A new theoretical framework jointly explains behavioral and neural variability across subjects performing flexible decision-making

Authors: Marino Pagan (1), Vincent D Tang (1), Mikio C. Aoi (1,2), Jonathan W. Pillow (1), Valerio Mante (3,4), David Sussillo (5,6), Carlos D. Brody (1,7)

Affiliations: (1) Princeton Neuroscience Institute, (2) Department of Neurobiology & Halıcıoğlu Data Science Institute, University of California, San Diego, (3) University of Zurich, (4) ETH Zurich, (5) Department of Electrical Engineering, Stanford University, Stanford, CA, USA, (6) Wu Tsai Neurosciences Institute, Stanford University, Stanford, CA, USA, (7) HHMI

Correspondence: mpagan@princeton.edu, brody@princeton.edu

Abstract

The ability to flexibly select and accumulate relevant information to form decisions, while ignoring irrelevant information, is a fundamental component of higher cognition. Yet its neural mechanisms remain unclear. Here we demonstrate that, under assumptions supported by both monkey and rat data, the space of possible network mechanisms to implement this ability is spanned by the combination of three different components, each with specific behavioral and anatomical implications. We further show that existing electrophysiological and modeling data are compatible with the full variety of possible combinations of these components, suggesting that different individuals could use different component combinations. To study variations across subjects, we developed a rat task requiring context-dependent evidence accumulation, and trained many subjects on it. Our task delivers sensory evidence through pulses that have random but precisely known timing, providing high statistical power to characterize each individual's neural and behavioral responses. Consistent with theoretical predictions, neural and behavioral analysis revealed remarkable heterogeneity across rats, despite uniformly good task performance. The theory further predicts a specific link between behavioral and neural signatures, which was robustly supported in the data. Our results provide a new experimentally-supported theoretical framework to analyze biological and artificial systems performing flexible decision-making tasks, and open the door to the study of individual variability in neural computations underlying higher cognition.

In our daily lives, we are often required to use context or top-down goals to select relevant information from within a sensory stream, ignore irrelevant information, and guide further action. For example, if we hear our name called in a crowded room and our goal is to respond based on the identity of the caller, the *frequencies* in the sound will be an important part of driving our actions; but if we wish to first turn towards to the caller, regardless of who they might be, information about *location,* of the very same sound, will be the most relevant to our actions. As with other types of decisions, when the evidence for or against different choices is noisy or uncertain, accumulation of many observations over time is an important strategy for reducing noise[1–4]. To study the neural basis of such context-dependent selection and accumulation of sensory evidence, we trained rats on a novel auditory task where, in alternating blocks of trials, subjects were cued to determine either the prevalent location ("LOC") or the prevalent frequency ("FRQ") of a sequence of randomly-timed auditory pulses (Fig. 1a). The relative rates of left vs. right and high vs. low pulses corresponded to the strength of the evidence about LOC and FRQ, respectively (Fig. 1b). These relative rates were chosen randomly and independently on each trial, and were used to generate a train of pulses that were maximally randomly-timed,

i.e., Poisson-distributed. Correct performance requires selecting the relevant feature for a given context, accumulating the pulses of evidence for that feature over time, and ignoring the irrelevant feature. Many rats were trained to good performance on this task using an automated training procedure (Fig. 1c; training code available at https://github.com/Brody-Lab/flexible_decision_making_training), with most rats learning the task in a timespan between 2 and 5 months (Extended Data Fig. 2g). Attained performances were similar to those of macaque monkeys performing analogous visual tasks[2]. Rats learned to associate the audio-visual cue presented at the beginning of each trial with the correct task context, and were able to switch between selected stimulus features within ~4 trials of a new context block (Extended Data Fig. 1e). We reasoned that the highly random yet precisely known stimulus pulses, together with large numbers of trials and subjects, would provide us with statistical power to characterize both behavioral[5] and neural responses.
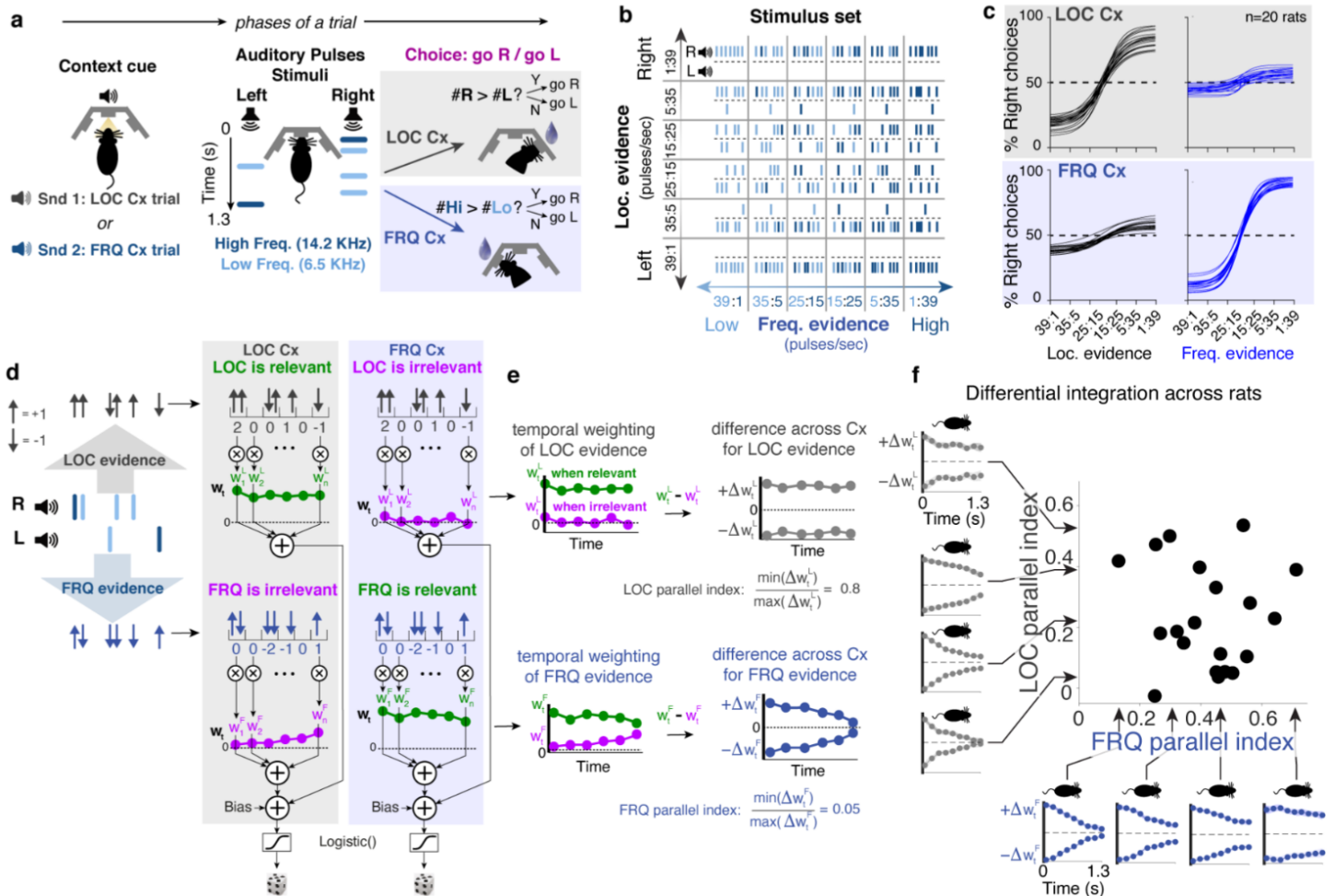
Figure 1. In a flexible information selection and integration task, rats trained to similarly good performance nevertheless display high variability in context-dependent behavioral features. a) Task design. Each trial starts with a sound indicating current context, either location (**LOC**) or frequency (**FRQ**). Rats are then presented with a 1.3 sec-long train of randomly-timed auditory pulses. Each pulse is played either from a Left or Right speaker, and is either Low- (6.5 KHz; light blue) or High- frequency (14 KHz; dark blue). In LOC trials, subjects are rewarded if they turn, at the end of the stimulus, towards the side that played the total greater number of pulses, ignoring pulse frequency; in FRQ trials, subjects should turn Right(Left) if there was a greater number of High(Low) pulses, ignoring pulse side location. Because the two features are independent, an identical stimulus can be associated with opposite responses in the two contexts. **b)** Schematic of the stimulus set used. The relative rates of left:right and high:low pulses correspond to the strength of the evidence about location and frequency, respectively. The relative rates are chosen randomly on each trial (while keeping the sum of the two rates fixed at 40 pulses/sec), and are then used to generate a Poisson-distributed pulse train. **c)** Psychometric curves showing good performance of 20 rats after training (n>120,000 trials for each rat): In the LOC context, location evidence strongly affects choices, while frequency evidence has only small impact; and in the FRQ context, frequency evidence strongly affects choices, while location evidence has limited impact. Curves are fits to a logistic function; see Methods. **d)** A given stimulus can be thought of as providing a train of positive and negative pulses of location evidence, as well as a train of pos/neg pulses of frequency evidence. Moreover, a stimulus can be presented in the LOC context (grey background) or in the FRQ context (blue background). A behavioral model of choices in a given context computes a weighted sum of net LOC evidence from each timepoint (50 ms bins) plus a weighted sum of net FRQ evidence from each timepoint, and then passes the net sum through a logistic function, to produce an estimate of the subject's probability of choosing Right. The values of the weights are fit to best match the experimental data, and indicate the strength with which that type of information, at that point in the stimulus, influences decisions. **e)** (Top:) The difference in the LOC weights when LOC evidence is relevant minus when it is irrelevant (green minus purple = Δw^L_t, the "differential behavioral kernel") is an estimate of how the temporal weighting of LOC evidence differs across the two contexts, i.e., its context-dependence. Both $+\Delta w^L_t$ and $-\Delta w^L_t$ are shown for better visibility of the shape of Δw^L_t as a function of time. The ratio of the smallest Δw^L_t over the largest Δw^L_t ("parallel index") is a measure of how parallel Δw^L_t is w.r.t. to the time axis. (Bottom:) Similarly for FRQ evidence. **f)** Example differential kernels for 4 rats, and scatterplot of LOC vs FRQ parallel indices across all n=20 well-trained rats. Despite good performance for all rats, there is high variability in the parallel indices across rats, and the parallel LOC index is not correlated with the parallel FRQ index.

Individual variability of behavioral kernels

To assess the temporal dynamics of context-dependent evidence accumulation, we measured the relative strength that pulses of evidence from each timepoint in the stimulus had on each subject's choice. We fitted a behavioral model in which a weighted sum of net LOC evidence from each timepoint, plus a weighted sum of net FRQ evidence from each timepoint, were passed through a logistic function to produce the probability that the subject will choose Right on that trial (Fig. 1d). The weights, chosen to best fit the experimental data, are estimates of the weight the subject placed on evidence from each timepoint. For a given rat, four time-dependent sets of weights were retrieved, for LOC and for FRQ evidence in each of the LOC and FRQ contexts. For good task performance, the weights when the evidence is relevant should overall be larger than when it is irrelevant. But the shape of the time-dependence could vary across individuals.

To focus on context-dependent effects, we examined "differential behavioral kernels", which we defined as the difference, across the two contexts, in the time-dependent weights of a given type of evidence (either LOC or FRQ; Fig. 1e). Indeed, even though performance was similarly high across our n=20 rats (Fig. 1c), we found a high degree of heterogeneity in the shape of the differential behavioral kernels (Fig. 1f; Ext. Data Fig. 3). The shape of the differential kernels spanned a continuum, from a temporally "flat" shape, meaning that context has a similar impact on pulses presented at any time point, to a shape that converged towards zero near the end of the stimulus, implying that the weight on choices of the latest pulses does not depend on whether they are relevant or irrelevant. These different kernel shapes were quantified by a "parallel index" (Fig. 1e). The kernels were consistent in individual rats across sessions (Ext. Data Fig. 4). Across rats, the shape of LOC differential kernels was not correlated with the shape of FRQ differential kernels (Fig. 1f). Thus, context-dependent temporal dynamics of evidence accumulation differed markedly across individual rats, and even within individual rats for the two different types of evidence, a result consistent with single-context findings in humans[6,7]. What underlies this variability across individuals?

Targeted dimensionality reduction (TDR) of trial-based neural dynamics in rat and macaque rule out two classes of models

What neural mechanisms allow different stimulus features to drive decisions in the different contexts? Some previous studies have proposed "early gating" of irrelevant information[8–12]. In this model, top-down signals to early sensory brain regions block irrelevant information from reaching more anterior, putative decision-making regions. This would predict that in trials where a given feature is irrelevant, little or no information about that feature would be found in decision-making regions-- whether at the single-neuron level, or whether at a more statistically sensitive, neural population level (Fig. 2a,b). However, monkey studies, recording from a cortical region closely associated with decision-making, the Frontal Eye Fields (FEF)[13,14], found no suppression of irrelevant feature information[2,15]. Fig. 2c shows this for the data of ref. 2 (p<0.001; reanalyzed using the methods of ref. [16]). In rats, the Frontal Orienting Fields (FOF) are a cortical region thought to be involved in decision-making for orienting choice responses[17,18], and have been suggested as homologous or analogous to macaque FEF[17,19,20]. Consistent with a key role for the FOF in our task, bilateral optogenetic silencing of rat FOF demonstrated that it is required for accurate performance of the task (Extended Data Fig. 5; n=3 rats). We implanted tetrodes into the FOF and also targeting another frontal region, the medial prefrontal cortex (mPFC), and we recorded from n=3532 putative single neurons during n=199 sessions from n=7 rats while they performed the task of Fig. 1. Carrying out the same analysis that had been applied to the monkey data[2,16], we also found no suppression of irrelevant feature information (p<0.01; Fig. 2d). The striking qualitative similarity between the rat (Fig. 2d) and monkey (Fig. 2c) traces suggests that the underlying neural mechanisms in the
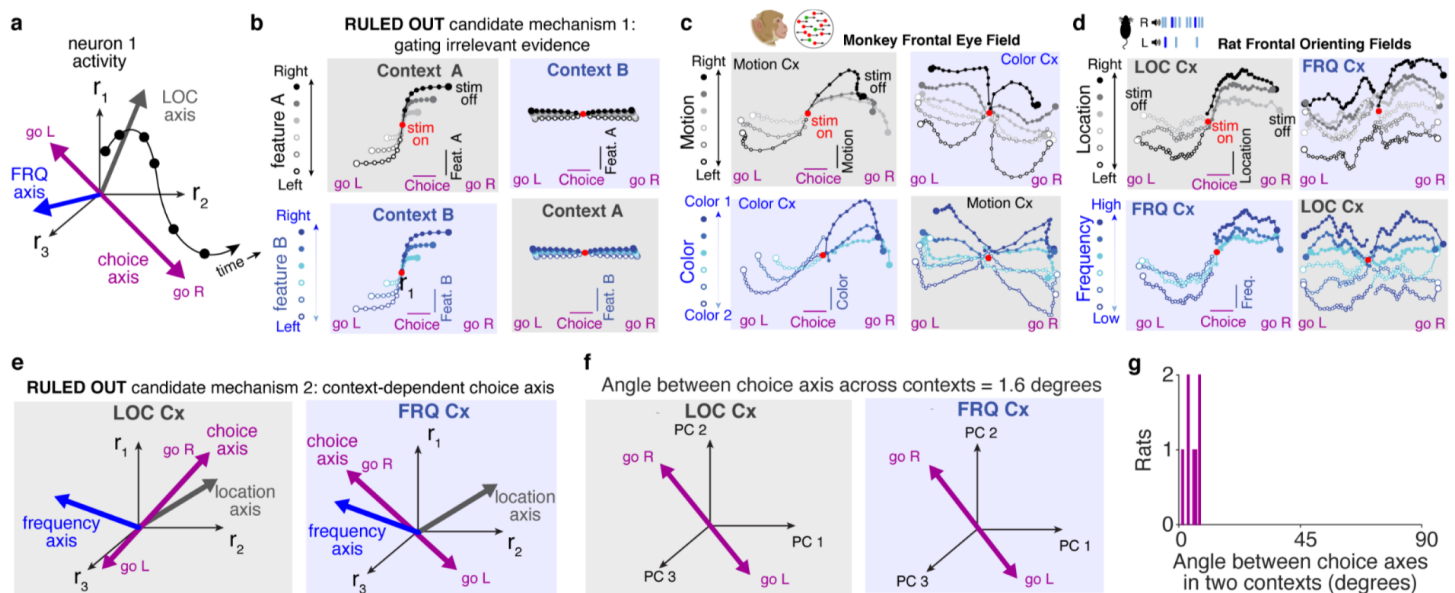
Figure 2. Frontal population dynamics in both rat and macaque are incompatible with early gating of irrelevant evidence, and instead support line attractor dynamics that are parallel across the two contexts. a) Illustration of analysis of population dynamics. The activity of many neurons at a given point in time corresponds to a point in high-dimensional "neural space". As activity evolves, this traces out a trajectory over time. This population trajectory can be projected onto axes chosen to best encode momentary LOC or FRQ evidence, or to best predict the subject's choice. In this figure, we will consider analyses in which the traces and their projections are computed after averaging over trials of similar evidence strengths. **b)** Conceptual illustration of predicted dynamics for a model in which early, context-dependent gating blocks irrelevant information from reaching frontal cortices. Top row: Population traces for groups of trials sorted by the strength of a feature A (e.g., LOC or FRQ), after projecting onto the choice and feature A axes. When A is the relevant feature (left, "Context A"), traces show large separation in the feature A axis, reflecting information about feature A in the population. In contrast, when feature A is irrelevant (right, "Context B"), no information about context A reaches frontal cortices, and separation along the feature A axis is negligible. Bottom row: same pattern for feature B. **(c,d)** In contrast to the early gating prediction (panel b), both monkey (panel c) and rat (panel d) data display similar feature evidence separation when the feature is irrelevant (right columns) as when it is relevant (left columns). **c)** Trial-based population dynamics in the Frontal Eye Fields (FEF) of macaque monkeys performing a visual task requiring context-dependent accumulation of either motion or color evidence (Mante et al., 2013). Same conventions as panel b, with feature A being motion and feature B being color. **d)** Trial-based population dynamics across contexts in the Frontal Orienting Fields (FOF) of rats performing the task depicted in Fig. 1. Same conventions as panel b, with feature A being LOC and feature B being FRQ. **e)** Illustration of an alternative model, not relying on early gating, but relying on context-dependent changes in the direction of the choice axis. In the LOC context (left), the choice axis is aligned to the direction of location input; in the FRQ context (right) the choice axis is aligned to the direction of frequency input. **f)** Close aligment of choice axes, computed separately for LOC and FRQ context trials, using data from the stimulus presentation period only. The two choice axes are shown in the two-dimensional plane that captures most variance about choice, across all trials regardless of context. Data is for all rats together. **g)** Angle between choice axes for the two contexts, as in panel f, for each individual rat (see also Extended Data Fig. 7).

two species may be similar enough that an active exchange of ideas between studies in the two species will be very fruitful.

The population analyses in Fig. 2 identify a "choice axis" as the direction in neural space that best predicts the subject's upcoming choice (Fig. 2a). A second potential mechanism underlying our task would have the choice axis reorienting across contexts[21], to align more with the relevant evidence for each context (Fig. 2e). To test this hypothesis, we estimated the direction of the choice axis separately in LOC and in FRQ contexts (see Methods). Contrary to the prediction of the context-dependent choice axis model, we found the two choice axes to be very closely aligned with each other, both in data across rats (angle between them = 1.6 degrees; Fig. 2f), and for each rat individually (Fig. 2g, Extended Data Fig. 7). These results are consistent with the notion that the choice axes are highly aligned across contexts, as previously reported in monkey FEF[2] (but see [21]). In sum, as with the monkey data and visual task of Mante et al. 2013, our data, using a different species (rat) and different sensory modality (audition), led us to rule out both the early gating model and the context-dependent choice axis model. These conclusions are thus neither modality- nor species- specific. We

note that although our own data did not support the early gating model, the framework that we develop next does not assume that early gating has been ruled out. Instead, the framework incorporates early gating as part of the solutions described in it.

The space of possible remaining solutions is spanned by three distinct components with different behavioral and anatomical implications.

We now consider the general set of possible solutions. To examine the mathematics behind how pulses of relevant evidence move the system along the choice axis, while pulses of the same evidence, when irrelevant, have a comparatively much smaller effect (Fig. 3a), we follow existing work in taking dynamics around the choice axis to be well approximated by a line attractor[22,23], i.e. a closely-packed sequence of stable points. This follows from the idea that the position of the system on the choice axis corresponds to net accumulated evidence towards Right vs Left choice, and in temporal gaps between pulses of evidence, an accumulator must be able to stably maintain accumulated values (hence the stable points). A line attractor strongly constrains the possible dynamics: linearized dynamics in the absence of external inputs for $\mathbf{r}$, the vector of neurons' firing rates,

$$\frac{d\mathbf{r}}{dt} = \mathbf{Mr}$$

necessarily have one eigenvalue of the matrix $\mathbf{M}$ equal to zero (corresponding to the eigenvector pointing along the line attractor, where all points are stable; by convention, this will be the zero[th] eigenvalue), and all other eigenvalues must have a negative real part (because it is an attractor; see Extended Discussion)[22]. In general, the eigenvectors of $\mathbf{M}$ need not be orthogonal to each other ("non-normal" dynamics).

An incoming pulse of evidence will perturb the system off the line attractor, to a position we will denote as $\mathbf{i}$, after which it will decay back (Fig. 3b). How far along the choice axis will the system have moved after the decay? Standard linear dynamics analysis suggests transforming $\mathbf{r}$ into eigencoordinates $\mathbf{r}^{(eigen)}$, in which each coordinate j evolves independently of the others, according to

$$r_j^{(eigen)}(t) = r_j^{(eigen)}(t=0) \, \exp(\lambda_j \, t),$$

where λ_j is that coordinate's eigenvalue, and $r_j^{(eigen)}(t=0)$ is the jth eigencoordinate of the initial perturbation $\mathbf{i}$. Since the zero[th] eigencoordinate has $\lambda=0$, it will remain constant. All others, with eigenvalues λ_j with negative real parts, will decay exponentially to zero, after which the system will have moved on the choice axis by a distance equal to that single remaining constant coordinate with $\lambda=0$. That zero[th] eigencoordinate is given by the dot product $\mathbf{s} \cdot \mathbf{r}$, where the row vector $\mathbf{s}$, known as the "selection vector[2]", is the first left eigenvector of $\mathbf{M}$. We note that since $\mathbf{s} \cdot \mathbf{r}(t) = \text{constant} = \mathbf{s} \cdot \mathbf{r}(t=0) = \mathbf{s} \cdot \mathbf{i}$, the dynamics of $\mathbf{r}(t)$ must be orthogonal to $\mathbf{s}$. That is, $\mathbf{s}$ characterizes the dynamics[2] (Fig. 3b; see Extended Discussion for full derivation).

For a pulse of evidence (LOC evidence, for example) to have a greater effect on choice in the LOC than the FRQ contexts, $\mathbf{s} \cdot \mathbf{i}$ must be greater in the LOC context than in the FRQ context. Across the two contexts, both $\mathbf{s}$ and $\mathbf{i}$ could vary. Thus, any network that solves the task must satisfy

$$\Delta(\mathbf{s} \cdot \mathbf{i}) = \mathbf{s}_{LOC} \cdot \mathbf{i}_{LOC} - \mathbf{s}_{FRQ} \cdot \mathbf{i}_{FRQ} > 0$$

(subscripts indicate context, and $\mathbf{i_{LOC}}$ and $\mathbf{i_{FRQ}}$ refer here to LOC evidence only; an equivalent analysis applies to FRQ evidence.) Our key theoretical insight is that this difference can be rewritten as the sum of three components (Fig. 3c).

$$\Delta(\mathbf{s}\bullet\mathbf{i}) = \frac{1}{2}\underbrace{(\mathbf{s}_{\text{Loc}}+\mathbf{s}_{\text{Frq}})}_{\overline{\mathbf{s}}}\bullet\underbrace{(\mathbf{i}_{\text{Loc}}-\mathbf{i}_{\text{Frq}})}_{\Delta\mathbf{i}} + \frac{1}{2}\underbrace{(\mathbf{i}_{\text{Loc}}+\mathbf{i}_{\text{Frq}})}_{\overline{\mathbf{i}}}\bullet\underbrace{(\mathbf{s}_{\text{Loc}}-\mathbf{s}_{\text{Frq}})}_{\Delta\mathbf{s}}$$

$$= \underbrace{\overline{\mathbf{s}}\bullet\Delta\mathbf{i}}_{\substack{\text{input}\\\text{modulation}}} + \underbrace{\Delta\mathbf{s}\bullet\overline{\mathbf{i}}}_{\substack{\text{selection vector}\\\text{modulation}}} \qquad (1)$$

$$= \underbrace{\overline{\mathbf{s}}\bullet\Delta\mathbf{i}_{\perp}}_{\substack{\text{indirect}\\\text{input}\\\text{modulation}}} + \underbrace{\overline{\mathbf{s}}\bullet\Delta\mathbf{i}_{\parallel}}_{\substack{\text{direct}\\\text{input}\\\text{modulation}}} + \underbrace{\Delta\mathbf{s}\bullet\overline{\mathbf{i}}}_{\substack{\text{selection}\\\text{vector}\\\text{modulation}}} \qquad (2)$$

where the overbar symbol represents the average over the two contexts and Δ indicates difference across contexts. "Indirect input modulation" is a change across contexts in the input $\mathbf{i}$, with the change orthogonal to the line attractor; "direct input modulation" is a change in the input $\mathbf{i}$ that is parallel to the line attractor; and "selection vector modulation" is a change in the selection vector $\mathbf{s}$. We note three parenthetical remarks that follow from this algebraic rewriting. First, early gating ($\mathbf{i}$=0 in the irrelevant context) is a special case within this framework (for examples of both direct and indirect input modulation from early gating, see examples 1 and 2 in Extended Discussion). Second, and despite the intuition implicit in Fig. 2e, the direction of the line attractor does not determine the impact on the final distance moved along it after a pulse: that distance is determined from the flow field (represented by $\mathbf{s}$) and the input vector $\mathbf{i}$ (equation (1); see also examples 4,5,6 in Extended Discussion)[2,22] . Third, the direction of the line attractor enters this framework in distinguishing indirect versus direct input modulation (equation 2). Assuming that the line attractors in the two contexts are parallel to each other (as follows from Fig. 2f,g and ref. 2), allows specifying a direction such that Δ$\mathbf{i}$ can be separated into components orthogonal and parallel to it. This distinction is shown below to be relevant to the speed of context-dependent effects and variability across individuals.

The action of each of the three components of equation (2) is illustrated in Figs. 3d,e,f. Since any network that solves the task can be described as a sum of these components, any network solution can be visualized as a point in barycentric coordinates, i.e. described in terms of distances from the vertices of a triangle (Fig. 3g). The position within this triangle identifies the relative weight of each of the three components in equation (2) for that solution.

Different solutions have different, and important, biological implications: for both indirect input modulation and selection vector modulation, the change on the choice axis after a pulse of evidence is initially the same for the two contexts, and a differential response develops only gradually (Fig. 3d,e, right panels); while for direct input modulation, the difference across contexts is immediate (Fig. 3f, right panels). This is indicated by the "fast" vs "slow" vertical axis in Fig. 3g, and has behavioral implications that will be examined in Fig 5. Separately, changes in the linearized input $\mathbf{i}$ can be achieved through changes in the activity of early sensory regions or modulations in communication from sensory to decision-making regions; while changes in

the selection vector (i.e., selection vector modulation) are changes purely in the dynamics of decision-making regions. Thus, where a network lies along the "input" vs "dynamics" tilted axis in Fig. 3g has anatomical implications as to the possible locus of context-dependence.

Seeking model networks that solve the task and have individual units with heterogeneous responses, as in the data (see Extended Data Fig. 6), Mante et al. (2013) used gradient descent methods to train recurrent neural networks (RNNs), and found that most successfully trained RNNs solved the task using context-dependent recurrent dynamics (i.e. the mechanism we call selection vector modulation). RNNs trained with gradient descent methods to solve our task are plotted in barycentric coordinates in Fig. 3h, which shows that indeed, successfully trained networks are densest near the selection vector modulation corner at bottom left (Fig. 3h), reproducing their result. Mante et al. observed a qualitative similarity in the targeted dimensionality reduction analysis of Fig 2b-d when applied to the RNNs, and when applied to their experimental data from macaque FEF. This prompted the suggestion of selection vector modulation as the leading candidate for how the brain implements context-dependent decision-making. However, the insight in eqn. (2) allowed us to develop methods to create distributed, heterogenous RNNs that lie at any chosen point within the barycentric coordinates (Fig. 3i; Methods). Surprisingly, we now find that when analyzed using targeted dimensionality reduction, RNNs at *any* point within the barycentric coordinates, not only those close to the selection vector modulation corner, produce traces that are qualitatively similar to the experimental data (Fig. 3j).

Thus, while targeted dimensionality reduction (TDR) trial-based analyses rule out early gating and context-dependent changes in the choice axis (Fig. 2), the large space of remaining solutions, spanned by the three components in equation (2) and with very significant differences and implications across the encompassed solutions, is not easily differentiated by trial-based analyses. In contrast, Figs. 3d,e,f suggest that a pulse-based analysis may distinguish some solutions, particularly across the fast vs slow vertical axis of Fig. 3g. Furthermore, the fact that all points within these barycentric coordinates are equally valid solutions, and are all qualitatively similar to previous data (see Extended Data Fig. 8), suggests that different individuals could use different solutions within this space. Could this explain the behavioral variability described in Fig. 1f?
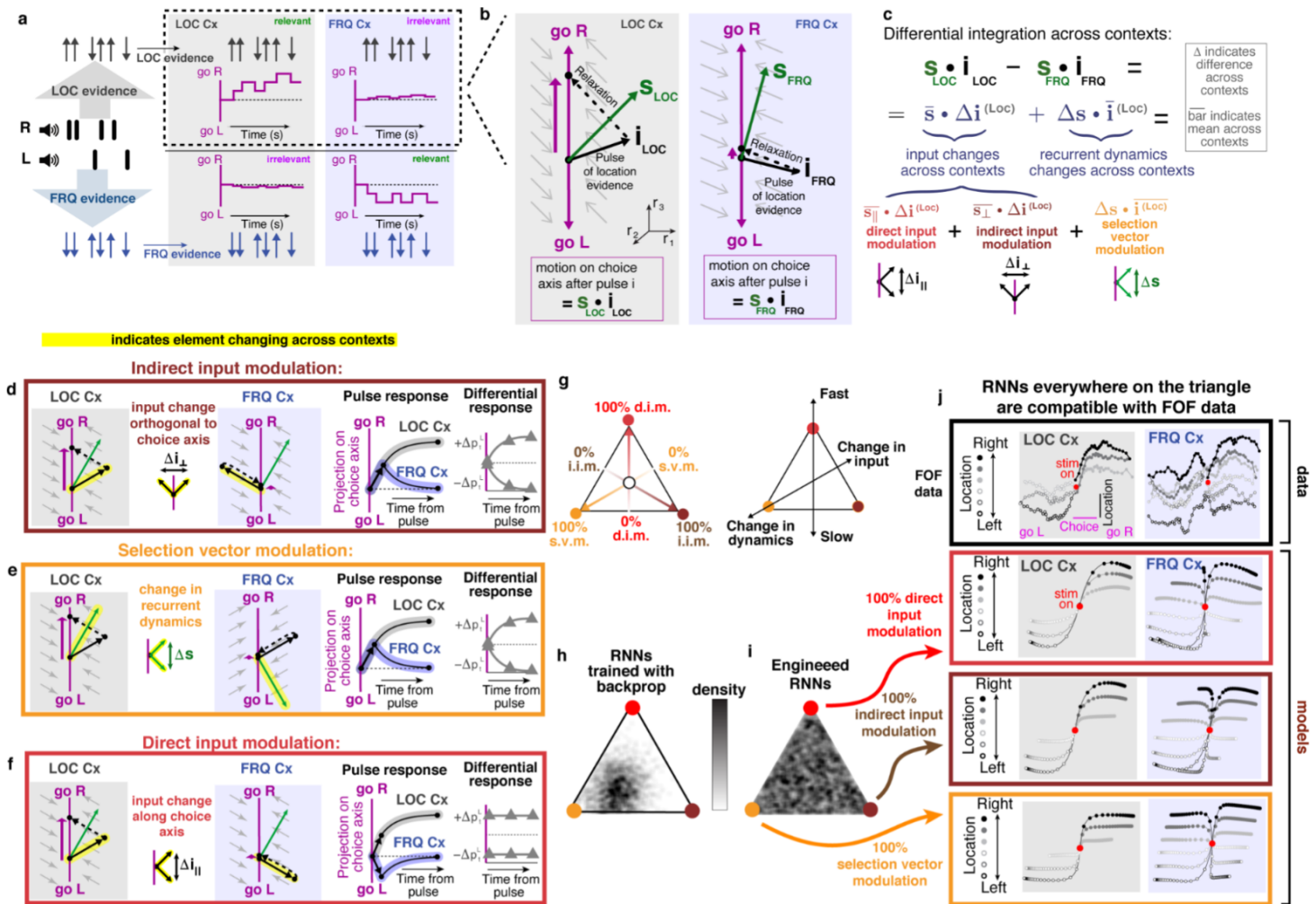
Figure 3. **Context-dependent evidence selection can be dissected into three distinct components, one of which is distinguished by its rapid selection, and all of which are consistent with previous data. a)** Conceptual illustration of task solutions. All solutions require pulses of relevant evidence to have an effect on the system's position along the choice axis, thus driving choices, while irrelevant evidence should have little effect on the choice axis. Subsequent panels use LOC evidence to illustrate analyses, but the same analyses apply to FRQ evidence. **b)** Assuming that the choice axis is a line attractor, the effect on choice axis position from a single pulse of evidence is given by the dot product $s \bullet i$ of the initial deflection caused by the pulse (input vector i), with a vector describing the relaxation dynamics back onto the line attractor, ("selection vector" s; dashed line shows the relaxation dynamics). Subscripts indicate context; since we are using LOC evidence as the example, the relevant context is the LOC context. The input vector and the selection vector can be different across contexts, but, following Fig. 2, the choice axis is taken to be parallel across contexts. **c)** Solving the task means that LOC evidence should have a bigger effect on choices in the LOC vs FRQ contexts, i.e., there must be a positive difference $s_{LOC} \bullet i_{LOC} - s_{FRQ} \bullet i_{FRQ}$. This difference can be rewritten as the sum of three components, each of which has distinct behavioral and anatomical implications, that together span the space of possible solutions. **d)** The "indirect input modulation" component describes changes in input, across the two contexts, that are orthogonal to the choice axis. Right: The effect on choice (i.e., projection onto the choice axis) is initially the same in the two contexts, differing only after the relaxation dynamics. The differential pulse response (relevant minus irrelevant pulse response for this type of evidence) grows from zero gradually. **e)** The "selection vector modulation" component describes changes in the recurrent dynamics. Right: Like indirect input modulation, the effect on choice is initially the same for the two contexts, and differs only after the relaxation dynamics. **f)** The "direct input modulation" component describes changes in the input vector that are parallel to the choice axis. A context-dependent effect on choice is visible immediately after the pulse of evidence, with an immediately large and sustained differential response. **g)** Any network that solves the task, after linearizing, can be expressed as a weighted sum of the three components, and visualized as a point in a triangle in barycentric coordinates. Right: the vertical axis on the triangle quantifies how quickly the choice response to a single pulse diverges according to context. A second axis captures whether context-dependent selection relies on changes of input or changes of recurrent dynamics. **h)** Distribution of 1000 100-unit recurrent neural networks (RNNs) trained to solve the task using backpropagation through time: starting from random initial points, after training the networks favored selection vector modulation, as found in Mante et al. (2013). **i)** Instead of using backprop, RNNs can be engineered to lie anywhere in the space of solutions. **j)** Networks across the entire space of solutions produce trial-averaged dynamics that are all qualitatively similar to the experimental neural data (computed as in **Fig. 2**; bottom three panels show analyses of three single example RNNs, taken from the three corners of the barycentric plot, top panel is reproduced from **Fig. 2d**; se also Extended Data Fig. 8). Trial-averaged dynamics can thus not distinguish across these different solutions, but panels d-g suggest that single pulse-evoked dynamics distinguish along the vertical axis of the barycentric plot.

Estimating pulse-evoked dynamics reveals heterogeneity in neural responses that is predicted by heterogeneity in "fast" vs "slow" network solutions.

A full characterization of the relevant linearized dynamics requires knowledge of the line attractor direction, the selection vector **s** in each context, and the components of the input vector **i** in each context that are parallel to **s** (since the effects of the input vector appear through **s●i**). The selection vector, however, is not directly observable. Nevertheless, we can leverage knowledge of the exact pulse timing on each trial to estimate the dynamics evoked by single pulses of evidence, and their projection onto the choice axis, which distinguishes solutions across the "fast" vs "slow" vertical axis of Fig. 3g. To do this, we modeled each neuron's firing rate as the sum of a time-dependent function (a "kernel") that is triggered and added by each pulse in the stimulus, plus time-dependent functions that accounted for other important factors: context, choice, and time (Fig. 4a). Four pulse-triggered kernels were retrieved, for LOC and for FRQ evidence, in each of the LOC and FRQ contexts. Recorded neuron firing rates were well-described with this approach (Ext. Data Fig. 9b). Analysis of the activity of units in model RNNs confirmed that pulse-triggered kernels estimated this way are highly similar to RNN responses to single evidence pulses presented in isolation (Ext. Data Fig. 9c,d), thus validating the approach. The pulse-triggered kernel for each neuron represents how its activity evolves over time due to a pulse. Similarly to how stimulus responses of many individual neurons can be described as a population trajectory in a joint space, where each axis corresponds to a neuron's activity (Fig. 2a), the pulse-triggered kernels of all individual neurons can be described as a *pulse-evoked* trajectory in the same neural space (Fig. 4b). The key difference is that in Fig. 2, the trajectory is aligned to the stimulus start (and is averaged across multiple different trains of evidence), whereas in Fig. 4b the trajectory is aligned to the presentation of a single pulse of evidence. As in the previous analysis, this trajectory was then projected onto the choice axis to obtain an estimate of the effect of a pulse of evidence on choice axis position (Fig. 4c, left). The difference in this response for evidence when it is relevant minus when it is irrelevant was defined as the "differential pulse response" (Fig. 4c, right). Application of this analysis to RNNs confirmed that estimated pulse responses closely approximate the actual responses to a pulse presented in isolation (Ext. Data Fig. 8c,d). The theory also predicts that the parallel index for the differential pulse response should be tightly linked to a network's position along the "fast" vs "slow" vertical axis in the barycentric coordinates (Fig. 3g); this was confirmed for RNNs engineered to employ different percentages of direct input modulation (Fig. 4d, Ext. Data Fig. 10a). Having validated the approach, we then applied it to the neural data we had recorded from the FOF. This revealed a large degree of heterogeneity in the parallel indices of differential pulse responses across rats and types of evidence (Fig. 4e), with the experimental traces resembling those produced by RNNs. We found no correlation across rats for LOC vs FRQ differential pulse kernels (Fig. 4e), an observation similar to our separate finding for differential behavioral kernels (Fig. 1f). These results are consistent with the notion that the variability in the observed pulse responses stems from individual animals implementing different combinations of solution components (Fig. 3g), even across the two types of evidence within individual rats.
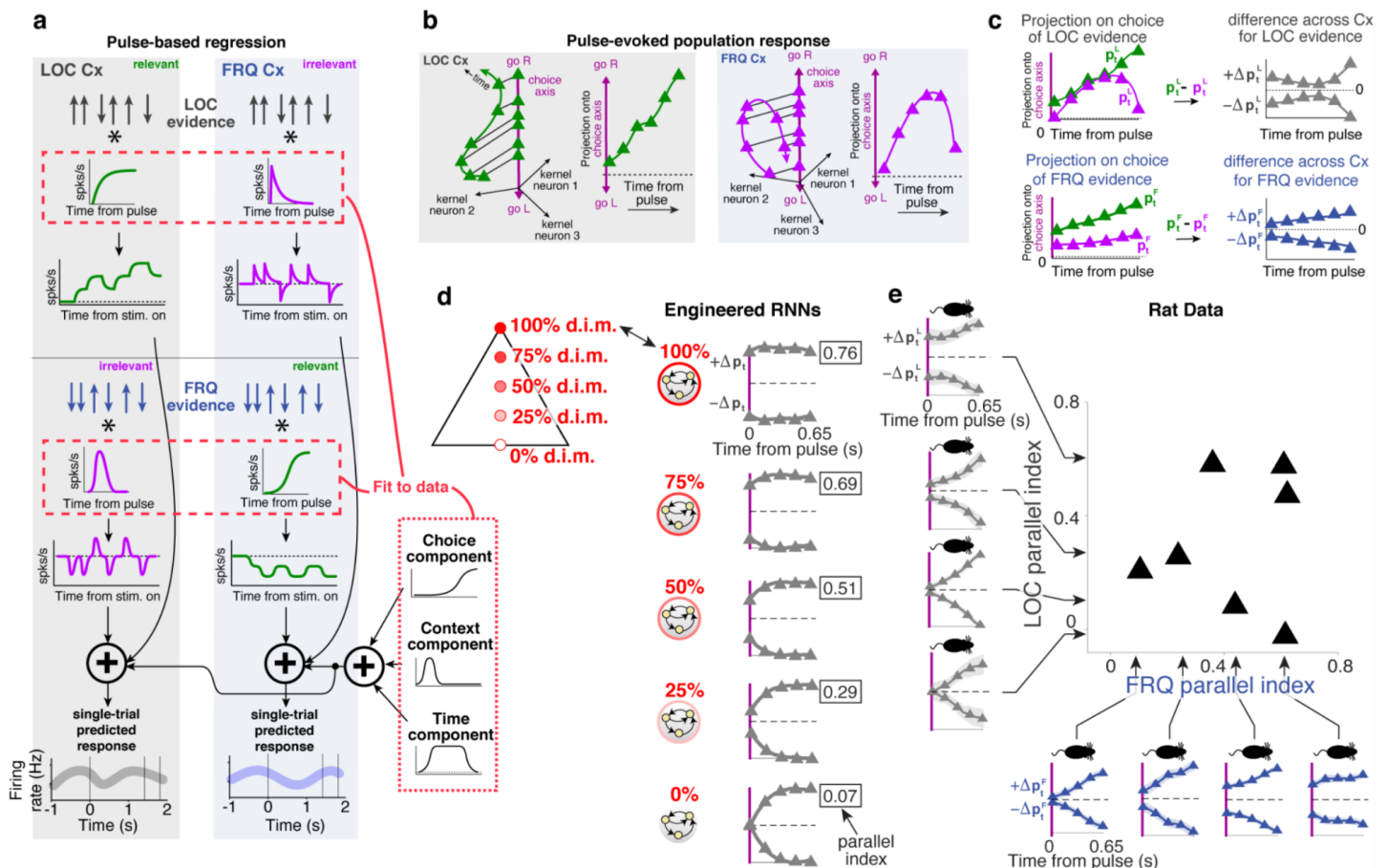
Figure 4. Experimental estimation of pulse-evoked dynamics on the choice axis. **a)** We first estimate each neuron's pulse-evoked response. To do this, the neuron's firing rate in a trial is modeled as the sum of pulse-response waveforms ("kernels," red dashed rectangles) that are triggered by each of the pulses in the stimulus, plus three time-dependent components (choice, context, and time elapsed during trial, dotted rectangles). Four different pulse-response kernels are fitted to the data, corresponding to LOC and FRQ evidence x LOC and FRQ context. Kernels are different for each neuron, fixed across trials, and fit to data from all trials. The best-fitting pulse-response kernels are the estimates of that neuron's response to each type of evidence pulse in each type of context on average. **b)** For each evidence x context combination, the pulse kernels across many individual neurons recorded from the same rat form a temporal trajectory in neural space. When projected onto the choice axis, this results in an estimate of the population's pulse-evoked dynamics on the choice axis. See Methods for estimation of the choice axis. **c)** Left: pulse-evoked dynamics on choice axis for LOC (top) and for FRQ evidence (bottom), when relevant (green) and irrelevant (magenta). Right: the difference between the relevant and the irrelevant dynamics is the "differential pulse response". **d)** Validation of the method on RNNs: the approach was applied to responses of units in RNNs engineered to lie along the vertical axis of the barycentric coordinates of **Fig. 3g**, ranging from 0% to 100% direct input modulation. The differential pulse response correspondingly ranges from slowly growing from zero (bottom; similar to **Fig. 3d** and **3f**) to having a large magnitude immediately after the pulse (top; similar to **Fig. 3e**). Number to right indicates parallel index for the corresponding plot, computed as in to Fig. 1e. **e)** Experimental data, with 4 example differential pulse responses for LOC evidence, 4 for FRQ evidence shown along the axes, and a scatterplot of LOC versus FRQ parallel indices from all n=7 rats. A broad range of parallel indices, with no immediately correlation between LOC and FRQ, can be seen across individuals.

Pulse-evoked neural dynamics recapitulate behavioral variability

Under the simplifying assumptions that the response to a pulse does not depend on time within a trial nor on previously presented evidence (assumptions supported by the good fit to the data of a model built with these assumptions, Figs. 4 and Extended Data Fig. 9b), the theory predicts a specific relationship between differential behavioral kernels and differential pulse responses: If T is the time at which position on the choice axis is read out to commit to a Right vs Left choice, then the impact on choices of a pulse at time t will follow the pulse-evoked movement along the choice axis after an interval T-t. For direct input modulation, with a differential pulse response that is immediate and sustained (Fig. 3f), the differential behavioral impact of a
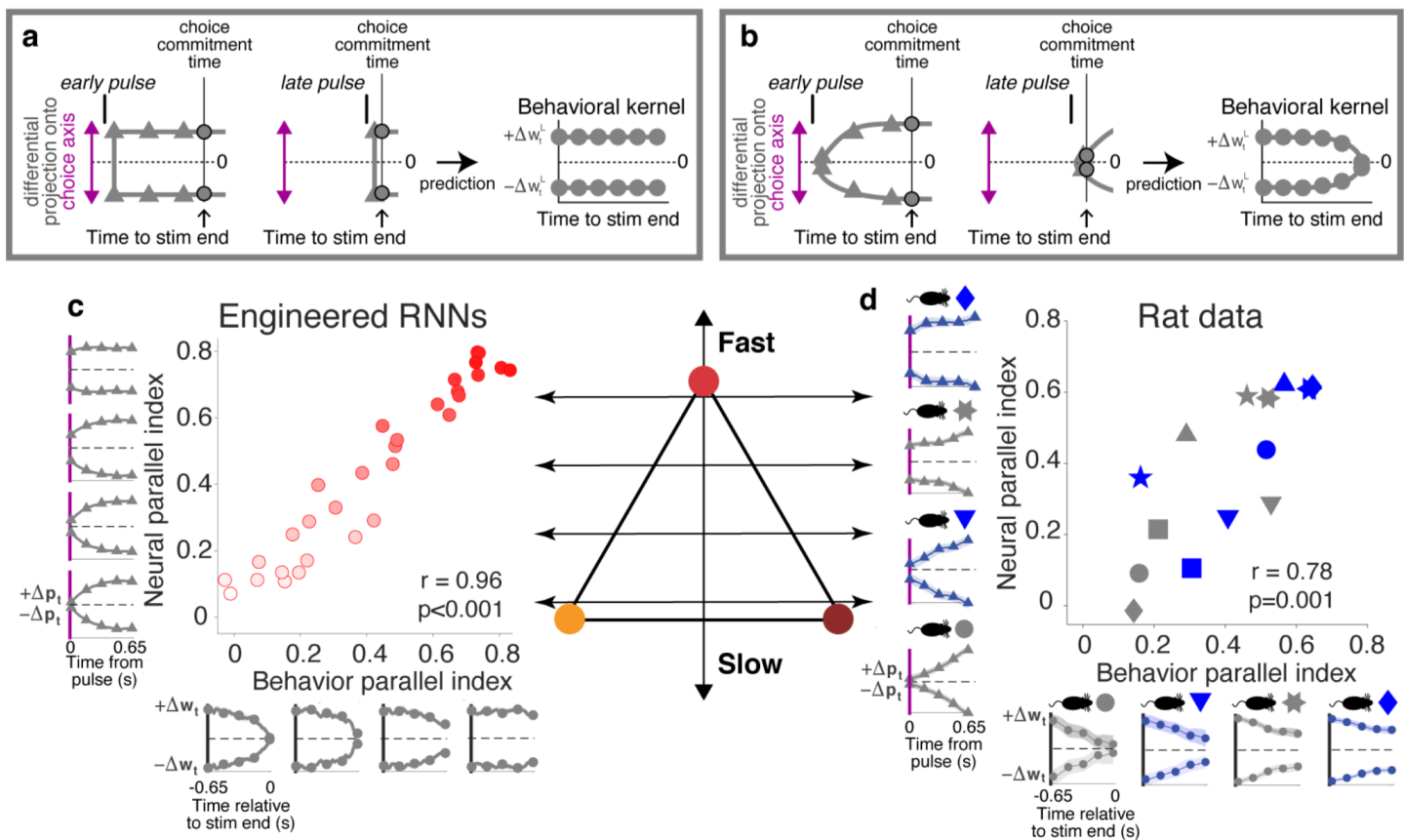
Figure 5. Theory predicts, and experimental data confirms, that variability in the neural parallel index should explain variability in the behavioral parallel index. a,b) Schematics of theoretical reasoning. **a)** In the case of a network using mostly direct input modulation, there will be immediate and sustained separation along the *neural* choice axis between relevant and irrelevant pulses (following **Fig. 3**). Thus the differential impact (across the two contexts) of a pulse on choice will not depend on whether the pulse is presented early or late relative to choice commitment. The temporally flat differential pulse response of the *neurons* (bottom left) thus results in a temporally flat differential *behavioral* kernel (bottom right). **b)** In contrast, if the network is using mostly selection vector modulation and/or indirect input modulation, pulses have a differential impact on choice only after the relaxation dynamics (**Fig. 3**). Thus pulses presented well before choice commitment will have a different impact in the relevant versus irrelevant contexts, while pulses presented immediately before choice commitment have no time to give rise to a differential impact. Thus, a gradually diverging differential pulse response (bottom left) results in a gradually converging differential behavioral kernel (bottom right). **c)** Synthetic data from n=30 engineered RNNs spanning the vertical axis of the barycentric coordinates (same colors as **Figs. 3,4**). Differential pulse responses were estimated as in **Fig. 4** and differential behavioral kernels as in **Fig. 1**. The models follow the theoretical prediction, with highly correlated parallel indices for the differential pulse responses (neural) and behavioral kernels (behavior). **d)** The theoretical (panel a) and modeling (panel b) prediction is also found in the experimental data: parallel indices for differential pulse responses, and for differential behavioral kernels, are highly correlated across rats (same conventions as panel c). Individual data point shapes indicate individual rats, LOC evidence analyses are in black, FRQ are in blue.

pulse should be the same whether it is presented close to, or long before, choice commitment. This should lead to a flat differential behavioral kernel (Fig. 5a). But for selection vector modulation or indirect input modulation, with differential pulse responses that grow only gradually from zero, the differential impact of a pulse will be small if presented shortly before choice commitment, and larger if presented longer before. This should result in a converging differential behavioral kernel (Fig. 5b). In other words, the shape of the differential behavioral kernel should be the reflection on the time axis of the differential pulse response. These two very different types of measures are thus predicted to have the same parallel index. We tested this prediction on RNNs engineered to solve the task using different amounts of direct input modulation. We computed both differential behavioral kernels, as in Fig. 1, and differential pulse responses, as in Fig. 4. As predicted, the parallel indices of the two were tightly correlated (Fig. 5c). We then tested whether a similar relationship existed for the rats' behavioral and neural experimental data. We found robust support in the data for the prediction that the two measures should be correlated (Fig. 5d, r=0.78, p<0.001), with the correlation also holding for LOC evidence alone (r=0.77, p<0.05) or for FRQ evidence alone (r=0.76, p<0.05). Thus, although

there is no correlation within the behavioral measure (Fig. 1f) or within the neural measure (Fig. 4e), the two are strongly correlated with each other (Fig. 5d). These results support both the overall theoretical framework, and the idea that variability in solutions across the barycentric coordinates of Fig. 3g is the common source underlying, and explaining, both neural and behavioral variability.

Discussion

In this work we combined high-throughput training of rats, a novel pulse-based task requiring context-dependent selection and accumulation of evidence, and pulse-based analyses of behavioral and neural data, to probe and confirm predictions of a new theoretical framework for flexible decision-making. The framework describes a space of possible solutions, and explains variability between and within individuals as variability within that space. Our theory stems from basic linearized dynamical systems analysis[23], where a simple algebraic rewriting of the mathematics behind context-dependent selection and accumulation (equations 1 and 2) led us to multiple insights: theoretical insights, defining the space of possible solutions (Fig. 3g); biological insights, describing the behavioral, neural, and anatomical implications of the different solutions; conceptual insights, identifying the underlying source that links neural and behavioral variability (Fig. 5d); and technical insights, allowing us to engineer recurrent neural networks that could not be constructed before, spanning the full space of solutions (Fig. 3h,i).

We describe our theoretical work as a "framework" because it does not specify particular network implementations. Instead, it defines a language describing the axes of the space of possible dynamical solutions and their characteristics. Each point in the space we have described could be implemented in multiple ways. Recent complementary work has focused on the study of network implementations[12,24,25].

Studies often center on findings that are common across subjects, and it is common practice to report the result for an "average subject". However, our results reveal a surprising degree of heterogeneity across, and even within, individual animals, underscoring the importance of characterizing the computations used by each individual subject. This issue may be of particular significance for cognitive computations, which are largely internal and therefore potentially subject to substantial covert variability across subjects. Here, studying how computations vary across subjects was made possible by two key methodologies. First, we used an efficient, automated procedure to train a sufficient number of rats to be able to observe and quantify cross-subject variability[5]. Second, we characterized each individual's computations by leveraging the statistical power afforded by a randomly-timed, pulse-based stimulus[5]. We used two independent methods that leveraged this pulsatile stimulus, one based on analysis of neural activity and another based on analysis of behavioral choices. The high correspondence between the two separate resulting measurements (Fig. 5d) indicates a link through a common underlying variable. Our framework describes how different, equally valid solutions can have different speeds with which the impact of an evidence pulse on a subject's choice becomes affected by the current context (vertical axis in Fig. 3g). We identify this speed as the common underlying variable linking the behavioral and neural measurements.

Whether or not early gating (i.e., blocking irrelevant evidence from reaching decision-making regions) accounts for context-dependent decision-making is an ongoing debate, with some studies providing evidence that it does (e.g., [11,12,26,27]), while others, including our own data (Fig. 2) provide evidence that it does not (e.g., [2,15,28,29]). Similar to variability across the vertical axis of the solution space of Fig. 3g, which we believe is a result of all of the encompassed solutions being capable of solving the task, solutions using or not using early gating, both of which are in the framework we have described (see examples 1 and 2 In Extended Discussion), are equally capable of solving the task. It is thus possible that there could be variability across tasks and individuals, perhaps even within them, regarding the use of early gating. Further work will resolve the relative

prevalence, or absence, of early gating. Further work will also be needed to resolve differences between our work and a recent study[21], using a different flexible decision-making task and focused on a different brain region, which found curved manifolds that rotated across different contexts.

Our work also provides a cautionary note, highlighting the fact that recurrent neural networks (RNNs) trained through gradient-descent methods, which are commonly used to model brain function[2,30–35], allow the discovery and exploration of some possible solutions, but need not comprise the full set of RNNs that are consistent with experimental data for a given phenomenon. In our work we found that gradient descent methods led towards only one corner of the full space of solutions (Fig. 3h). It was a deeper understanding of the mathematics behind solutions (equations 1 and 2), not gradient descent, that allowed us to engineer data-compatible RNNs across the full space of solutions (Fig. 3i,j).

Even though our experiments were carried out in rats, the similarity in the results of behavioral and neural analyses that could be carried out in common across two species, rats and monkeys (Fig. 1c, Fig. 2), suggests that conclusions reached from rat data may generalize to other species as well. A recent study[36] indicates that human subjects performing context-dependent decision-making process different stimulus features independently, in line with our result that subjects can use separate mixtures of components to select and accumulate each of the two features (Fig. 1f, Fig. 4e). Drawing parallels across species was greatly facilitated by adopting highly similar behavioral paradigms (Extended Data Fig. 1), and targeting putatively similar brain regions (e.g. monkey FEF and rat FOF).

Our theoretical framework indicates that a key part to understanding context-dependent, flexible behavior consists in unraveling the context-dependent interactions between sensory inputs and recurrent dynamics. For example, observing large context-dependent changes in the representation of sensory evidence, either in sensory or decision-making regions (leading to large $\triangle \mathbf{i}$ in equation 1) is not sufficient to conclude that these drive context-dependent decision-making: only input changes aligned to the direction in neural space capturing decision-making region recurrent dynamics ($\triangle \mathbf{i} \cdot \bar{\mathbf{s}}$) produce a context-dependent effect. Thus, while some studies have argued that early gating is indicated by a representation of evidence in decision-making regions that is weaker in the irrelevant context (i.e., smaller |**i**| in our terminology)[26], which in many cases may be correct, example 3 in the Extended Discussion demonstrates that it need not be so: a context with smaller |**i**| can counterintuitively be the one where **i** has the greater impact on decisions, because it has the larger $\mathbf{s} \cdot \mathbf{i}$. The emphasis on the interaction between **i** and **s** is closely related to the alignment of input and dynamics recently observed in the context of sensory learning[37]. Similarly, large changes in the recurrent dynamics across contexts ($\triangle \mathbf{s}$) will play no role in the context-dependent behavior if they occur in directions orthogonal to the average input direction $\bar{\mathbf{i}}$. For clarity, we must emphasize that in our framework, **i** represents the sensory input to decision-making brain regions when linearized around the context-dependent state of the system before sensory pulses arrive. Thus, in terms of their anatomical locus, context-dependent changes in the linearized input ($\triangle \mathbf{i}$) can be potentially created by context-modulation of early sensory regions, but could also occur purely in higher-order decision-making regions, through context-modulation of the point around which linearization will be calculated. In contrast, $\triangle \mathbf{s}$ refers exclusively to context-dependent changes in decision-making regions.

A limitation of our approach is that we are currently unable to discriminate between mechanisms relying on context-dependent changes of recurrent dynamics versus changes in the linearized sensory inputs (i.e. the oblique axis in Fig. 3g, right). A full characterization of the relevant neural dynamics will require estimation of the selection vector **s** (the vector that summarizes the key aspects of the network dynamics), for each context. Simultaneous recordings from large neural populations, combined with the application of recently-developed latent-based methods, such as LFADS[38] or PLNDE[39], which are designed to capture the dynamics underlying high-dimensional neural trajectories, may prove instrumental in future work in this direction. Another potential limitation stems from the possibility that recurrent dynamics might evolve more rapidly[40] than the current time

resolution in our measurements, leaving us unable to discriminate between contextual input modulation versus fast recurrent modulation. However, our results indicate that our analyses quantified the speed of evidence selection as smoothly varying across subjects (Fig. 1f; Fig. 4e), suggesting that in most subjects dynamics are slow enough to be captured with our method.

In sum, our work provides a new, general framework to describe and investigate neural mechanisms underlying flexible decision-making, and opens the door to the cellular-resolution study of individual variability in neural computations underlying higher cognition.

Acknowledgments

Author Contributions

M.P. and C.D.B. designed the experiment. M.P., V.M. and C.D.B. designed the automated training procedure. M.P. and V.D.T. performed the experiments. M.C.A. and J.W.P. developed the mTDR analysis. M.P., M.C.A. and J.W.P. designed the pulse-based analysis of neural data. All authors contributed to the conceptual development of the theory. M.P. and C.D.B. developed the mathematical framework. M.P. and D.S. trained and analyzed artificial neural networks. M.P. and C.D.B. wrote the manuscript after discussions among all authors. C.D.B. supervised the project.
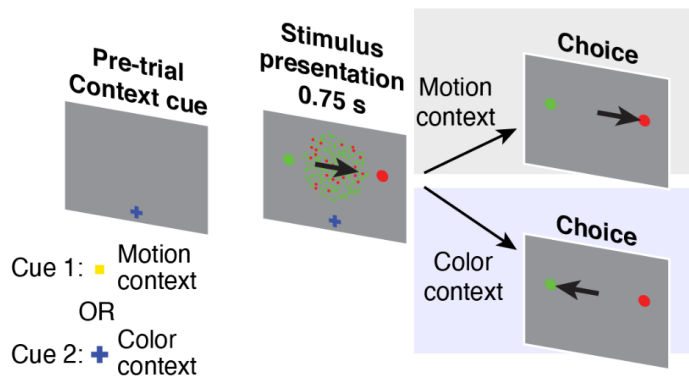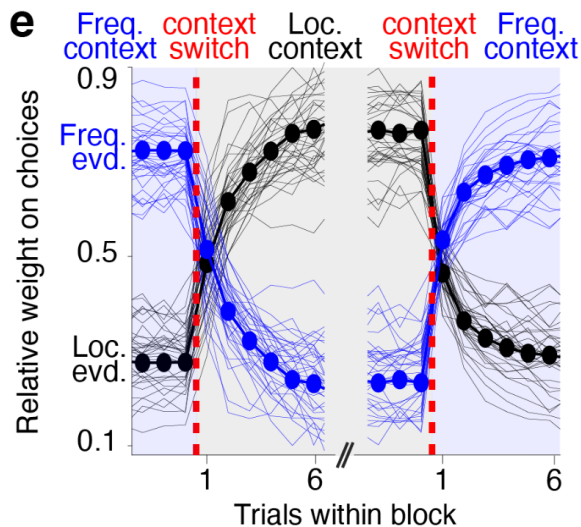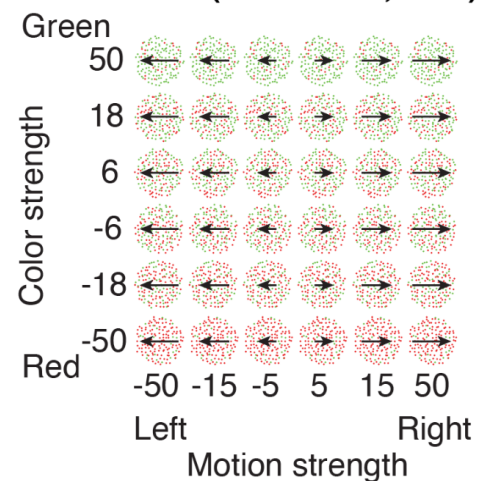
Extended Data Figures
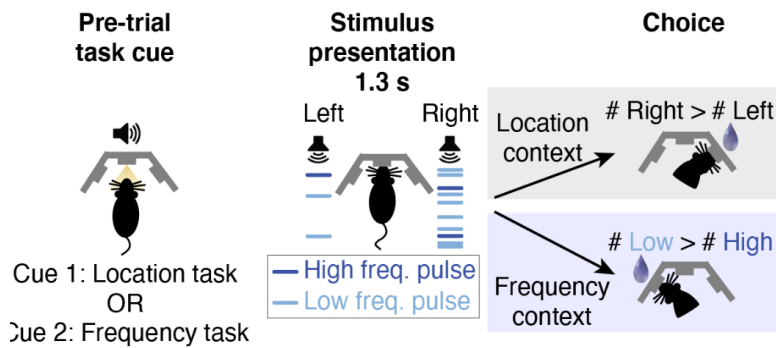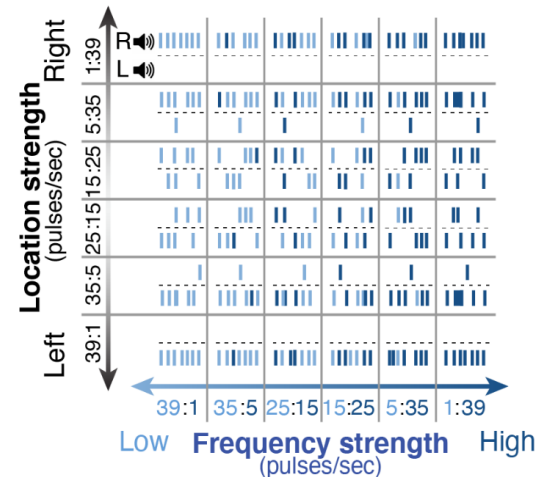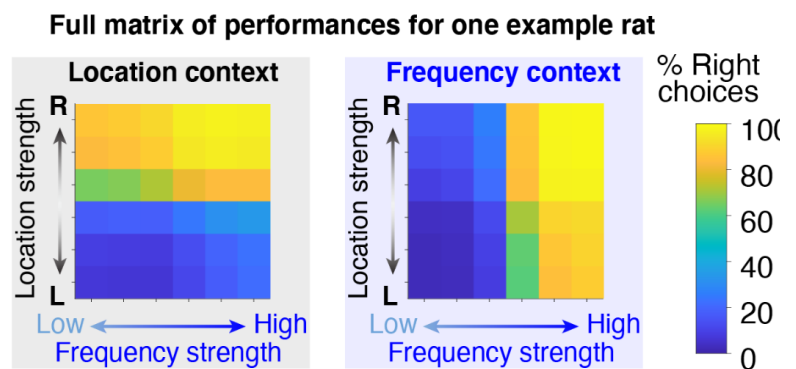
Extended Data Figure 1. **a-d)** Comparison of rat task and monkey task[2]. **a)** In the rat task, the subject is cued using an audiovisual stimulus, and is presented with a train of randomly-timed auditory pulses varying in location and frequency. In different contexts, the subject determines the prevalent location or the prevalent frequency of the pulses. **b)** Stimulus set for the rat task: strength of location and prevalent frequency are varied independently on each trial. **c)** In the monkey task, the subject is cued using the shape and color of a fixation dot, and is presented with a field of randomly-moving red and green dots. In different contexts, the subject determines the prevalent color or the prevalent motion of the dots. **d)** Stimulus set for the monkey task: strength of motion and prevalent color are varied independently on each trial. **e)** Rats rapidly switch between contexts. Performances saturate within the first 4-5 trials in the block. The weight of location and frequency evidence is computed using a logistic regression (see methods). Thin lines indicate individual rats, thick lines indicate the average across rats. **f)** Full matrix of behavioral performances for one example rat across the two contexts.
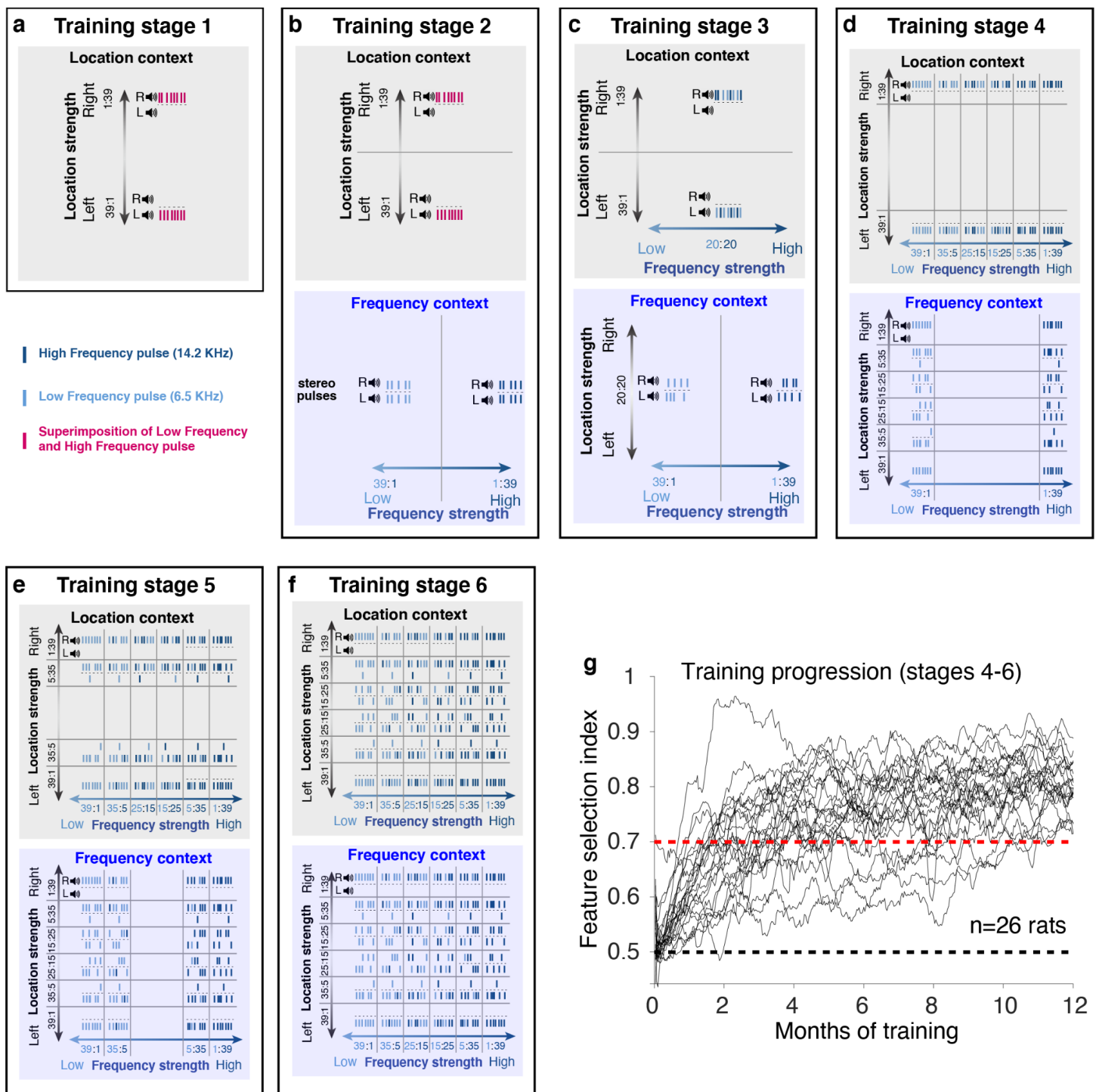
Extended Data Figure 2. **a-f)** Training procedure. **a)** Stage 1: rats are trained only on the location task, with strong location evidence and no frequency evidence (pulses consist of superimposed low and high frequency). The context cue is played before each trial. **b)** Stage 2: rats learn to alternate between the location and frequency context. In the frequency context rats are presented with strong frequency evidence and no location evidence (stereo pulses). **c)** Stage 3: introduction of pulse modulation. In the frequency context, pulses are now presented on either side (but with no prevalent side). In the location context, pulses are either high-frequency or low-frequency (but with no prevalent frequency). **d)** Stage 4: irrelevant information is introduced, but the relevant information is always at maximum strength. **e)** Stage 5: relevant information can have intermediate strength. **f)** Stage 6: relevant information can have low strength. **g)** Training progression. Most rats learn stages 1-3 in approximately 2 weeks, but it takes a much longer time to learn stages 4-6 because of the introduction of irrelevant evidence. The feature selection index quantifies whether rats attend to the correct feature and ignore the irrelevant feature (see methods). The black dashed line indicates chance, the red dashed line indicates the threshold performance to consider a rat trained. Most rats learn the task within 2-5 months.
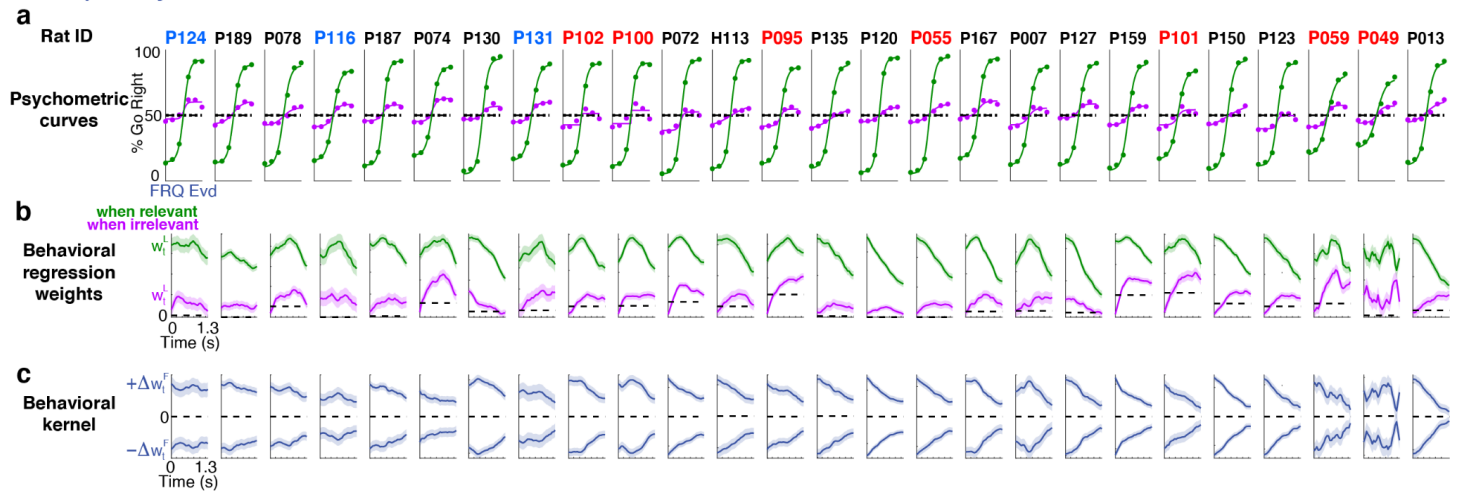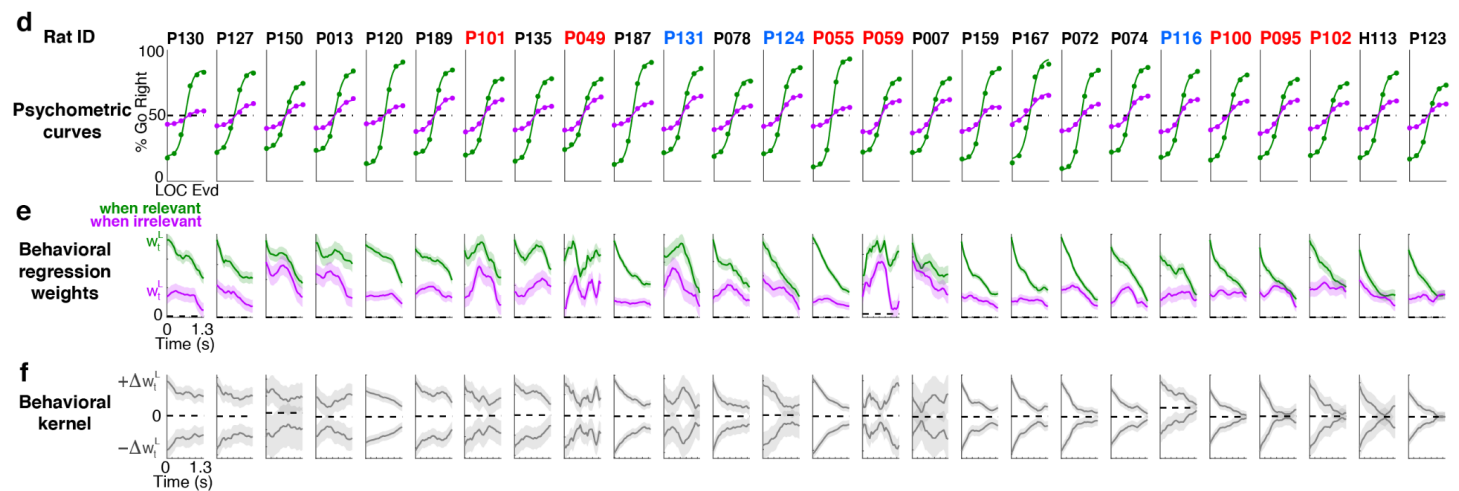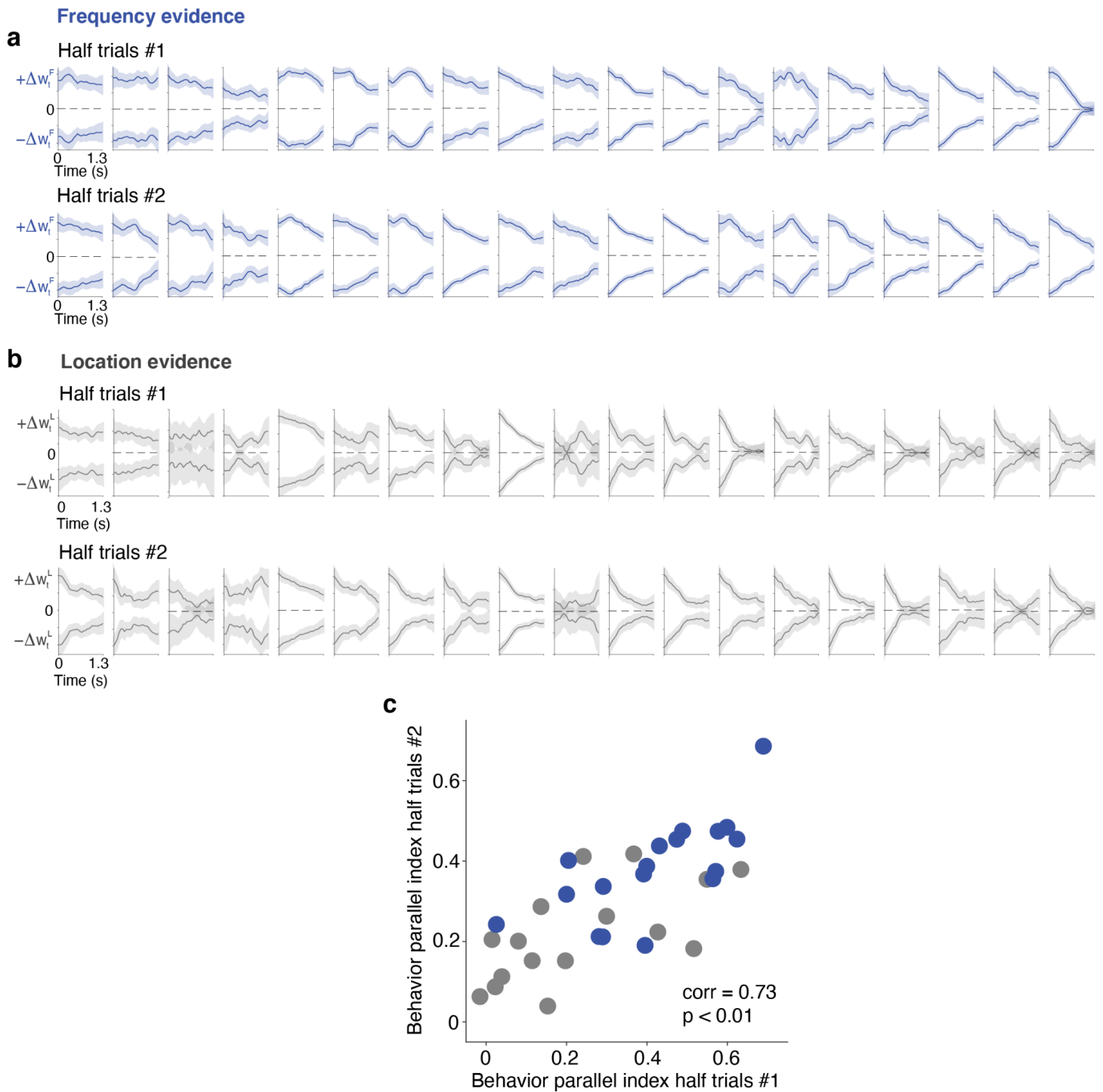
**Extended Data Figure 3**. Behavioral data for all rats. Rat ID color indicates whether rat was used for electrophysiology (red), optogenetics (cyan) or only for behavior (black). **a)** Psychometric curves for frequency evidence, measuring the fraction of right choices as a function of strength of frequency evidence (6 levels of strength, see Fig. 1b). Green indicates frequency context (relevant), purple indicates location context (irrelevant). **b)** Weights for frequency evidence computed using the behavioral logistic regression for each rat (see Fig. 1d); colors as in panel a. **c)** Differential behavioral kernel for frequency evidence across all rats. **d)** Psychometric curves for location evidence, measuring the fraction of right choices as a function of strength of location evidence (6 levels of strength, see Fig. 1b). Green indicates location context (relevant), purple indicates frequency context (irrelevant). **e)** Weights for location evidence computed using the behavioral logistic regression for each rat (see Fig. 1d); colors as in panel d. **f)** Differential behavioral kernel for location evidence across all rats.

Extended Data Figure 4. Stability of behavioral kernels. **a-b)** All trials collected for each rat were randomly split into two halves, and behavioral kernels were recomputed independently for each split half. **a)** Half split behavioral kernels for frequency evidence. **b)** Half split behavioral kernels for location evidence. **c)** The parallel index computed behavioral trials in the first half split is highly correlated with the parallel index computed using the second half split.
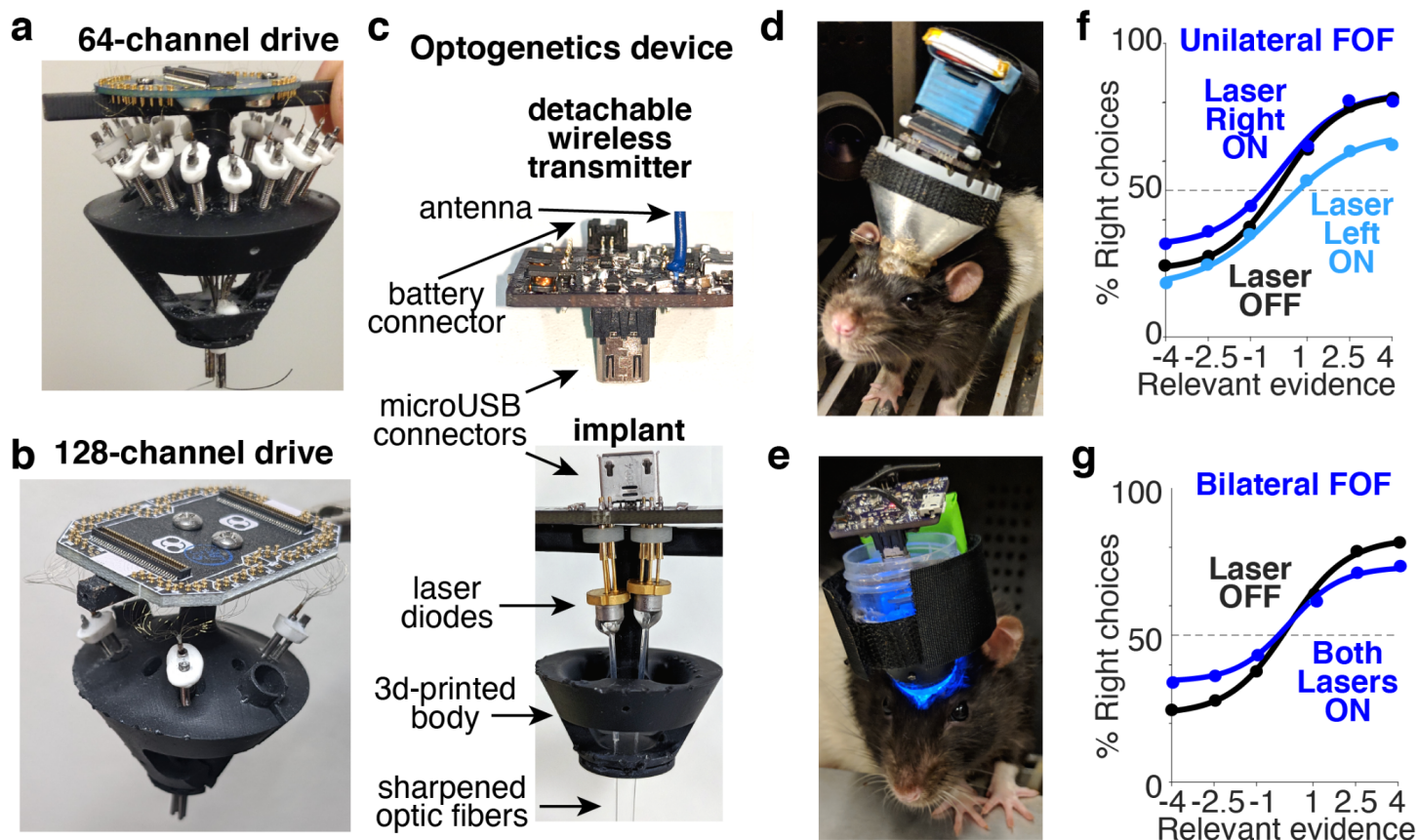
Extended Data Figure 5. **(a)** 64-channel custom-made multi-tetrode drive, allowing independent movement of 16 tetrodes. This drive was used in one rat for wired recordings. **(b)** 128-channel custom-made multi-tetrode drive, allowing independent movement of 4 bundles with 8 tetrodes each. This drive was used in six rats for wireless recordings. **(c)** Device for wireless optogenetic perturbation. In the implant, two chemically sharpened optic fibers targeting both hemispheres are attached using optical glue to two laser diodes. The laser diodes are controlled independently by a control board, which communicates wirelessly with the computer controlling the behavior. The control board can be attached/detached using a microUSB connector. **(d)** Example rat with wireless electrophysiology implant and headstage. **(e)** Example rat with wireless optogenetic implant and control board. **(f-g)** Result of inactivation of FOF. 3 rats expressed AAV2/5-mDlx-ChR2-mCherry and were stimulated with blue light (450 nm, 25mW) for the full duration of the stimulus. **(f)** Result of unilateral inactivation on rats' choices as a function of strength of relevant evidence (averaged across the two contexts). Activation of each laser was randomized across trials. **(g)** Result of bilateral FOF inactivation on rats' choices as a function of strength of relevant evidence (averaged across the two contexts).
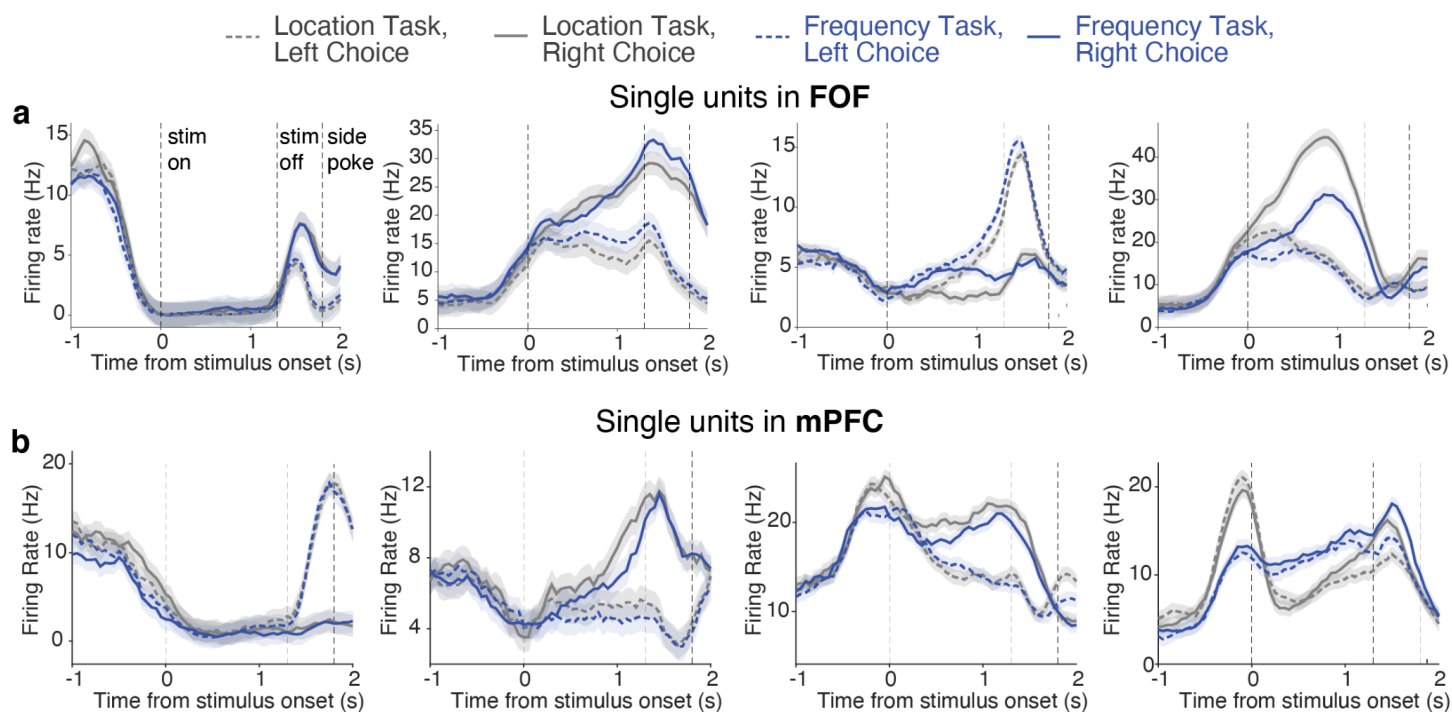
Extended Data Figure 6. Example responses of single units recorded in FOF **(a)** and in mPFC **(b)**. Shown are the peri-stimulus time histograms of responses for correct trials, averaged according to context and choice. Units in both areas exhibit significant heterogeneity and large modulation according to combinations of the rat's upcoming choice and the current context. The dashed vertical lines indicate the beginning of the pulse-train stimulus presentation, the end of the pulse-train stimulus presentation, and the average time when the rat performed a poke in one of the two side ports to indicate his choice.
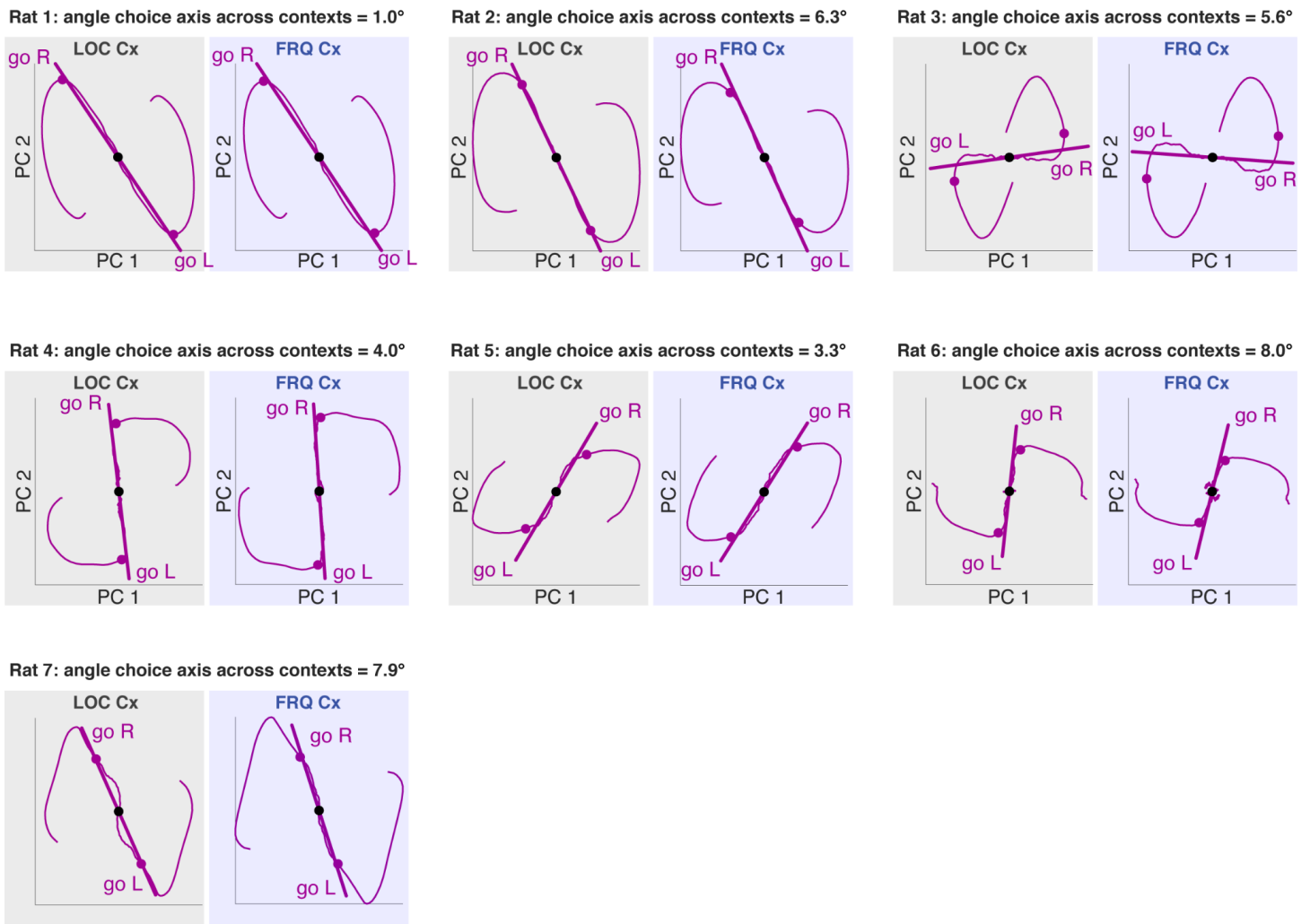
Extended Data Figure 7. Choice-related dynamics, computed independently for each rat, and across the two contexts. For each rat, the horizontal and vertical axes in the two subpanels are the same across the two panels, and are computed using data from both contexts. The dynamics in each context are computed using the choice kernels of the pulse-based regression (Fig. 4a). The black dot indicates the time of the start of stimulus presentation, the purple dots indicate the end of stimulus presentation. The line indicates the choice axis computed in the given context, and above the panels is indicated the angle between the choice axes computed across the two contexts.
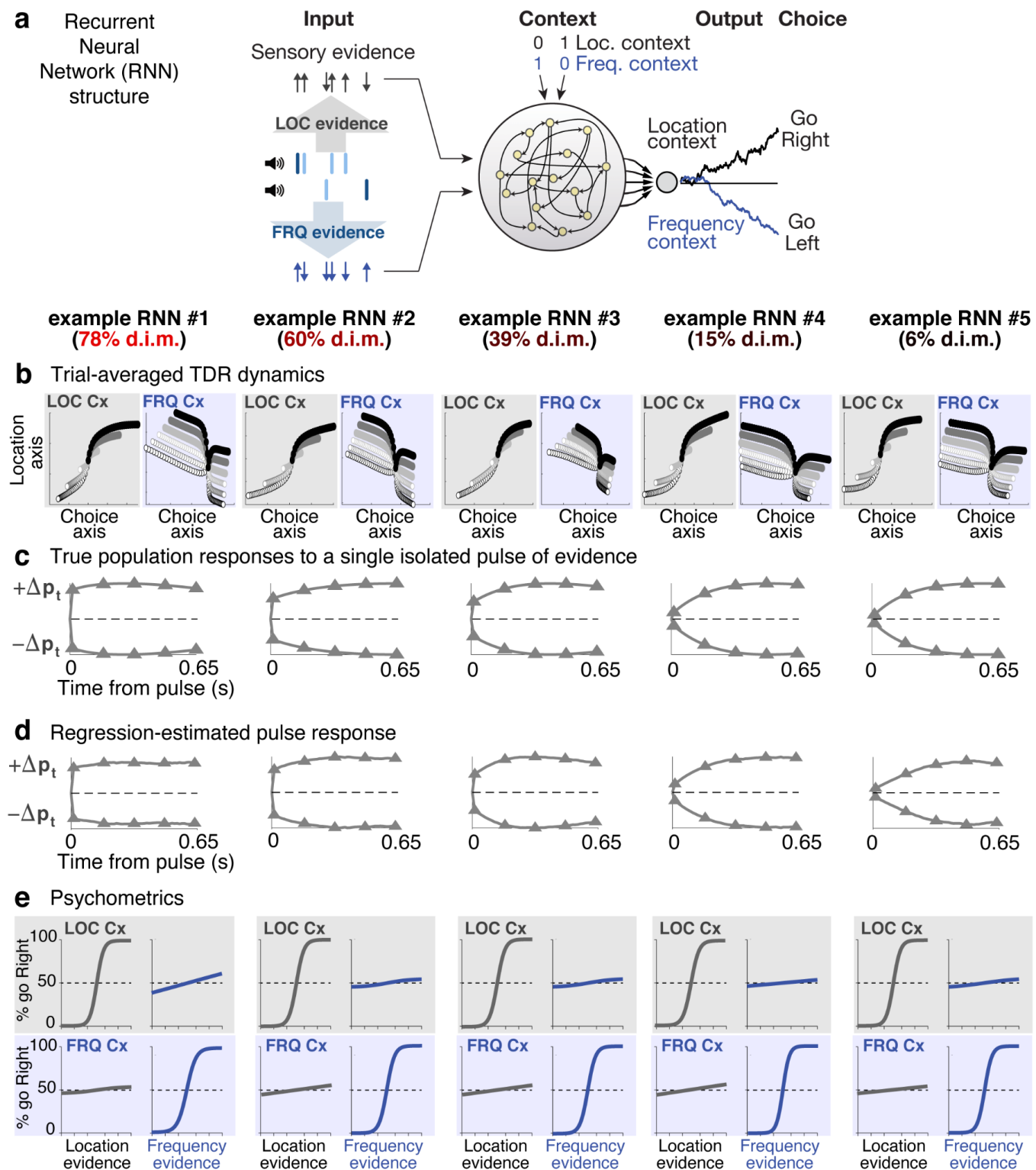
Extended Data Figure 8. Engineered recurrent neural networks (RNNs) across the entire vertical axis of the solution space (Fig. 3g) all qualitatively reproduce rat TDR trial-based dynamics and psychometric results but are distinguished by pulse-based analysis. **a)** Architecture of the RNNs. **b-f)** Activity and behavior of five example RNNs, engineered with different percentages of direct input modulation (d.i.m.; same notation as in Fig. 4; each column in b-f corresponds to an RNN with a given d.i.m. percentage). **b)** TDR trial-based dynamics (as in Fig. 2c-d) are similar across the full d.i.m. range. **c)** Differential network response to a single isolated pulse of location evidence across the two contexts ("true" single-pulse responses). As predicted by the theory, networks with larger d.i.m. components display larger initial differential responses. **d)** Estimation of the differential network response using the pulse-based regression method of Fig. 4. The pulse-based regression accurately captures the true pulse responses of panel b. **e)** Psychometric curves (Fig. 1c) show uniformly good performance across the d.i.m. range.
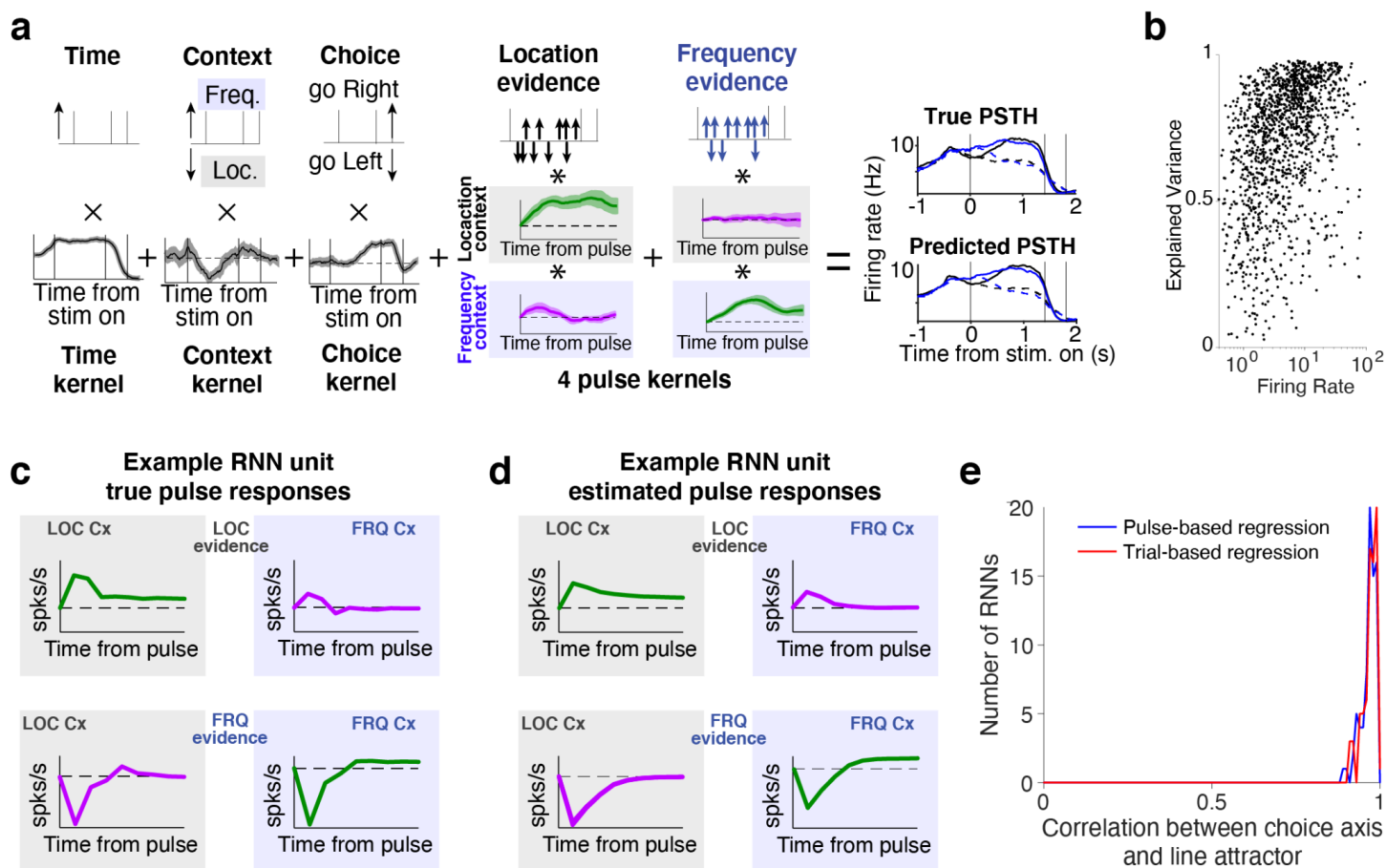
Extended Data Figure 9. Validation of pulse regression method. **a)** Example application of the pulse regression to one example recorded unit. **b)** Fraction of explained variance as a function of firing rate across all recorded units. **c-d)** The pulse-regression kernels provide an accurate estimate of the response to a single isolated pulse. In **c)** are shown the responses to a single isolated pulse of either location or frequency evidence in both contexts for an example RNN unit. In **d)** are shown the estimates of these pulses from the dynamics of the RNN solving the task with regular trials featuring many consecutive pulses presented at 40Hz. **e)** Comparison of the direction of the true line attractor (computed by finding the RNN's fixed points, see methods) with the choice axis estimated by the trial-based regression (Fig. 2c,d) and the pulse-based regression (Fig. 4a). The choice axis closely approximates the direction of the true line attractor.
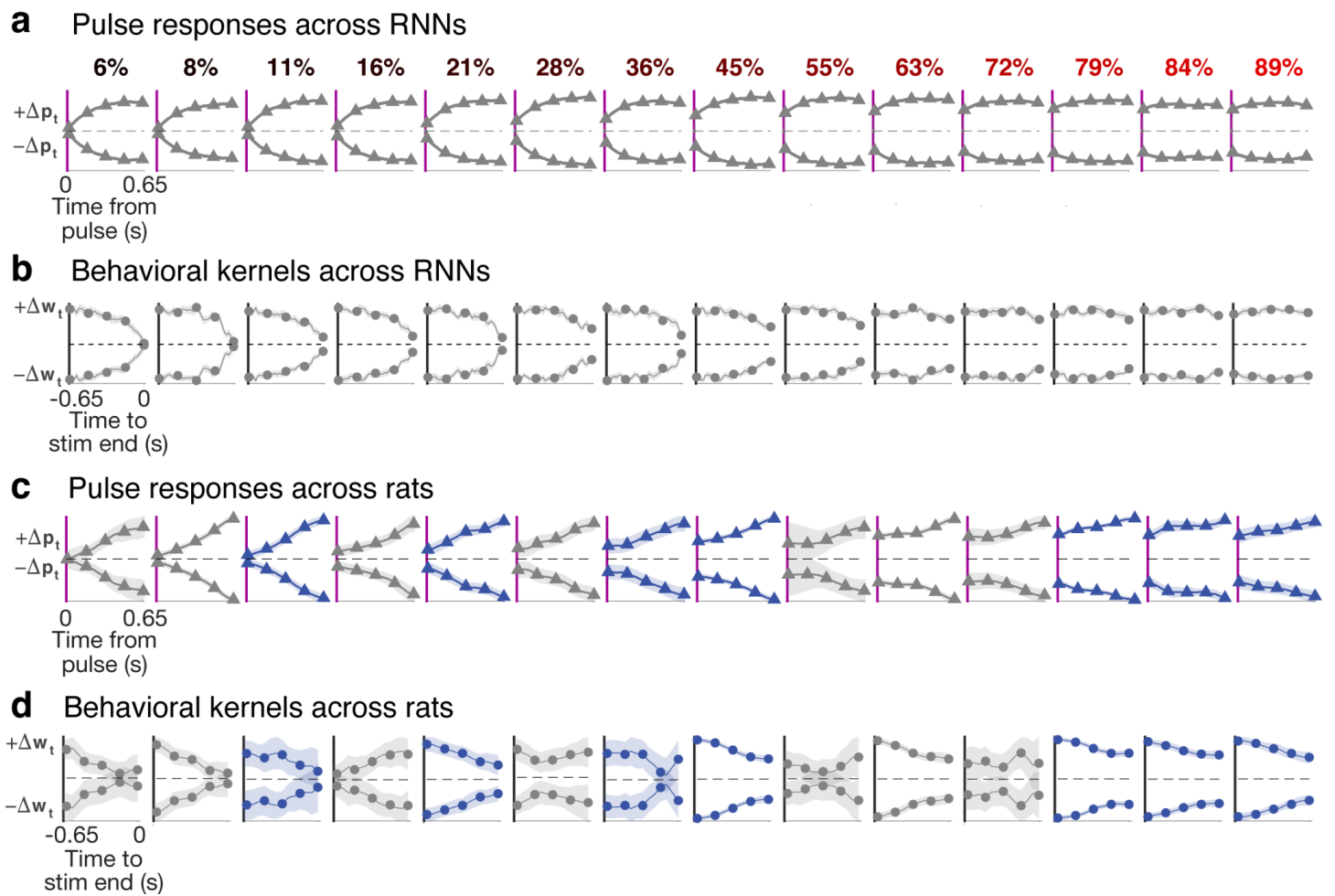
Extended Data Figure 10. **(a)** Differential pulse responses across all RNNs shown in Fig. 5c (n=30). The number above each behavioral kernel indicates the fraction of direct input modulation for the associated RNN (same notation as in Extended Data Figure 8). **(b)** Corresponding behavioral kernel for each RNN. **(c)** Differential pulse responses across all rats shown in Fig. 5d (n=7 rats, two features per rat). Gray indicates location feature, blue indicates frequency feature. **(d)** Corresponding behavioral kernels for each rat and feature.

Methods

Subjects

All animal use procedures were approved by the Princeton University Institutional Animal Care and Use Committee (IACUC) and were carried out in accordance with NIH standards. All subjects were adult male Long-Evans rats that were kept on a reversed light-dark cycle. All training and testing procedures were performed during the dark cycle. Rats were placed on a restricted water schedule to motivate them to work for a water reward. A total of 26 rats were used for the experiments presented in this study. Of these, 7 rats were used for electrophysiology recordings, and 3 rats were implanted with optical fibers for optogenetic inactivation.

Behavior

All rats included in this study were trained to perform a task requiring context-dependent selection and accumulation of sensory evidence (Figure 1a). The task was performed in a behavioral box consisting of three straight walls and one curved wall with three "nose ports". Each nose port was equipped with an LED to deliver visual stimuli, and with an infrared beam to detect the rat's nose entrance. In addition, above the two side ports were speakers to deliver sound stimuli, and water cannulas to deliver a water reward.

At the beginning of each trial, rats were presented with an audiovisual cue indicating the context of the current trial, either "location context" or "frequency context". The context cues consisted of 1s-long, clearly distinguishable FM modulated sounds, and in addition the "location context" was signaled by turning on the LEDs of all three ports, while in the "frequency context" only the center LED was turned on. After the end of the context cue, the rats were required to place their nose into the center port. While maintaining fixation in the center port, rats were presented with a 1.3s-long train of randomly-timed auditory pulses. Each pulse was played either from the speaker to the animal's left or from the speaker to their right, and each pulse a 5 ms pure tone with either low-frequency (6.5 KHz) or high frequency (14 KHz). The pulse trains were generated by Poisson processes with different underlying rates. The strength of the location evidence was manipulated by varying the relative rate of right vs left pulses, while the strength of the frequency evidence was manipulated by varying the relative rate of high vs low pulses (Fig. 1b). The overall pulse rate was kept constant at 40 Hz.

In the "location context", rats were rewarded if they turned, at the end of the stimulus, towards the side that had played the greater total number of pulses, ignoring the frequency of the pulses. In blocks of "frequency" trials, rats were rewarded for orienting left if the total number of low frequency pulses was higher than the total number of high frequency pulses, and orienting right otherwise, ignoring the location of the pulses. The context was kept constant in blocks of trials, and block switches occurred after a minimum of 30 trials per block, and when a local estimate of performance reached a threshold of 80% correct. Behavioral sessions lasted 2-4 hours, and rats performed on average 542 trials per session.

Electrophysiology

Tetrodes were constructed using nickel/chrome alloy wire, 12.7 μm (Sandvik Kanthal), and were gold-plated to 200 kΩ at 1 kHz. Tetrodes were mounted onto custom-made drives (Ext. Data Fig. 5a)[41], and the microdrives were implanted using previously described surgical stereotaxic implantation techniques[18]. Five rats were implanted with bilateral electrodes targeting FOF, centered at +2 anteroposterior (AP), ±1.3 mediolateral (ML) from bregma, while two rats were implanted with bilateral electrodes targeting the prelimbic (PL) area of mPFC, with coordinates +3.2 anteroposterior (AP), ±0.75 mediolateral (ML) from bregma. In one rat with an implant in FOF, 16 tetrodes were connected to a 64-channel electronic interface board (EIB), and recordings

were performed using a wired setup (Open-Ephys). In the other six rats, 32 tetrodes per rat were connected to a 128-channel EIB and recordings were performed using wireless headstages (Spikegadgets; Ext. Data Fig. 5b).

Optogenetics

Preparation of chemically-sharpened optical fibers (0.37 NA, 400 μm core; Newport) and basic virus injection techniques were the same as previously described[18]. At the targeted coordinates (FOF, +2 AP mm, ±1.3 ML mm from bregma), injections of 9.2nl of adeno-associated virus (AAV) (AAV2/5-mDlx-ChR2-mCherry, three rats) were made every 100 μm in depth for 1.5mm. Four additional injection tracts were completed at coordinates 500 μm anterior, 500 μm posterior, 500 μm medial and 500 μm lateral from the central tract. In total, 1.5μl of virus was injected over approximately 30min. Chemically sharpened fibers were lowered down the central injection track. Virus expression was allowed to develop for 8 weeks before optogenetic stimulation began. Optogenetic stimulation was delivered at 25mW using a customized wireless system derived from the "Cerebro" system (https://karpova-lab.github.io/cerebro ; Ext. Data Fig. 5c,d)[42,43].

Analysis of behavior

Data was extracted from all behavioral sessions in which rats' fraction of correct responses was equal or above 70%, feature selection index (see below) was equal or above 0.7, and in which rats performed at least 100 trials. Analysis of behavior was performed for all rats with electrophysiology or optogenetics implants, as well as for all other rats that performed at least 120,000 valid trials, i.e. where the rat maintained fixation for the full duration of the pulse train before making a decision. Psychometric curves (Fig. 1c; Extended Data Figure 3) were used to display the fraction of rightward choices as a function of the difference between the total number of right pulses and left pulses (location evidence strength), and as a function of the difference between the total number of high pulses and low pulses (frequency evidence strength). These curves were fit to a 4-parameter logistic function[5]:

$$y(x) = y_0 + \frac{a}{1+exp\left(\frac{-(x-x_0)}{b}\right)} \tag{3}$$

To quantify whether a rat selected the contextually relevant evidence to form his decisions on a given session, we computed a "feature selection index". For this purpose, we performed a logistic regression for each of the two contexts, where the rat's choices were fit as a function of the strength of location and frequency evidence. For each context, we considered all valid trials, and we compiled the rat's choices, as well as the strength of location and frequency evidence. The vector of choices was parameterized as a binary vector (Right = 1; Left = 0), the strength of location evidence was computed as the natural logarithm of the ratio between the rate of right and the rate of left pulses, while the strength of frequency evidence was computed as the natural logarithm of the ratio between the rate of high-frequency and the rate of low-frequency pulses. In the location context, we fit the probability of choosing right on trial k using the logistic regression:

$$logit\left(P(right)_k\right) = s_{LOC\ EVD,\ k}^{LOC\ CTX} \cdot w_{LOC\ EVD}^{LOC\ CTX} + s_{FRQ\ EVD,\ k}^{LOC\ CTX} \cdot w_{FRQ\ EVD}^{LOC\ CTX} + \beta^{LOC\ CTX} \tag{4}$$

where $s_{LOC\ EVD,\ k}^{LOC\ CTX}$ indicates the strength of location evidence on trial k, $s_{FRQ\ EVD,\ k}^{LOC\ CTX}$ indicates the strength of frequency evidence on trial k, $w_{LOC\ EVD}^{LOC\ CTX}$ is the weight of location evidence on the rat's choices, $w_{FRQ\ EVD}^{LOC\ CTX}$ is the

weight of frequency evidence on the rat's choices, and $\beta^{LOC\ CTX}$ is a bias term. The relative weight of location evidence in the location context was computed as:

$$\text{relative weight location} = \frac{w_{LOC\ EVD}^{LOC\ CTX}}{w_{LOC\ EVD}^{LOC\ CTX}+w_{FRQ\ EVD}^{LOC\ CTX}} \tag{5}$$

Similarly, in the frequency context we fit the rat's choices as:

$$logit\left(P(right)_k\right) = s_{LOC\ EVD,\ k}^{FRQ\ CTX} \cdot w_{LOC\ EVD}^{FRQ\ CTX} + s_{FRQ\ EVD,\ k}^{FRQ\ CTX} \cdot w_{FRQ\ EVD}^{FRQ\ CTX} + \beta^{FRQ\ CTX} \tag{6}$$

where $s_{LOC\ EVD,\ k}^{FRQ\ CTX}$ indicates the strength of location evidence on trial k, $s_{FRQ\ EVD,\ k}^{FRQ\ CTX}$ indicates the strength of frequency evidence on trial k, $w_{LOC\ EVD}^{FRQ\ CTX}$ is the weight of location evidence on the rat's choices, $w_{FRQ\ EVD}^{FRQ\ CTX}$ is the weight of frequency evidence on the rat's choices, and $\beta^{FRQ\ CTX}$ is a bias term. The relative weight of frequency evidence in the frequency context was computed as:

$$\text{relative weight frequency} = \frac{w_{FRQ\ EVD}^{FRQ\ CTX}}{w_{LOC\ EVD}^{FRQ\ CTX}+w_{FRQ\ EVD}^{FRQ\ CTX}} \tag{7}$$

Finally, the feature selection index was then computed as the average between the relative weight of location in the location context (Eq. 5), and the relative weight of frequency in the frequency context (Eq. 7):

$$\text{feature selection index} = \frac{1}{2} \cdot \frac{w_{LOC\ EVD}^{LOC\ CTX}}{w_{LOC\ EVD}^{LOC\ CTX}+w_{FRQ\ EVD}^{LOC\ CTX}} + \frac{w_{LOC\ EVD}^{FRQ\ CTX}}{w_{LOC\ EVD}^{FRQ\ CTX}+w_{FRQ\ EVD}^{FRQ\ CTX}} \tag{8}$$

The feature selection index was used to precisely quantify the rats' learning during training, as this metric allows to compare data across stages with different evidence strength (Extended Data Fig. 2). In addition, the relative weight of location and frequency were computed for each rat as a function of the position of a trial within the block (e.g. immediately after a block switch, one trial after a block switch etc.), providing a measure of the rats' ability to rapidly switch attended feature upon context switching (Extended Data Fig. 1e).

Behavioral logistic regression

To quantify the dynamics of evidence accumulation, behavioral data was analyzed using another logistic regression. Importantly, in Eq. 5 and 7 we quantified the rat's weighting of evidence using a single number, because we considered the generative rates, i.e. the expected strength of location and frequency evidence on a given trial. Now, we seek instead to quantify how these weights vary throughout stimulus presentation, by taking advantage of the knowledge of the exact pulse timing. For each rat, data across all sessions was compiled into a single vector of choices (Right vs Left), and two matrices detailing the pulse information presented on every trial. More specifically, the choice vector was parameterized as a binary vector (Right = 1; Left = 0), with dimensionality N, where N is the total number of valid trials. Pulse information was split into location evidence and frequency evidence, and was binned into 26 bins with 50 ms width. For a given bin, the amount of location evidence was computed as the natural logarithm of the ratio between the number of right and the number of left pulses, and was compiled in a location pulse matrix X^L with dimensionality N x 26. Similarly, frequency evidence was computed as the logarithm of the ration between high-frequency and low-frequency pulses, and was compiled into a frequency pulse matrix X^F with dimensionality N x 26. To

quantify the impact on choices of evidence presented at different time points we fit a logistic regression, where the probability of choosing right at trial k was given by:

$$logit\left(P(right)_k\right) = \sum_{t=1}^{26} X_{k,t}^L \cdot w_t^L + X_{k,t}^F \cdot w_t^F + \beta \tag{9}$$

where $X_{k,t}^L$ indicates the location evidence at time t on trial k, $X_{k,t}^F$ indicates the frequency evidence at time t on trial k, w_t^L indicates the location weight at time t, w_t^F indicates the frequency weight at time t, and β indicates the bias to one particular side. Weights were fit using ridge regression, and the ridge regularizer was chosen to optimally predict cross-validated choices. The regression was applied separately for trials in the location context, and trials in the frequency context, resulting in four sets of weights computed for each rat (Fig. 1d).

To study how evidence was differentially integrated across the two contexts, we then computed a differential behavioral kernel. The location differential kernel was equal to the difference between the location weights computed in the location context, and the location weights computed in the frequency context. Similarly, the frequency differential kernel was equal to the difference between the frequency weights computed across the two contexts (Fig. 1e).

To quantify the shape of the differential behavioral kernels, we computed a behavioral parallel index. This was defined as the ratio between the minimum difference between the weights across the two contexts, over the maximum difference, computed across all time points. As a result, a parallel index = 1 indicates that the difference between the two sets of weights is constant at all time points (i.e. the maximum difference is equal to the minimum difference), while a parallel index = 0 indicates that the two sets of weights fully converge for some time point. Note that the parallel index does not specify the direction of convergence, although empirically we found that differential behavioral kernels only displayed convergence towards the end of the pulse stimulus presentation (Fig. 1f, Extended Data Fig. 3).

Analysis of neural data

Spike sorting was performed using MountainSort[44], followed by manual curation of the results. 3280 putative single units were recorded from 5 rats in FOF, while 252 units were recorded from 2 rats in mPFC. To measure the responses of individual neurons, peri-stimulus time histograms (PSTH) were computed by binning spikes in 20 ms intervals, and averaging responses for trials according to choice and context. Responses of single neurons in both areas were highly heterogeneous and multiplexed multiple types of information (Extended Data Fig. 6), and no systematic difference was found in the encoding of task variables across the two regions (see e.g. Extended Data Fig. 7), so all studies of neural activity were carried out at the level of neural populations, and pooling data from FOF and from mPFC.

Trial-based targeted dimensionality reduction (TDR) analysis of neural population dynamics

To study trial-averaged population dynamics, we applied model-based targeted dimensionality reduction (mTDR)[16], a dimensionality-reduction method which seeks to identify the dimensions of population activity that carry information about different task variables. This method was applied to our rat dataset, and to reanalyze a dataset collected while macaque monkeys performed a similar visual task (Extended Data Fig. 1)[2]. In brief, the goal of mTDR is to identify the parameters of a model where the activity of each neuron is described as a linear combination of different task variables (choice, time, context, stimulus strength). For each of these task variables, the model retrieves a weight vector specifying how that variable influences neural activity at each time, and the collection of these weight vectors across all neurons are constrained to form a low-rank matrix. Singular Value Decomposition of this low-rank weight matrix is then used to identify basis vectors that

maximally encode each of the task variables. Using this method, we identified one axis maximally encoding information about the upcoming choice of the animal ("choice axis"), one axis maximally encoding information about the momentary strength of the first stimulus feature (location for rat data, motion for monkey data), and one axis maximally encoding information about the momentary strength of the second stimulus feature (frequency for rat data, color for monkey data). To study how neural dynamics evolved in this reduced space, we first averaged the activity of each neuron across all correct trials according to the strength of location evidence, strength of frequency evidence, context, and choice. For this analysis, spike counts were computed in 50 ms non-overlapping bins with centers starting at the beginning of the pulse train presentation and ending 50 ms after the end of the pulse train presentation. For any given trial condition, a "pseudo-population" (i.e. including non-simultaneously recorded neurons) was computed for each time point by compiling the responses of all neurons into a single vector. The trajectory of this vector over time was then projected on the retrieved task-relevant axes to evaluate population dynamics (Fig. 2).

Pulse-based TDR analysis of neural population dynamics

To estimate the impact of evidence pulses and other task variables on neural responses, we fit the activity of each recorded unit using a pulse-based linear regression (Fig. 4a). The critical difference between our previous trial-based application of TDR and the current pulse-based analysis is that in the trial-based analysis, neural responses are described as a function of a single number, the expected strength of location and frequency evidence over the entirety of a trial, and the analysis ignores the precise timing of pulses. In contrast, the pulse-based analysis leverages knowledge of the precise timing of evidence presentation, a feature made possible by the pulse-based nature of our task. For each neuron, spike counts were computed in 20-ms non-overlapping bins with centers starting 1 second before the beginning of the pulse train presentation, and ending 700 ms after the end of the stimulus presentation. The activity of neuron i at time t on trial k was described as:

$$r_{i,t}(k) = \beta_{choice;\,i,t} \cdot choice\,(k) + \beta_{context;\,i,t} \cdot context(k) + \beta_{time;\,i,t} + \beta_{LOC,LOC;\,i} * pulses_{LOC,LOC}(k) +$$

$$+ \beta_{LOC,FRQ;\,i} * pulses_{LOC,FRQ}(k) + \beta_{FRQ,LOC;\,i} * pulses_{FRQ,LOC}(k) + \beta_{FRQ,FRQ;\,i} * pulses_{FRQ,FRQ}(k) \quad (10)$$

where $x_{choice}(k)$ indicate the rat's choice on trial k (Right = 1, Left = 0), $x_{context}(k)$ indicate the context on trial k (Location = 1, Frequency = 0), $pulses_{LOC,LOC}(k)$ indicates the signed location evidence (number of right pulses minus number of left pulses) presented at each time bin on trial k in the location context, $pulses_{LOC,FRQ}(k)$ indicates location evidence in the frequency context, $pulses_{FRQ,LOC}(k)$ indicates frequency evidence (number of high pulses minus number of low pulses) in the location context, and $pulses_{FRQ,FRQ}(k)$ indicates frequency evidence in the frequency context. The first three regression coefficients $\beta_{choice;\,i}$, $\beta_{context;\,i}$ and $\beta_{time;\,i}$ account for modulations of neuron i across time according to choice, context and time. The other four sets of regression coefficients $\beta_{LOC,LOC;\,i}$, $\beta_{LOC,FRQ;\,i}$, $\beta_{FRQ,\,LOC;\,i}$ and $\beta_{FRQ,FRQ;\,i}$ indicate the impact of a pulse on the subsequent neural activity, and the symbol * indicates a convolution of each kernel with the pulse train; for example, in the case of location evidence in the location context:

$$\beta_{LOC,LOC;\,i} * pulses_{LOC,LOC}(k) = \sum \beta_{LOC,LOC;\,i,\tau} \cdot pulses_{LOC,LOC}(k;\,t - \tau) \quad (11)$$

meaning that the element at position τ of kernel $\beta_{LOC,LOC;\,i}$ represents the impact of a pulse of location evidence in the location context on the activity of unit i after a time τ. The three kernels for choice, context and time describe modulations from 1 second before stimulus start to 0.7s after stimulus end in 20-ms non-overlapping bins, resulting in 151-dimensional vectors. The four pulse kernels describe modulations from the time of pulse presentation to 0.65s after pulse presentation resulting in 33-dimensional vectors. To avoid overfitting, this regression was regularized using a ridge regularizer, as well as an L2 smoothing prior[45].

Pulse kernels were regarded as an approximation of the neural response to each pulse type (an assumption confirmed by analysis of recurrent neural networks, Ext. Data Fig. 9c,d), and pulse-evoked population responses were computed by compiling pulse kernels across all N neurons recorded from the same rat (Fig. 4b). We then studied the evolution of the projection p_t of this N-dimensional pulse-evoked population response onto a single "choice axis". To compute the choice axis, we evaluated the dynamics of choice-related activity across all neurons (Ext. Data Fig. 7), and we computed the first principal component of the matrix obtained by compiling the choice kernels across all neurons, limited to a time window during the presentation of the pulse train stimulus (0 to 1.3s after stimulus start).

To study the differential evolution of pulse-evoked population responses across the two contexts, we computed a "differential pulse response". For location evidence, the differential pulse response was defined as the difference between the projection onto the choice axis of the response to location pulses in the location context, and the response to location pulses in the frequency context. For frequency evidence, the differential pulse response was computed as the difference between the projection onto the choice axis of the frequency pulse response in the frequency context, minus the frequency pulse response in the location context (Fig. 4c).

To quantify the shape of differential pulse responses, we computed a neural parallel index. This was defined as the ratio between the minimum difference between the pulse responses across the two contexts, over the maximum difference, computed across all time points. As a result, a parallel index = 1 indicates that the difference between the two pulse responses is constant at all time points (i.e. the maximum difference is equal to the minimum difference), while a parallel index = 0 indicates that the two pulse responses fully converge for some time point. Note that the parallel index does not specify the direction of convergence, although empirically we found that differential pulse responses only displayed a shape that diverged over time, i.e. further amplifying the effect of relevant over irrelevant evidence onto the choice axis (Fig. 4e, Extended Data Fig. 10).

Recurrent neural networks (RNNs)

To validate our analyses of behavior and neural dynamics, and to gather a deeper understanding of the mathematical mechanisms that could underlie our rats' context-dependent behavior, we trained Recurrent Neural Networks (RNNs) to perform a pulse-based context-dependent evidence accumulation task analogous to that performed by the rats.

The activity of the N=100 hidden units of each network (Ext. Data Fig. 8a) was defined by the equations:

$$\tau \cdot dx/dt = -x + w_R \cdot r + w_u^{LOC} \cdot u^{LOC} + w_u^{FRQ} \cdot u^{FRQ} + w_C \cdot c + k \tag{12}$$

$$r = tanh(x) \tag{13}$$

where x(t) is a N-dimensional vector indicating the activation of each unit in the network, which are passed through a tanh nonlinearity to obtain the activity vector r(t); w_R indicates the N x N matrix of recurrent weights; u^{LOC} is the location input vector indicating at each time point the amount of location evidence (right minus left

pulses), u^{FRQ} is the frequency input vector indicating at each time point the amount of frequency evidence (high minus low pulses), w_u^{LOC} indicates the 1 x N matrix of weights applied to the location input, w_u^{FRQ} indicates the 1 x N matrix of weights applied to the frequency input, c is a 2-dimensional vector with a one-hot encoding of the current context, w_c indicates the 2 x N matrix of weights applied to the context, and k is a scalar indicating a bias term. In the location context, the first element of c is 1, and the second element is 0; in the frequency context, the first element of c is 0, and the second element is 1. The output of the network was determined by a single output unit performing a linear readout of the activity of the hidden units:

$$ z = w_O \cdot r + k_O \tag{14} $$

where w_O indicates the N x 1 vector of weights assigned to each hidden unit, and k_O is a scalar representing the output bias. The choice of the network on a given trial was determined by the sign of z at the last time point (T = 1.3s). During training and analysis, evolution of the network was computed in 10 ms time bins. During training, the time constant τ was set to 10ms, but in subsequent analyses this value was varied to replicate the autocorrelation observed in neural data to τ = 100ms.

Training of RNNs using backpropagation

Recurrent neural networks were trained using back-propagation-through-time with the Adam optimizer and implemented in the Python JAX framework. The weights of the network were initialized using a standard normal distribution, modified according to the number of inputs to a unit, and then rescaled. If η is drawn from a standard normal distribution $\eta \sim N(0, 1)$, input weights were chosen as $\eta \cdot \frac{1}{\sqrt{U}}$; recurrent weights were chosen as $\eta \cdot \frac{0.8}{\sqrt{N}}$; output weights were chosen as $\eta \cdot \frac{1}{\sqrt{N}}$; where U indicates the number of inputs (U=4), and N indicates the number of hidden units (N=100). All the biases of the network were initialized at 0. The initial conditions were also learned, and were also initialized randomly from a standard normal distribution, with each element of the initial condition initialized as $0.1 \cdot \eta$. The Adam parameters for training were: b1=0.9; b2=0.999; epsilon=0.1 . The learning rate followed an exponential decay with initial step size = 0.002, and decay factor = 0.99998. Training occurred over 120,000 batches with a batch size of 256 trials. Using this procedure, we trained 1000 distinct RNNs to solve the task using different random initializations on each run (Fig. 3h). All networks learned to perform the task with high accuracy (see e.g. Ext. Data Fig. 8e).
All the code for training, analysis, and engineering of RNNs will be made available before the time of publication at: https://github.com/Brody-Lab/flexible_decision_making_rnn.

Analysis of RNN mechanisms

To analyze the mechanism implemented by each RNN to perform context-dependent evidence accumulation, we first identified the fixed points of each trained network using a previously described optimization procedure[2,46].
While the network dynamics are in general described by a nonlinear function F (equations 12, 13), around fixed points these dynamics can be approximated as a linear system:

$$ dr/dt = F\left(r, u^{LOC}, u^{FRQ}\right) \approx \frac{\partial F}{\partial r} \cdot r + \frac{\partial F}{\partial u^{LOC}} \cdot u^{LOC} + \frac{\partial F}{\partial u^{FRQ}} \cdot u^{FRQ} = M \cdot r + i^{LOC} \cdot u^{LOC} + i^{FRQ} \cdot u^{FRQ} \tag{15} $$

The jacobian matrix M indicates the state-transition matrix of the linear dynamical system, capturing the partial derivative of each unit's activity with respect to change to any other unit's activity. The "effective input" i is defined as the partial derivative of each unit's activity with respect to changes to the input, and it can be computed independently for pulses of location evidence (i^{LOC}), or for pulses of frequency evidence (i^{FRQ}).

In our analysis for simplicity we will only focus on the effect of pulses of location evidence (the same considerations hold for frequency evidence), so we will drop the superscript and simply write "i " to indicate the "effective input" for pulses of location evidence. We will instead use a subscript to indicate whether this effective input for location evidence is computed in the location context (i_{LOC}) or in the frequency context (i_{FRQ}). Likewise, we will drop the superscript from the input weights and simply use w_u to indicate the input weights for location evidence.

For each trained RNN, we focused on the analysis of the linearized dynamics corresponding to the fixed point with the smallest absolute network output (i.e. where the network is closest to the decision boundary), but results were similar when considering different fixed points (i.e. linearized dynamics were mostly similar across different fixed points). Similar to previous reports, we found that in every network fixed points were roughly aligned to form a "line attractor" for each of the two contexts, and that eigendecomposition of the jacobian matrix M reveals a single eigenvalue close to 0, and all other eigenvalues negative, reflecting the existence of a single stable direction of evidence accumulation (i.e. the line attractor) surrounded by stable dynamics.

The right eigenvector associated with the eigenvalue closest to 0 defined the direction of the line attractor ρ, while the corresponding left eigenvector defined the direction of the selection vector **s**. For each network, we computed these quantities separately for the two contexts, i.e. by setting the contextual input c as (1,0) for the location context, or as (0,1) for the frequency context before computing the fixed points and the eigendecomposition. As a result, for each network we computed the line attractor in each of the two contexts, which we name ρ_{LOC} and ρ_{FRQ} , the selection vector in each of the two contexts (s_{LOC} and s_{FRQ}), and the effective input in each of the two contexts (i_{LOC} and i_{FRQ}). Using these quantities, we directly computed the terms in equation 2 to quantify how much each of the three components contributed to differential pulse accumulation, and we plotted the results for 1000 RNNs in barycentric coordinates (Fig. 3h).

Engineering of RNNs to implement arbitrary combinations of components

To engineer recurrent neural networks that would implement arbitrary combinations of components, we started from the RNN solutions obtained from standard training using backpropagation-through-time. For a given trained network, we first computed the fixed points of the network and the linearized network dynamics, and we identified the line attractor, selection vector and effective input across the two contexts (see above).

Because the RNN dynamics are known (Equations 12 and 13), the linearized dynamics can be expressed in closed form as a function of the network weights:

$$M^j = \frac{\partial F^j}{\partial r} = w_R^j \odot tanh'(w_R \cdot r_{fixed} + w_C \cdot c + k) \tag{16}$$

$$i = \frac{\partial F}{\partial u} = w_u \odot tanh'(w_R \cdot r_{fixed} + w_C \cdot c + k) \tag{17}$$

where M^j indicates the j-th column of the jacobian matrix, w_R^j indicates the j-th column of the matrix of recurrent weights, r_{fixed} indicates the network activity at the fixed point, $tanh'$ indicates the first derivative of the hyperbolic tangent nonlinearity, and $\odot$ indicates the Hadamard product or element-wise multiplication, where

the elements of two vectors are multiplied element-by-element to produce a vector of the same size. We further define the "saturation factor" for each of the two contexts as:

$$sat_{LOC} = tanh'(w_R \cdot r_{fixed, LOC} + w_C \cdot c_{LOC} + k) \tag{18}$$

$$sat_{FRQ} = tanh'(w_R \cdot r_{fixed, FRQ} + w_C \cdot c_{FRQ} + k) \tag{19}$$

where $r_{fixed, LOC}$ indicates the fixed point with the smallest absolute network output in the location context, $r_{fixed, FRQ}$ indicates the fixed point with the smallest absolute network output in the frequency context, c_{LOC} indicates the context input in the location context (1,0), and c_{FRQ} indicates the context input in the frequency context (0,1). The effective input for the two contexts can therefore be computed as:

$$i_{LOC} = w_u \odot sat_{LOC} \quad ; \quad i_{FRQ} = w_u \odot sat_{FRQ} \tag{20}$$

The three components of context-dependent differential integration defined in Equation 2 can therefore be rewritten as a function of the input weights w_u.

Selection vector modulation, which is equal to the dot product between the difference in the selection vector and the average effective input, can be rewritten as:

$$\Delta s \cdot \bar{i} = \Delta s \cdot (i_{LOC} + i_{FRQ})/2 = \Delta s \cdot w_u \odot (sat_{LOC} + sat_{FRQ})/2 = \Delta s \cdot w_u \odot \overline{sat} = w_u \cdot (\Delta s \odot \overline{sat}) \tag{21}$$

where $\overline{sat}$ indicates the average saturation factor across contexts, and the last step took advantage of the associative property of the Hadamard and dot product.

Direct input modulation, which is equal to the dot product between the difference in the effective input and the line attractor, can be rewritten as:

$$\Delta i \cdot \bar{\rho} = (i_{LOC} - i_{FRQ}) \cdot \bar{\rho} = w_u \odot (sat_{LOC} - sat_{FRQ}) \cdot \bar{\rho} = w_u \odot \Delta sat \cdot \bar{\rho} = w_u \cdot (\Delta sat \odot \bar{\rho}) \tag{22}$$

where Δsat indicates the difference between the saturation factor across the two contexts.

Indirect input modulation, which is equal to the dot product between the difference in the effective input and the average selection vector orthogonal to the line attractor $\bar{s}_{\perp}$, can be rewritten as:

$$\Delta i \cdot \bar{s}_{\perp} = (i_{LOC} - i_{FRQ}) \cdot \bar{s}_{\perp} = w_u \odot (sat_{LOC} - sat_{FRQ}) \cdot \bar{s}_{\perp} = w_u \odot \Delta sat \cdot \bar{s}_{\perp} = w_u \cdot (\Delta sat \odot \bar{s}_{\perp}) \tag{23}$$

Knowledge of equations 21, 22 and 23 allow us to identify input vectors that produce network dynamics relying on any arbitrary combinations of the three components. For example, producing a network using exclusively selection vector modulation requires the first component (Eq. 21) to be large , while the second (Eq. 22) and third (Eq. 23) components must be 0. In other words, the input weights w_u must satisfy:

$$w_u \cdot (\Delta s \odot \overline{sat}) > 0 \quad ; \quad w_u \cdot (\Delta sat \odot \bar{\rho}) = 0 \quad ; \quad w_u \cdot (\Delta sat \odot \bar{s}_{\perp}) = 0 \tag{24}$$

In addition, we must also require that the network does not accumulate the pulse in the irrelevant context. Because we are conducting this analysis for pulses of location evidence, this means that the dot product between the effective input and the selection vector in the frequency context should be 0 :

$$i_{FRQ} \cdot s_{FRQ} = 0 \qquad \Rightarrow \qquad i_{FRQ} \cdot s_{FRQ} = w_U \odot sat_{FRQ} \cdot s_{FRQ} = w_U \cdot \left(sat_{FRQ} \odot s_{FRQ} \right) = 0 \qquad (25)$$

Finally, we then use the Gram-Schmidt process to find the set of weight w_u maximally aligned to the vector $\Delta s \odot \overline{sat}$, and orthogonal to vectors $\Delta sat \odot \overline{\rho}$, $\Delta sat \odot \overline{s_\perp}$ and $sat_{FRQ} \odot s_{FRQ}$. Similar considerations can be applied to produce networks using different mechanisms. For example, to engineer a network that uses only direct input modulation the input weight must be maximally aligned to $\Delta sat \odot \overline{\rho}$ and orthogonal to $\Delta s \odot \overline{sat}$, $\Delta sat \odot \overline{s_\perp}$ and $sat_{FRQ} \odot s_{FRQ}$. Engineering networks implementing combinations of mechanisms can be obtained by choosing the input vector as a linear combination between extreme network solutions. Finally, we emphasize that the mechanism chosen for one stimulus feature (e.g. location) is entirely independent from the mechanism chosen for the other stimulus feature (e.g. frequency).

Statistical methods

Comparison of the strength of the encoding of relevant vs irrelevant information (Fig. 2c,d) was performed by quantifying the variability across responses to different stimulus strengths, normalized by trial-by-trial variability, limiting the analysis to the subspace orthogonal to choice encoding. Error bars for neural and behavioral kernels were computed using bootstrapping[47].

References

1. Gold, J. I. & Shadlen, M. N. The neural basis of decision making. *Annu. Rev. Neurosci.* **30**, 535–574 (2007).

2. Mante, V., Sussillo, D., Shenoy, K. V. & Newsome, W. T. Context-dependent computation by recurrent dynamics in prefrontal cortex. *Nature* **503**, 78–84 (2013).

3. Brody, C. D. & Hanks, T. D. Neural underpinnings of the evidence accumulator. *Curr. Opin. Neurobiol.* (2016) doi:10.1016/j.conb.2016.01.003.

4. Okazawa, G. & Kiani, R. Neural Mechanisms that Make Perceptual Decisions Flexible. *Annu. Rev. Physiol.* (2022) doi:10.1146/annurev-physiol-031722-024731.

5. Brunton, B. W., Botvinick, M. M. & Brody, C. D. Rats and humans can optimally accumulate evidence for decision-making. *Science* **340**, 95–98 (2013).

6. Keung, W., Hagen, T. A. & Wilson, R. C. Regulation of evidence accumulation by pupil-linked arousal processes. *Nat Hum Behav* **3**, 636–645 (2019).

7. Keung, W., Hagen, T. A. & Wilson, R. C. A divisive model of evidence accumulation explains uneven weighting of evidence over time. *Nat. Commun.* **11**, 2160 (2020).

8. Reynolds, J. H. & Chelazzi, L. Attentional modulation of visual processing. *Annu. Rev. Neurosci.* **27**, 611–647 (2004).

9. Noudoost, B., Chang, M. H., Steinmetz, N. A. & Moore, T. Top-down control of visual attention. *Curr. Opin. Neurobiol.* **20**, 183–190 (2010).

10. Maunsell, J. H. R. & Treue, S. Feature-based attention in visual cortex. *Trends Neurosci.* **29**, 317–322 (2006).

11. Wimmer, R. D. *et al.* Thalamic control of sensory selection in divided attention. *Nature* **526**, 705–709 (2015).

12. Barbosa, J., Proville, R., Rodgers, C. C., Ostojic, S. & Boubenec, Y. Flexible selection of task-relevant features through across-area population gating. *bioRxiv* 2022.07.21.500962 (2022) doi:10.1101/2022.07.21.500962.

13. Ding, L. & Gold, J. I. Neural correlates of perceptual decision making before, during, and after decision commitment in monkey frontal eye field. *Cereb. Cortex* **22**, 1052–1067 (2012).

14. Gold, J. I. & Shadlen, M. N. Representation of a perceptual decision in developing oculomotor commands. *Nature* **404**, 390–394 (2000).

15. Siegel, M., Buschman, T. J. & Miller, E. K. Cortical information flow during flexible sensorimotor decisions. *Science* **348**, 1352–1355 (2015).

16. Aoi, M. C., Mante, V. & Pillow, J. W. Prefrontal cortex exhibits multidimensional dynamic encoding during decision-making. *Nat. Neurosci.* **23**, 1410–1420 (2020).

17. Erlich, J. C., Bialek, M. & Brody, C. D. A cortical substrate for memory-guided orienting in the rat. *Neuron* **72**, 330–343 (2011).

18. Hanks, T. D. *et al.* Distinct relationships of parietal and prefrontal cortices to evidence accumulation. *Nature* **520**, 220–223 (2015).

19. Leonard, C. M. The prefrontal cortex of the rat. I. Cortical projection of the mediodorsal nucleus. II. Efferent connections. *Brain Res.* **12**, 321–343 (1969).

20. Sinnamon, H. M. & Galer, B. S. Head movements elicited by electrical stimulation of the anteromedial cortex of the rat. *Physiol. Behav.* **33**, 185–190 (1984).

21. Okazawa, G., Hatch, C. E., Mancoo, A., Machens, C. K. & Kiani, R. Representational geometry of perceptual decisions in the monkey parietal cortex. *Cell* **184**, 3748–3761.e18 (2021).

22. Seung, H. S. How the brain keeps the eyes still. *Proc. Natl. Acad. Sci. U. S. A.* **93**, 13339–13344 (1996).

23. Vyas, S., Golub, M. D., Sussillo, D. & Shenoy, K. V. Computation Through Neural Population Dynamics. *Annu. Rev. Neurosci.* **43**, 249–275 (2020).

24. Dubreuil, A., Valente, A., Beiran, M., Mastrogiuseppe, F. & Ostojic, S. The role of population structure in computations through neural dynamics. *Nat. Neurosci.* **25**, 783–794 (2022).

25. Langdon, C. & Engel, T. A. Latent circuit inference from heterogeneous neural responses during cognitive tasks. *bioRxiv* 2022.01.23.477431 (2022) doi:10.1101/2022.01.23.477431.

26. Takagi, Y., Hunt, L. T., Woolrich, M. W., Behrens, T. E. & Klein-Flügge, M. C. Adapting non-invasive human

recordings along multiple task-axes shows unfolding of spontaneous and over-trained choice. *Elife* **10**, (2021).

27. Flesch, T., Juechems, K., Dumbalska, T., Saxe, A. & Summerfield, C. Orthogonal representations for robust context-dependent task performance in brains and neural networks. *Neuron* **110**, 1258–1270.e11 (2022).

28. Sasaki, R. & Uka, T. Dynamic readout of behaviorally relevant signals from area MT during task switching. *Neuron* **62**, 147–157 (2009).

29. Rodgers, C. C. & DeWeese, M. R. Neural correlates of task switching in prefrontal cortex and primary auditory cortex in a novel stimulus selection task for rodents. *Neuron* **82**, 1157–1170 (2014).

30. Duan, C. A. *et al.* Collicular circuits for flexible sensorimotor routing. *Nat. Neurosci.* **24**, 1110–1120 (2021).

31. Perich, M. G. & Rajan, K. Rethinking brain-wide interactions through multi-region 'network of networks' models. *Curr. Opin. Neurobiol.* **65**, 146–151 (2020).

32. Orhan, A. E. & Ma, W. J. Publisher Correction: A diverse range of factors affect the nature of neural representations underlying short-term memory. *Nat. Neurosci.* **22**, 505 (2019).

33. Wang, J., Narain, D., Hosseini, E. A. & Jazayeri, M. Flexible timing by temporal scaling of cortical responses. *Nat. Neurosci.* **21**, 102–110 (2018).

34. Sohn, H., Narain, D., Meirhaeghe, N. & Jazayeri, M. Bayesian Computation through Cortical Latent Dynamics. *Neuron* **103**, 934–947.e5 (2019).

35. Remington, E. D., Narain, D., Hosseini, E. A. & Jazayeri, M. Flexible Sensorimotor Computations through Rapid Reconfiguration of Cortical Dynamics. *Neuron* **98**, 1005–1019.e5 (2018).

36. Ritz, H. & Shenhav, A. Humans reconfigure target and distractor processing to address distinct task demands. *bioRxiv* 2021.09.08.459546 (2022) doi:10.1101/2021.09.08.459546.

37. Chadwick, A. *et al.* Learning shapes cortical dynamics to enhance integration of relevant sensory input. *bioRxiv* 2021.08.02.454726 (2021) doi:10.1101/2021.08.02.454726.

38. Pandarinath, C. *et al.* Inferring single-trial neural population dynamics using sequential auto-encoders. *Nature Methods* vol. 15 805–815 Preprint at https://doi.org/10.1038/s41592-018-0109-9 (2018).

39. Kim, T. D., Luo, T. Z. & Pillow, J. W. Inferring latent dynamics underlying neural population activity via neural differential equations. *Conference on Machine …* (2021).

40. Murphy, B. K. & Miller, K. D. Balanced amplification: a new mechanism of selective amplification of neural activity patterns. *Neuron* **61**, 635–648 (2009).

41. Aronov, D. & Tank, D. W. Engagement of neural circuits underlying 2D spatial navigation in a rodent virtual reality system. *Neuron* **84**, 442–456 (2014).

42. Tervo, D. G. R. *et al.* The anterior cingulate cortex directs exploration of alternative strategies. *Neuron* **109**, 1876–1887.e6 (2021).

43. Brown, J. *et al.* Expanding the Optogenetics Toolkit by Topological Inversion of Rhodopsins. *Cell* **175**, 1131–1140.e11 (2018).

44. Chung, J. E. *et al.* A Fully Automated Approach to Spike Sorting. *Neuron* **95**, 1381–1394.e6 (2017).

45. Pillow, J. W. *et al.* Spatio-temporal correlations and visual signalling in a complete neuronal population. *Nature* **454**, 995–999 (2008).

46. Sussillo, D. & Barak, O. Opening the black box: low-dimensional dynamics in high-dimensional recurrent neural networks. *Neural Comput.* **25**, 626–649 (2013).

47. Efron, B. & Tibshirani, R. J. *An Introduction to the Bootstrap*. (CRC Press, 1994).

48. Druckmann, S. & Chklovskii, D. B. Neuronal circuits underlying persistent representations despite time varying activity. *Curr. Biol.* **22**, 2095–2103 (2012).