buprenex (0.05 mg/kg) and meloxicam (0.1 mg/kg) at the conclusion of the surgery. Mice were individually housed following surgery and were given at least 48 h to acclimate to the head plate before the recording.

On the day of the recording, we again anesthetized the mice with isoflurane in oxygen and made three very small (~1 × 1-1.5 mm) craniotomies over the right hemisphere using a scalpel. The IC craniotomy was made just lateral to the midline and just caudal to the lambdoid suture, taking the utmost care to avoid the sinuses running underneath. The MGB craniotomy was made 1.5 mm rostral to the lambdoid suture and 2-3 mm lateral to midline. To expose A1, we first retracted the temporalis muscle and then made a craniotomy centered on the temporal suture line approximately 1.5 mm rostral to the lambdoid suture. Finally, we built a well around each craniotomy with UV-cured cement (Flow-It ALC Flowable Composite) and filled them with lubricating ointment (Paralube Vet Ointment), to keep the brain moist over the entire recording session. As the mice recover from anesthesia, we head-fixed them by attaching their head plate to a rigid clamp while their body rested on a cradle inside a dimly lit single-wall sound attenuating recording chamber (Acoustic Systems). We waited for at least 30 min before starting the recording.

Simultaneous multi-regional recordings

We used three linear 64-channel silicon probes (Cambridge Neurotech), including a single shank (H3 probe, 20 mm spacing) electrode for recording from IC, and a dual shank (H2 probe, 25 mm spacing between contacts within a shank, and 200 mm spacing between shanks) electrode for recording from MGB and A1, and inserted each using a micromanipulator (Narishige) and a hydraulic microdrive (FHC). The IC recording probe was inserted slightly off the dorsal to ventral axis to capture the tonotopic gradient of the central nucleus. We inserted the MGB electrode perpendicular to the brain surface with the tip of the probe resting 3.5 - 4 mm below the surface and the two shanks oriented medio-laterally. We started at the lateral edge of the craniotomy and then retracted and reinserted the probe at progressively more medial positions until we had broadband noise-evoked spiking activity in most channels of both shanks, and in so doing, we were assured of recording from the lateral bank of the MGB, which contains the ventral subdivision and, depending on the caudal-rostral coordinates, might also contain recording sites in the dorsal subdivision and supra-geniculate nucleus. Finally, we inserted the A1 probe perpendicular to the brain surface until the tip of the probe was approximately 1.2 mm below the brain surface, resting just above the white matter, with the two shanks oriented rostro-caudally. After all three probes were inserted, we let the brain settle for at least 15 min before the recordings began.

The laminar position of the A1 recording probes was estimated online by measuring the current source density (CSD) evoked by single broadband noise bursts (50 ms duration, 4 ms onset/offset cosine ramps, 1 s inter-stimulus interval). We also measured frequency response areas using pure tone stimuli (50 ms duration, 4 ms onset/offset cosine ramps, 0.5 s inter-stimulus interval, 4 - 45 kHz with 0.1 octave steps, 0 – 70 dB SPL with 5 dB steps, 2 repetitions of each stimulus, pseudo-randomized). We ensured recordings were made from A1 based on the best frequency gradient along the rostro-caudal axis and the tonotopic reversal toward the rostral edge of the craniotomy.

QUANTIFICATION AND STATISTICAL ANALYSIS

Data acquisition and online analysis

Raw neural signals were digitized at 32-bit, 24.4 kHz and stored in binary format (PZ5 Neurodigitizer, RZ2 BioAmp Processor, RS4 Data Streamer; Tucker-Davis Technologies). To eliminate artifacts, electrical signals were notch filtered at 60 Hz, the common mode signal (channel-averaged trace) was subtracted from the raw signals from all channels, independently for each probe. Signals were then band-pass filtered (300–3000 Hz, second-order Butterworth filters), from which multiunit activity was extracted as negative deflections in the electrical trace with an amplitude exceeding 4 standard deviations of the baseline hash. To extract local field potentials, signals were down-sampled to 1000 Hz and spatially smoothed along the channels with a triangle filter (5-point Hanning window). The CSD was calculated as the second spatial derivative of the local field potential signal. The noise-evoked columnar CSD patterns were used to determine the laminar location of A1 recording sites: A brief current sink first occurs approximately 10 ms after noise onset, which was used to determine the lower border of L4, and other layers were defined relative to the location of L4 (Figure S3) as described previously.³⁸

Single units sorting

We used Kilosort2⁵¹ to sort spikes into single unit clusters. We concatenated all data files from a given recording session so that the same unit could be tracked over the full course of the experiment (~90 min). We ensured our units were well isolated clusters and had refractory period violations for less than 5% of the total spikes (inter-spike interval < 1.5 ms for IC and < 2 ms for MGB and A1). Once isolated, cortical single units were classified as putative excitatory regular spiking versus putative inhibitory fast spiking and assigned to a layer based on the CSD analysis. For the regular versus fast spike waveform classification, the trough to peak interval of the mean spike waveform formed a bimodal distribution, which was used to subdivide neurons with intervals exceeding 0.6 ms as regular spiking, while neurons with intervals shorter than 0.6 ms were fast spiking (Figure S4A). To be included for analysis, the peak firing rate in the post-stimulus response period (0 – 50 ms) to isolated noise bursts had to be³ 1.5 standard deviations above the spontaneous firing rate.

Spike timing-based template matching classifier

To classify the interval separating paired pulses, we used a decoder that works by creating a set of templates based on trial-averaged neural responses to stimuli from all trials except one, and then assigning the raw spiking response in the remaining trial to the stimulus with the ‘closest’ template in the Euclidean distance sense. Hence, it’s a template matching, minimum Euclidean distance classifier that is based on first-order statistics.⁵² A matrix with $R \times S$ rows and B columns was constructed, where R is the number of stimulus repeats ($R = 25$), S is the number of stimuli ($S = 20$), B is the number of bins that contain spikes ($B = 1000$; 1 ms bins were used). Let v_{ij} represent the spike count in i^{th} row and j^{th} column of the matrix, where i goes from 1 to SR and j goes from 1 to B . Templates for each stimulus were defined as $t^s = [t_1^s, \dots, t_B^s]$, where the j^{th} element is calculated as $t_j^s = \frac{1}{R} \sum_{i \in S} v_{ij}$. For each trial $v_i = [v_{i,1}, \dots, v_{i,B}]$, the Euclidean

distance between that trial and each stimulus template t^s was defined as $d_i^s = \sqrt{\sum_{j=1}^B (v_{ij} - t_j^s)^2}$. Using these distances, the spike train

was classified as being generated by the stimulus class represented by the closest template, resulting in an outcome vector of $c_i = [c_1, \dots, c_{SR}]$ where $c_i = \text{argmin}(d_i^s)$, and each element c_i indicated to which stimulus the i^{th} trial is assigned. Decoder threshold was calculated as the shortest interval after which classifier accuracy is always above chance.

Ensemble decoding analysis using dimensionality reduction and a support vector machine classifier

We decoded whether individual cycles occurred in rhythmic or random epochs based on ensemble spike trains. From the simultaneously recorded ensemble of neurons within each region, we reduced dimensionality using principal component analysis. We considered a 100 ms window following each noise burst with 1 ms bin size, averaged the neural response across every 4 noise bursts within each cycle, and then smoothed using a 5 ms Gaussian filter. The smoothing is to be able to capture the shared variance across neurons while discarding the stochastic fluctuations particular to individual neurons. We then fed the responses from the top principal components that capture 75% the variance in the data to a SVM with a linear kernel. We trained and cross-validated the classifier using 10-fold cross-validation which was iterated 100 times with a different random split for each iteration. The decoder output is further parameterized by fitting a sigmoidal function to be able to infer the decoder accuracy within each context and the time it takes for the change in context to be decoded. As a control, we performed the same analysis after shuffling the ordering of cycles and we repeat this process for each recording in all three brain regions. The SVM training and cross validation was performed in MATLAB using functions ‘cvpartition’, ‘fitcsvm’ and ‘predict’.

Neural timescales

We defined the neural timescale (τ) for each unit as the decay time constant of an exponential fit to the temporal autocorrelation function. We first compute the post-stimulus time histogram (PSTH) as the trial averaged neural response to the frozen noise burst in a 100 ms window. We then compute the autocorrelation of the PSTH for various time lags, and fit a decaying exponential to this autocorrelation function as $f(t) = ae^{-bt}$, where we fixed $a = 1$ since the autocorrelation at time lag 0 is always 1. The decay constant

was extracted from the fit as $\tau = \frac{1}{b}$, the time it takes for the autocorrelation exponential fit to decrease by a factor of e. For units where the single-exponential fit explains less than 75% of the total variance in the data, we use a double-exponential fit $f(t) = ae^{-bt} + ce^{-dt}$ and estimate τ as the weighted sum of the two time constants, $\tau = \frac{a(\frac{1}{b}) + c(\frac{1}{d})}{a + c}$ consistent with prior work.⁶ All fits were verified by eye.

Modulation indices

Modulation indices are computed as $\frac{p_{\text{random}} - p_{\text{rhythm}}}{p_{\text{random}} + p_{\text{rhythm}}}$ where p is a parameter such as neural timescale, mean firing rate, or first-spike jitter.

Statistical Analysis

All statistical analysis was performed with MATLAB (Mathworks). Non-parametric statistical tests were used in select cases where data samples did not meet the assumptions of parametric statistical tests. The Lilliefors test was used to check for normality of the data. Statistical significance was determined as the conjunction of separable variances between samples (i.e., p value) and the magnitude of the difference between samples (i.e., effect size). Effect sizes were estimated with Cohen's d for normally distributed data and with Cliff's delta for samples that did not conform to a normal distribution. Both p value and effect size changed in clear and predictable ways between hierarchical stages of processing. To the extent that assigning a threshold for statistical significance is useful, we chose the combination of a p value less than 0.05 and a Cohen's d greater than 0.5 (or Cliff's delta >0.3), which are traditionally assigned to medium-sized effects or greater. Multiple post hoc comparisons were corrected using Bonferroni correction.

Current Biology, Volume 31

Supplemental Information

**Inverted central auditory hierarchies
for encoding local intervals
and global temporal patterns**

Meenakshi M. Asokan, Ross S. Williamson, Kenneth E. Hancock, and Daniel B. Polley

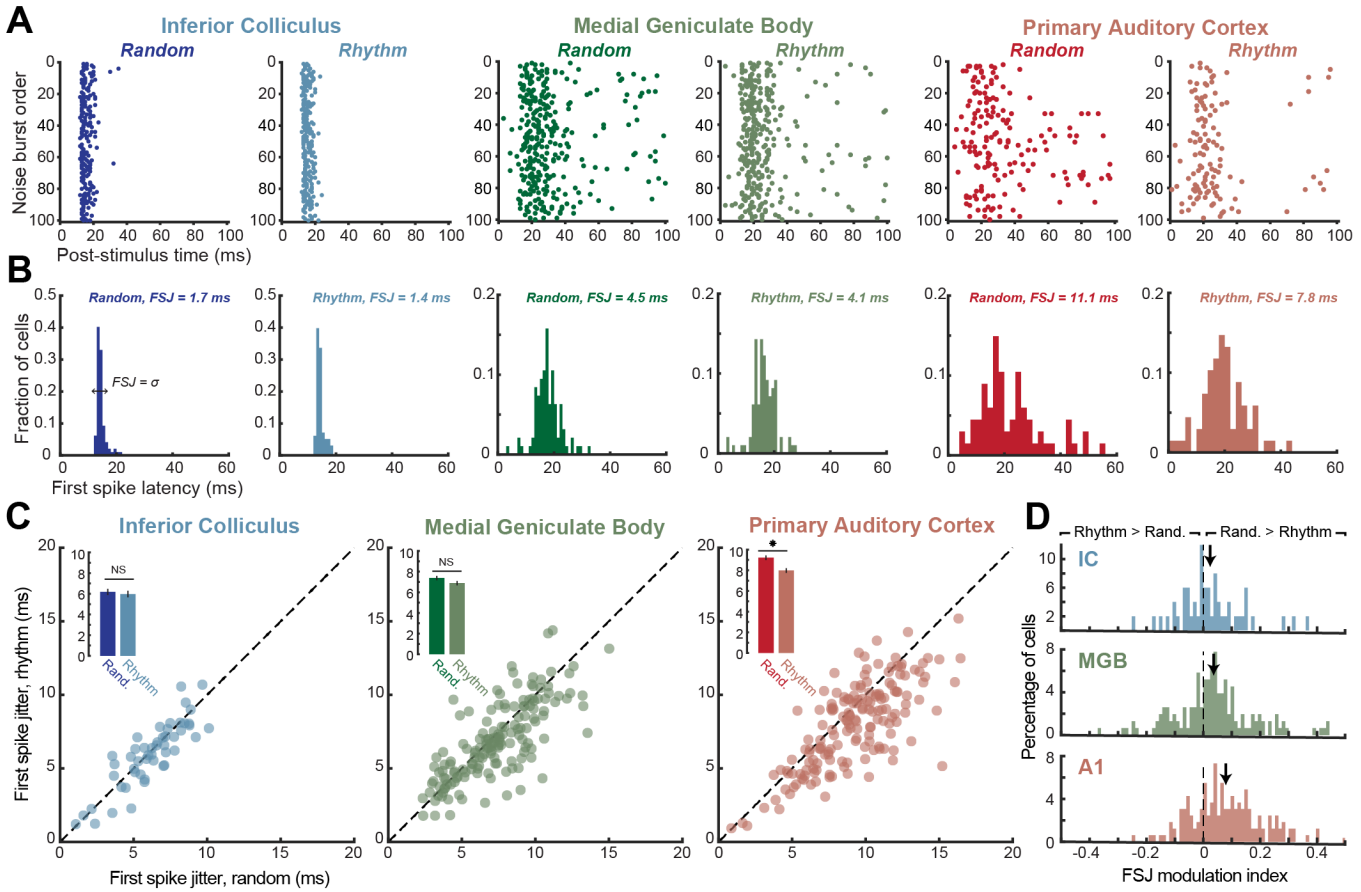


Figure S1: Variability in sound-evoked first spike latency is reduced during rhythmic epochs in A1, but is not changed subcortically, related to Figure 3.

A) Spike rasters from representative single units in the IC, MGB and A1 for 100 noise bursts presented in the random or rhythmic epochs (25 cycles \times 4 bursts per cycle). Gray shaded area denotes the timing of the 20ms white noise burst.

B) Histogram of first spike latencies for each noise burst from the units shown in a. For each unit, the first spike latency jitter (FSJ) was calculated as one standard deviation of the first spike latency distribution.

C) FSJ measurements from the random and rhythmic epochs are presented for each single unit from the IC, MGB, and A1. *Inset:* Mean \pm SEM of the neural timescale from the same sample. NS = not significant. Asterisk denotes statistical significance with a paired t-test ($p < 0.05$ and Cohen's $d > 0.5$); (IC, 0.21/0.18; MGB, 0.001/0.27; A1, 3×10^{-12} /0.59 for p-value/Cohen's d).

D) Histogram of the FSJ asymmetry index ($[\text{random} - \text{rhythm}] / [\text{random} + \text{rhythm}]$), where values < 0 indicate less jitter responses during random intervals and values > 0 indicate less jitter during rhythmic intervals. Arrows denote sample means. Dashed gray line indicates equivalent FSJ during random and rhythmic epochs. FSJ is significantly reduced during rhythmic noise burst arrangements in A1, but not MGB or IC (one-sample t-test against a population mean of 0, IC, 0.15/0.2; MGB, 0.001/0.26; A1, 8×10^{-14} /0.64 for p-value/Cohen's d).

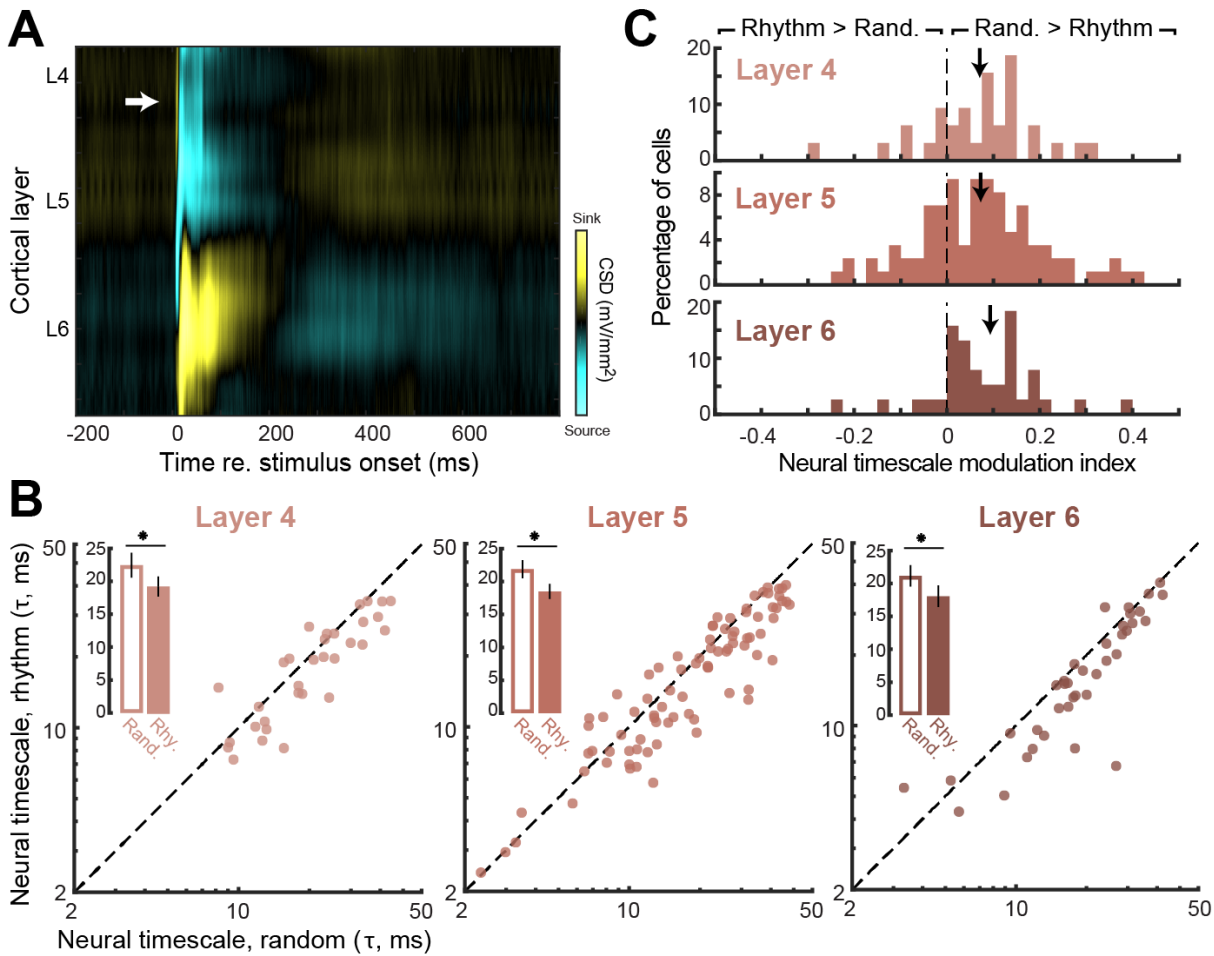


Figure S2: Equivalent dampening of neural timescales during rhythmic interval arrangements across cortical layers, related to Figure 3.

A) Extracellular recordings were made from layer 4, 5 and 6 of A1 with a dual shank silicon probe with 32 channels on each shank. The current source density (CSD) was measured from each shank to assign each isolated single unit to a layer. White arrow in CSD trace identifies the early current sink in layer 4 that was used to assign units to layers.

B) Neural timescale measurements from the random and rhythmic epochs are presented for each single unit from layer 4 ($n = 32$), layer 5 ($n = 85$) and layer 6 ($n = 38$). *Inset:* Mean \pm SEM of the neural timescale from the same sample. Asterisk denotes statistical significance with a paired t-test ($p < 0.05$ and Cohen's $d > 0.5$); (layer 4, $0.002/0.6$; layer 5, $4 \times 10^{-7}/0.6$; layer 6, $6 \times 10^{-5}/0.74$ for p-value/Cohen's d).

C) Histogram of neural timescale asymmetry index ($[\text{random} - \text{rhythm}] / [\text{random} + \text{rhythm}]$), where values < 0 indicate more dampened responses during random intervals and values > 0 indicate more dampened responses during rhythmic intervals. Arrows denote sample means. Dashed gray line indicated equivalent neural timescales during random and rhythmic epochs. Neural timescales are significantly reduced when noise bursts form rhythms than during random arrangements in all layers (one-sample t-test against a population mean of 0, layer 4, $0.003/0.57$; layer 5, $5 \times 10^{-6}/0.53$; layer 6, $2 \times 10^{-4}/0.67$ for p-value/Cohen's d).

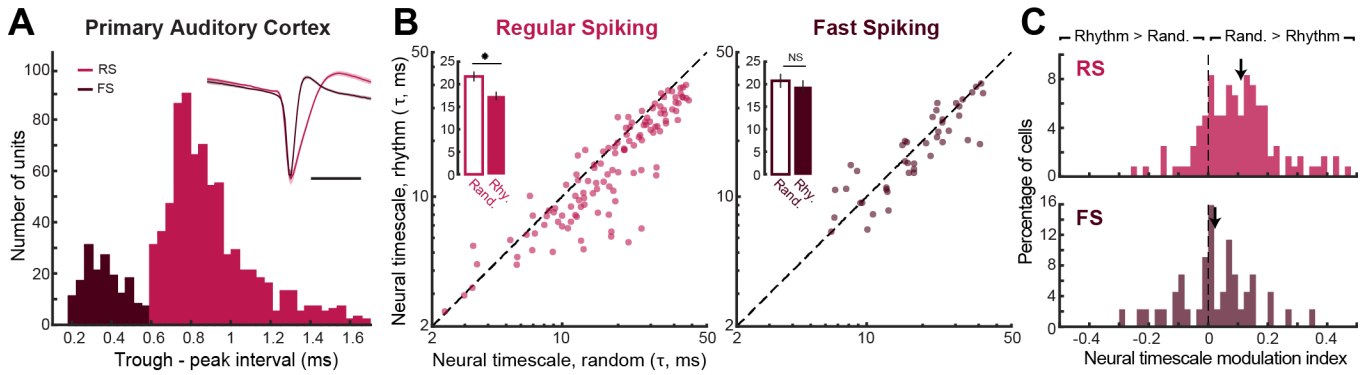


Figure S3: Neural timescales are dampened during rhythmic interval arrangements for regular spiking units, not fast spiking units, related to Figure 3.

A) Isolated A1 single units were classified as fast spiking (FS) or regular spiking (RS) (trough-to-peak delay < 0.6 ms or > 0.6 ms, respectively). Waveforms reflect mean \pm SEM FS and RS waveforms. Scale bar = 1ms.

B) Neural timescale measurements from the random and rhythmic epochs are presented for each single RS ($n = 120$) and FS ($n = 44$) unit. *Inset:* Mean \pm SEM of the neural timescale from the same sample. NS = not significant and the asterisk denotes statistical significance with a paired t-test ($p < 0.05$ and Cohen's $d > 0.5$); (RS, $3 \times 10^{-15}/0.83$; FS, $0.12/0.24$ for p-value/Cohen's d).

C) Histogram of neural timescale asymmetry index ($[\text{random} - \text{rhythm}] / [\text{random} + \text{rhythm}]$), where values < 0 indicate more dampened responses during random intervals and values > 0 indicate more dampened responses during rhythmic intervals. Arrows denote sample means. Dashed gray line indicated equivalent neural timescales during random and rhythmic epochs. Neural timescales are significantly reduced when noise bursts form rhythms than during random arrangements in RS units, but not FS units (one-sample t-test against a population mean of 0, RS, $1 \times 10^{-13}/0.76$; FS, $0.24/0.17$, for p-value/Cohen's d).

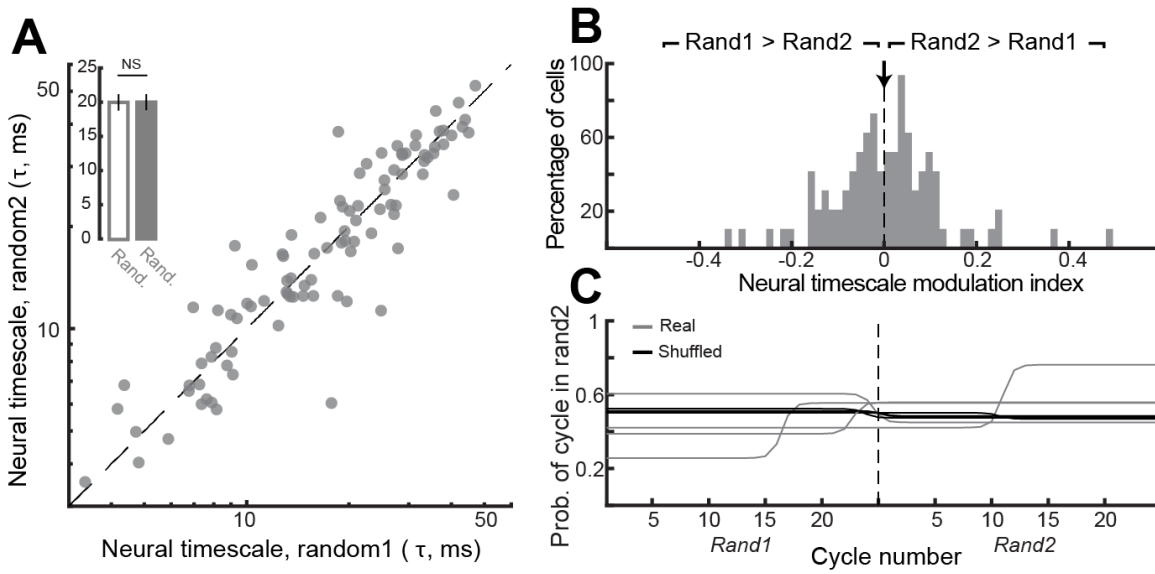


Figure S4: No change in A1 spiking timescales or decoder output when random interval arrangements are used throughout, related to Figure 3.

A) Neural timescale measurements from the two random epochs are presented for 96 single units, each recorded with a cycle size of 4 burst intervals. *Inset:* Mean \pm SEM of the neural timescale from the same sample. NS = not significant (0.9/-0.01 for p-value/Cohen's d).

B) Histogram of neural timescale asymmetry index. Arrows denote sample means. Dashed gray line indicated equivalent neural timescales during random and rhythmic epochs. Neural timescales are not modulated (one-sample t-test against a population mean of 0, 0.45/0.01 for p-value/Cohen's d).

C) Decoder output (Mean \pm SEM indicated by the thicker line with shaded error bar) for ensembles in A1 (4 mice; repeated measures ANOVA, $F = 0.75$, $p = 0.87$ for main effect for cycle number and $F = 1.94$, $p = 0.001$ for cycle number \times trial type interaction term)