



**Supplementary Material for**  
**Spiking neurons can discover predictive features by aggregate-label learning**

Robert Gütig

E-mail: [guetig@em.mpg.de](mailto:guetig@em.mpg.de)

Published 4 March 2016, *Science* **351**, aab4113 (2016)  
DOI: 10.1126/science.aab4113

**This PDF file includes:**

Materials and Methods  
Supplementary Text S1 and S2  
Figs. S1 to S9  
Tables S1 and S2  
Full Reference List

**Text S1: Continuous tuning curves** Many sensory clues, such as the spectral locations of peaks within a communication call or the color of a strawberry, are not confined to points in sensory space but instead extend along one or several continuous dimensions. When measured along such dimensions, responses of sensory neurons often exhibit smoothly varying tuning curves. At the same time, also the feedback that follows a sensory clue, such as the urgency of the communication call or the sugar content of the strawberry is typically a continuous quantity. Can continuous aggregate teaching signals subserve neural learning of continuous tuning curves such that the graded neural response to a particular clue realization reflects its predicted value for the animal? We addressed this question by replacing the discrete set of activity patterns that represented clues and distractors in the previous scenarios with a continuous manifold of activity patterns (Fig. S4A, Methods: Continuous feature dimensions). As a result, each feature occurrence was now parametrized by a continuous value  $\alpha$  (sampled randomly), and the aggregate label  $\ell$  given by the sum of all “feedback contributions”  $f(\alpha)$  that occurred within a trial, i.e.

$$\ell = \sum_i f(\alpha_i). \quad (41)$$

This scalar feedback was sufficient for multi-spike tempotrons to reliably learn a broad class of generic tuning curves  $f$  (Methods: Continuous tuning curves) as well as the arbitrarily selected gray level profiles of Leonardo da Vinci’s Mona Lisa (Figs. S4B and C). The deterministic neuron model learned to utilize the random background activity in which features were embedded to approximate the continuous valued desired output spike numbers. The fidelity of the learned neural tuning remained high even for target tuning functions that were substantially more complex than the typically reported uni-modal tuning curves of neurons in primary sensory areas (Fig. S4D). Furthermore, multi-spike tempotrons were able to learn to decode a broad range of neural encoding schemes of the clue parameter, including spike-latency (48, 49), spike-count (50, 51), and mean firing rate (52, 53) based representations (Methods: Encoding

schemes). Learning and processing exhibited considerable robustness to temporal spike jitter (Fig. S4E) or random deletion of presynaptic spikes (Fig. S4F).

**Text S2: Self-supervised neural networks** Self-supervised learning was insensitive to the exact functional form of the internally generated feedback signal: In addition to the mean, also the, non-linear, median and mode of the processing layer activity resulted in similar convergence times (Fig. S6G). Importantly, the supervisory signal could also be implemented on the basis of the spike count of a model supervisor neuron (Fig. S6G, Methods: Scalar teaching signals in self-supervised networks, Stability of self-supervised learning dynamics).

Convergence times and feature selection in self-supervised networks were strongly influenced by the total number of spikes within the Poissonian activity patterns that represented individual features (Fig. S6C). In individual input architectures these spike counts were independent across the processing layer neurons, such that the saliency of a given feature typically differed between individual neurons. This disagreement between the neurons' individual biases slowed down the convergence of the learning dynamics in individual input networks when multiple features were competing for a processing layer's responses (Figs. 4B inset and S6G). Indeed, this increase in learning time disappeared when the relative saliencies of all features agreed across the processing layer, i.e. when the activity patterns of all features were ordered (for each neuron) such that their spike count ranks were preserved from neuron to neuron (Figs. 4B inset and S6G). Self-supervised networks could reliably detect feature spike patterns as long as their number of spikes was not substantially lower (approximately 80%) than present in a corresponding interval of background activity (Fig. S6E). While spike counts are an intuitive, long established, determinant of feature salience, it has also been suggested that spike synchrony might serve as an alternative potent saliency cue for downstream neural processing layers (54, 55). For instance, a recent study has implicated rate and synchronicity as complementary codes for the

perceived brightness of visual features (56). Consistent with these findings, we found that convergence times of self-supervised networks were substantially accelerated even when only a moderate fraction of spikes within a feature’s activity pattern arrived in synchrony (Fig. S6E, inset). Finally, we found that the mechanism of self-supervision was fully operational even with as few as 3 processing layer neurons (Fig. S6D), however, larger processing layers showed faster convergence when feature representations were corrupted by noise (Fig. S6F).

Also in the more general scenario of continuous sensory features, self-supervised networks converged to a state in which all processing neurons responded with the same number of spikes to any feature (Fig. S7A). While rare feature occurrences within trials resulted in wide tuning curves of the processing layers, competition between more frequently occurring features led to populations of specialized neurons with narrower tuning (Fig. S7B). The dependence of the tuning width on the processing layer size was mostly confined to a regime of low feature frequencies where smaller networks converged to sharper tuning curves (Fig. S7C). Self-supervised learning dynamics tended to amplify initial biases within processing layer responses (Fig. S7D). However, consistent with experimental findings (39), initial biases could be overwritten by non-uniform feature statistics during training, that were matched by the emerging tuning curve statistics (Fig. S7E).

**Fig. S1** Multi-spike tempotron learning **A**, Raster plot (as in Fig. 1A) of an example input activity pattern and **B**, the corresponding voltage trace (blue curve) elicited in a postsynaptic integrate-and-fire neuron. For the neuron’s biological firing threshold at  $\vartheta = 1$ , the input activity elicits five output spikes (red vertical lines). **C**, The spike-threshold-surface (STS) (blue line), i.e. the number of output spikes (y-axis) as a function of the cell’s (hypothetical) firing threshold (x-axis). Corresponding to the 5 spikes shown in B, the cell’s biological threshold value (vertical gray line) intersects the STS at 5 spikes. The STS can be characterized by the positions of the steps in spike counts along the x-axis (dotted vertical lines). Each of these critical threshold values,  $\vartheta_k^*$ , is defined as the highest threshold value (supremum) that yields  $k$  output spikes. The highest threshold value that would result in a single output spike corresponds to the global maximum of the unresetted postsynaptic potential, i.e.  $\vartheta_1^* = V_{\max}$ . For a given input activity pattern, each  $\vartheta_k^*$ , and hence the entire STS, is a function of the neuron’s synaptic efficacies (inset). **D**, Example learning kernels of the multi-spike tempotron (blue and red lines). To change a neuron’s response, i.e. to modify the number of spikes at which the biological threshold crosses the STS, synaptic efficacies are adjusted such that the locations of specific  $\vartheta_k^*$  are shifted. For example, to increase or decrease the number of output spikes of the shown neuron to 6 or 4, respectively, the multi-spike tempotron learning rule would modify the neuron’s synaptic efficacies along the gradient that shifts  $\vartheta_6^*$  or  $\vartheta_5^*$  towards the biological threshold, respectively. The times  $t_6^*$  and  $t_5^*$  correspond to the times at which the voltage, i.e. Eq. 5 evaluated with the critical threshold values, reaches the corresponding critical thresholds. The learning kernels show the spike-timing dependence of a synapse’s contribution to the respective  $\vartheta_k^*$ , i.e.  $dV(t_k^*)/d\omega$ , as a function of its firing time within the given trial: In blue for a potentiating shift towards 6 output spikes ( $k = 6$ ) and in red for a depressing shift towards 4 output spikes ( $k = 5$ ). The shape of the learning kernels is dominated by the temporal profile of the PSP kernel (Methods: Neuron model) but incorporates corrections from preceding output

spikes.

**Fig. S2** Supervised learning of sensory clues. **A**, Spike-threshold-surfaces (solid lines, cf. Fig. S1C) of the example neurons underlying the top three voltage traces in Fig. 1B. Starting from the initial synaptic efficacies (blue histogram and blue STS) that gave rise to 6 output spikes (Fig. 1B, top trace) in response to the example activity pattern (Fig. 1A), learning modified the neuron’s efficacies (red histogram) such that the resulting STS (red) developed a plateau with three output spikes (Fig. 1B, second trace) at the intersection with the biological threshold value (vertical gray line). The elevation of the plateau results from the presence of three clues in the input activity. When the neuron was trained to fire five spikes in response to each clue occurrence (Fig. 1B, third trace) the resulting STS (black) intersected the neural spike threshold at  $3 \times 5 = 15$  output spikes. Dotted lines and shaded areas (same color code) show the mean and mean  $\pm 1$  s.d. intervals of the STSs for ensembles of 1000 independent simulation runs. Insets (same color code) show the efficacy distributions of the three individual examples (bar histograms) and the corresponding ensemble distributions (solid lines). **B**, Ensemble means (solid lines) and mean  $\pm 1$  s.d. intervals (shaded areas) of the neural responses for the five clues task shown in Figs. 1B (bottom trace) and C (right) over 1000 independent simulation runs. After the 1000th cycle the valencies of the individual clues were changed from one to four (mid-blue), two to three (light-blue), three to five (green), four to two (yellow) and five to one (red) spikes per clue occurrence. **C**, Mean convergence times (blue circles and red diamonds) and exponential fits (blue and red solid lines with slopes of 1.23 and 1.04, respectively) obtained with the “relative” (as shown in Fig. 2A) and “absolute” multi-spike tempotron learning rules (Methods: Multi-spike tempotron), respectively. Black  $\times$ -symbols denote mean convergence times for the relative learning rule without the default initialization of synaptic efficacies (Methods: Initialization). The inset shows the dependence of the mean convergence time on the

learning rate (x-axis) in units of the default parameter  $\lambda = 10^{-5}$  for mean trial durations of 5s, 20s and 50s (bottom to top). **D**, Random symmetry breaking between  $N_f = 10$  features with perfectly correlated occurrence numbers, i.e. the task of Fig. 2C with  $\rho = 1$ . Learned neural responses (gray scale) of 1000 independently simulated neurons (x-axis, sorted by learned clue preference). Each neuron randomly selected one of the features (y-axis) as its clue, which then blocked learning of any of the other features.

**Fig. S3** Comparison of the multi-spike tempotron (left) and the correlation-based learning rule (right) in supervised and unsupervised learning tasks. **A**, and **B**, show population data for the learning curves of Figs. 1C (same color codes) left and right, respectively. Solid lines depict the mean neural responses (shaded areas correspond to  $\pm 1$  s.d.) over 1000 independent simulations. In the extra panel (B, bottom right) correlation-based learning was stabilized by a divisive weight normalization (Methods: Probabilistic reward contingencies) that limited the norm of the synaptic weights to twice its initial value. **C**, Convergence times of the unsupervised feature detection tasks as in Fig. 4B, inset. Note that while correlation based learning was stable in A and C, the more difficult task in B required early stopping. While the shown transient state exhibits clear acquisition of the desired responses for all clues, responses to background activity did not decrease to zero, probably due to a lack in the approximate learning rule’s synaptic specificity. Continued learning caused a gradual increase of the responses to background activity and a subsequent degradation and loss of the desired feature responses (cf. Discussion). As shown in panel B (bottom right), this instability disappeared when the learning included a divisive normalization.

**Fig. S4** Supervised learning of continuous tuning curves. **A**, Neural encoding stage (Methods: Continuous feature dimensions): A circular continuous feature dimension (x-axis) is mapped to a manifold of activity patterns. The number (top) and timing (middle) of spikes within the

50 ms feature interval (y-axis) are shown for three example inputs (color coded) as a function of a continuous feature parameter  $\alpha$ . For each feature parameter, spike times linearly interpolate between their times (circles) in adjacent templates (vertical solid lines). If the number of spikes in two adjacent templates differs, excess spikes disappear at intermediate feature parameters depicted by transitions from solid to dotted lines that connect “existing” (filled circles) with “non-existing” (open circles) template spikes. (bottom) Mean number of output spikes as a function of the feature parameter after neurons were trained with a single feature parameter  $\alpha = \pi$  with an aggregate-label (Eq. 41) that was arbitrarily set to 8.7 (gray horizontal line) spikes per feature. The plot shows the decay of the neurons’ mean responses (average over 1000 simulations) with increasing distance from the training feature (solid lines), for different correlation lengths used in the encoding stage, i.e. the parameter  $\gamma$  (Methods: Continuous feature dimensions) set to 2 (blue), 5 (red), and 10 (black), respectively. Shaded regions depict intervals of  $\pm 1$  s.d.. **B**, Learned mean neural responses (open circles) for example target tuning curves (red lines): simple harmonic function (top), random periodic function with 5 non-zero Fourier modes (middle) and the gray values of one row (black arrow in c) of Leonardo da Vinci’s Mona Lisa. Responses shown in the top ( $R^2 = 0.998$ ) and middle ( $R^2 = 0.996$ ) panels realized the median  $R^2$  values over 1000 independent simulations. **C**, Mean responses (gray-level coded) of 100 neurons each trained with one row of the Mona Lisa as tuning function. **D-F**, Fidelity ( $R^2$  value) of learned mean responses in capturing random periodic tuning functions as a function of the number of non-zero Fourier modes (D), input spike jitter (E), or spike deletion probability (F). In E and F, target tuning curves are simple harmonic functions. Colors refer to different encoding schemes (Methods: Encoding schemes) by which the feature parameter is encoded: mean Poisson firing rates (green), spike latencies (red), spike counts (black), or both, spike latencies and counts (blue). All responses were measured after 3000 cycles of annealed training.



**Fig. S5** Supervised learning with probabilistic reward contingencies. **A**, Neural responses as a function of learning cycles in the presence (red) and absence (blue) of false rewards. Solid lines show the mean responses to a single clue with a reward contingency of  $\pi_R = 0.8$  averaged over 1000 independent simulations. False rewards (red) were caused by occurrences of an additional clue with the same reward contingency but whose occurrences were not visible to the neuron. Their main effect is a delay in the acquisition of the clue. In addition, false rewards caused a slight elevation of the neural responses, including those to background activity as shown in the inset in Hz. Responses to distractor features remained negligible in both scenarios (not shown). Shaded areas depict  $\pm 1$  s.d. All input spike patterns had a spike deletion probability of 0.25. **B**, Histogram of mean responses to three clues with reward contingencies and noise levels as in as in Fig. 2F, i.e.  $\pi_1 \approx 1$  (black),  $\pi_2 = 0.8$  (white), and  $\pi_3 = 0.6$  (gray), respectively. The figure includes all 820 out of 1000 independently simulated neurons whose mean responses exceeded one spike for all three clues. Mean responses were averaged over 1000 sequential cycles following  $10^5$  cycles of training. **C**, Scatter plot of individual mean responses. Each circle corresponds to one simulation and is color coded (color bar) according to the neurons mean response to the most reliable feature. The position of each circle denotes the relative response strengths to the second (x-axis) and third features (y-axis), respectively. All points lie below the diagonal (solid black line). The  $\times$ -symbol marks the simulation whose mean responses to all three features are closest to the population mean (Euclidean distance) and underlies the learning curves shown in Fig. 2F.

**Fig. S6** Self supervised neural networks learning discrete features. **A**, Population data corresponding to Fig. 4B. Mean processing layer responses as a function of learning time. Feature responses (left y-axis) were averaged within two groups depending on whether a feature's final processing layer response was greater (blue) or smaller (red) than 0.5 spikes per feature.

Background responses are shown in black (right y-axis). Solid lines depict averages over 1000 simulations and shaded areas correspond to  $\pm 1$  s.d. around the mean. Dotted lines depict the learning curves of the simulation shown in Fig. 4B. **B**, Time evolution of the mean correlation coefficients of the synaptic efficacies within the processing layer for shared input networks (blue) and independent input networks (red). Line styles and shaded areas as in panel A. **C**, Scatter plot of the spike count (y-axis) of the feature activity pattern versus the convergence time (in cycles) of shared input networks with 10 processing neurons when the number of features was one ( $N_f = 1$ ). Circles depict data from 999 out of 1000 converged simulations. For better visibility, convergence times (x-axis) are also represented by colors. The gray square represents a single simulation whose feature had only 82 spikes and did not converge within the simulation's run time limit of  $10^4$  cycles. For the same network architecture but with  $N_f = 10$ , the inset shows the histogram (over 1000 simulations) of the spike count ranks of all features that the processing layer responded to after convergence. **D-F**, Mean convergence times, CT, over 1000 simulations of shared (solid lines) and individual (dotted lines) input architectures as functions of the processing layer size (D), feature spike count (FSC) (E) and spike deletion probability  $p_{\text{del}}$  (F). Panel D compares convergence times of the single feature task ( $N_f = 1$ ) that resulted from a strict, single neuron based convergence criterion (blue) with a population based alternative (red) (Methods: Convergence in unsupervised tasks). Each feature activity pattern contained 125 spikes, i.e.  $\text{FSC} = 1$  (see below). In E, the convergence time is plotted as a function of the input patterns' feature spike count FSC. Networks have 10 processing neurons and the number of features is  $N_f = 1$  (blue and black) or  $N_f = 10$  (red and green). Feature spike counts are plotted in units of the average background spike count within a feature interval, i.e.  $rNT_f = 125$  spikes. The inset shows the acceleration in the mean convergence times for  $N_f = 1$  and  $\text{FSC} = 0.84$  as a function of the fraction of spikes within the feature activity pattern that are synchronous. In F, all features contain 125 spikes, i.e.  $\text{FSC} = 1$ , and the

number of processing layer neurons is 3 (red), 10 (blue) and 50 (black), respectively (Methods: Fixed feature spike counts, Input noise). Throughout this figure the background duration was  $T_0 = 1$  s. **G**, Mean convergence times, CT, over 1000 independent simulations obtained with different feedback signals corresponding to the mode, median, and mean (cf. Fig. 4B, inset) of the processing layer spike counts or the normalized output spike count of an integrate-and-fire supervisor neuron model. Dotted bars correspond to a single feature task, i.e.  $N_f = 1$  and solid bars to  $N_f = 10$ . Dark colors refer to shared input networks and light colors to individual input architectures. Open and filled bars for the  $N_f = 10$  feature task with the individual input architecture refer to random vs. spike-count rank ordered input features, respectively (see supplementary online text for details).

**Fig. S7** Self supervised neural networks learning continuous features. **A**, Example of the evolution of the mean processing layer responses (colorbar) to features from a continuous feature dimension. From a random initial state, a population tuning curve emerges with all processing layer neurons responding to input features within the same region of the continuous dimension. The mean feature occurrence count was  $c_f^{\text{mean}} = 5$ . **B**, Width of the resulting tuning curve (Methods: Neural tuning curves) as a function of the mean feature occurrence count  $c_f^{\text{mean}}$ . **C**, Mean tuning width (solid lines) as in B as a function of the processing layer size for mean feature counts of  $c_f^{\text{mean}} = 1$  (blue),  $c_f^{\text{mean}} = 2$  (red) and  $c_f^{\text{mean}} = 5$  (black). Shaded regions extend between the 10<sup>th</sup> and 90<sup>th</sup> percentile of 1000 simulations. **D**, (top) Time evolution of the correlations between the evolving and final population tuning curves after 200 cycles (black). Control correlations between shuffled independent simulations are shown in red. Solid lines depict means over 1000 simulations and shaded areas correspond to  $\pm 1$  s.d.. The time resolved tuning curve correlations from the simulation shown in A is represented by the black dotted line. To disentangle the effects of the specific realization of the activity patterns underlying

the feature dimension from those of the initial synaptic efficacies, the cyan dashed line depicts the mean of the shuffled correlations measured within 10 blocks of 100 simulations that used independent initial efficacies but identical feature dimensions within each of the blocks. On average the feature dimension contributed roughly half of the correlation between the initial and final tuning curves. (bottom) Time evolution of the spike count coherence (SCC) within the processing layer, i.e. the fraction of processing layer neuron spike counts within a learning cycle that corresponded to the mode of the processing layer activity on a given trial. **E**, Coverage of the continuous feature dimension for different feature parameter distributions (color coded) that were used during learning. For each feature parameter the curves show the fraction of 1000 simulations in which the feature parameter was inside the networks response region after convergence. Feature distributions (inset) were: uniform (blue) and van Mises distributions with dispersion parameters  $\kappa = 0.25$  (red),  $\kappa = 0.5$  (green) and  $\kappa = 1$  (black), respectively. The mean feature occurrence count was  $c_f^{\text{mean}} = 2$ . Throughout this figure networks had individual input architectures and the background duration was  $T_\emptyset = 1$  s.

**Fig. S8** Self supervised neural networks with lateral inhibition. **A**, Mean responses of 50 neurons (y-axis) organized in 5 equally sized self-supervised networks (gray horizontal lines) without (left) and with (right,  $\omega_I = -0.1$ ) lateral inhibition (Methods: Lateral inhibition). **B**, Tuning curve overlap as a function of the strength of the lateral inhibition (x-axis) between 5 self-supervised networks. Mean fractions of the continuous feature dimension for which the population responses of 1 (blue), 2 (green), 3 (yellow), 4 (orange) or 5 (red) of the groups exceeds 0.5 spikes. For IPSP amplitudes below approximately  $\omega_I < -0.075$  the overlap between individual groups practically vanished and any feature parameter was encoded by exactly one group. **C**, Tuning width histograms for 5 non-interacting groups (gray) versus tuning width histograms of 3 (green), 5 (red) and 7 (blue) groups with lateral inhibition strength of

$\omega_I = -0.1$ . The number of groups did not affect the almost perfect coverage fractions (B) of the groups' non-overlapping tuning curves: 0.9978 (3 groups), 0.998 (5 groups), and 0.9941 (7 groups). Instead, increases in the number of groups resulted in a more pronounced specialization of individual responses, i.e. narrower tuning curves. Measurements in B and C are based on 100 independent simulations of networks with individual input architectures,  $c_f^{\text{mean}} = 2$  and  $T_\emptyset = 1$  s.

**Fig. S9** Supervisor mediated neural interactions. **A**, Time evolution of processing layer responses of the simulations shown in Fig. 5B. The processing layer extended over 50 neurons that were supervised by 50 supervisory units with “random” (top) and circulant (bottom) connectivities (cf. Fig. 5B, insets). After an initial phase of 50 cycles (green background) the single inhomogeneous eigenvalue of the supervisory matrix was set to  $\lambda_{\text{inhom}} = 4$  for one critical cycle (red background) (Methods: Neural population tuning for discrete features). In the following consolidation phase (150 cycles) neural responses converged to the inhomogeneous eigenvector (Fig. 5B). The red dashed lines depict the mean processing layer response. **B**, Scatter plots of baseline and amplitude fit parameters between the learned neural responses and the inhomogeneous mode of the supervisory matrix (Methods: Supervisory matrices). Harmonic fits (bottom) had an additional phase parameter (resulting from the linear combination of the two first order harmonic eigenvectors of  $S$ ) that followed a uniform distribution (not shown). All fits had  $R^2$  values above 0.995. Red circles mark selected simulations whose  $R^2$  values were closest to the population means of 0.9979 and 0.9985, respectively, and underlie the data shown in (A) and Fig. 5B. **C**, Population statistics (Methods: Continuous feature maps) of the emerged continuous maps (Fig. 6B): Phases (left), slopes (middle) and  $R^2$  values (right). Red vertical lines denote the values of the simulation underlying Fig. 6A and B (left) whose  $R^2$  value was closest to the population mean.



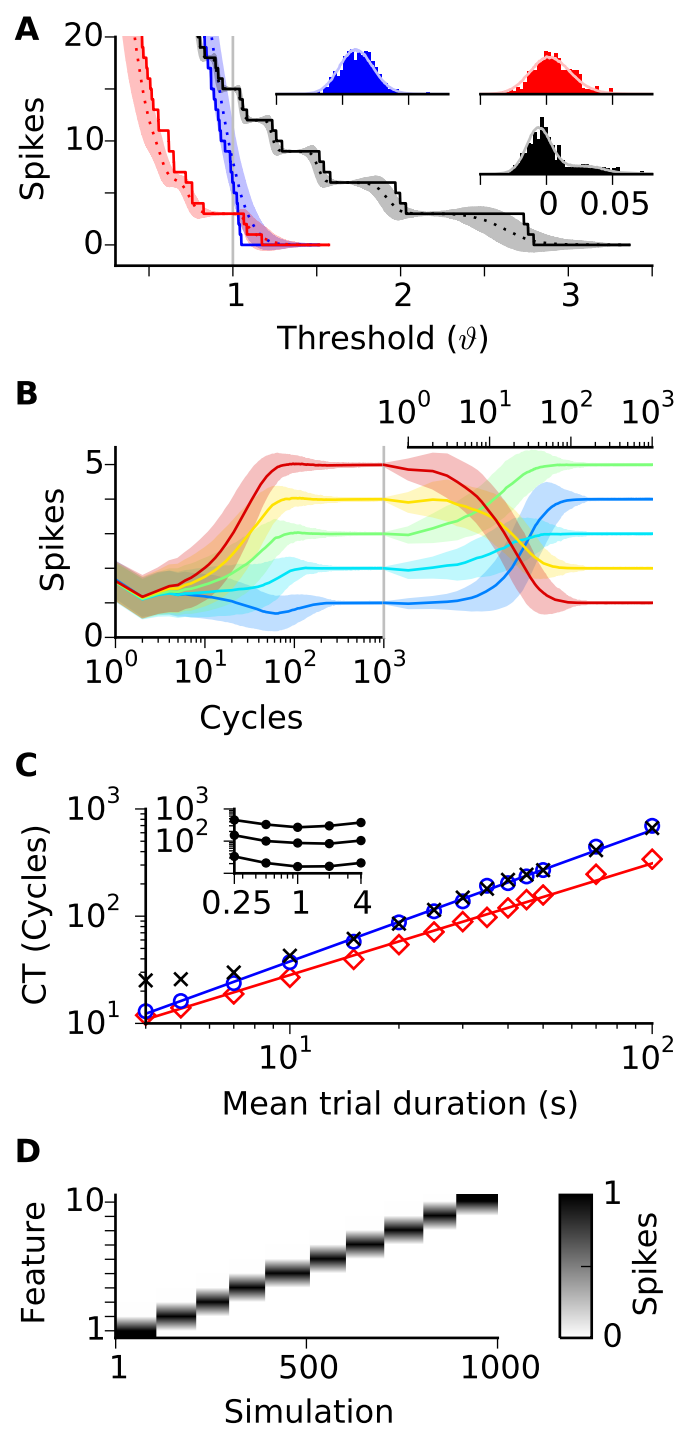


Fig. S2

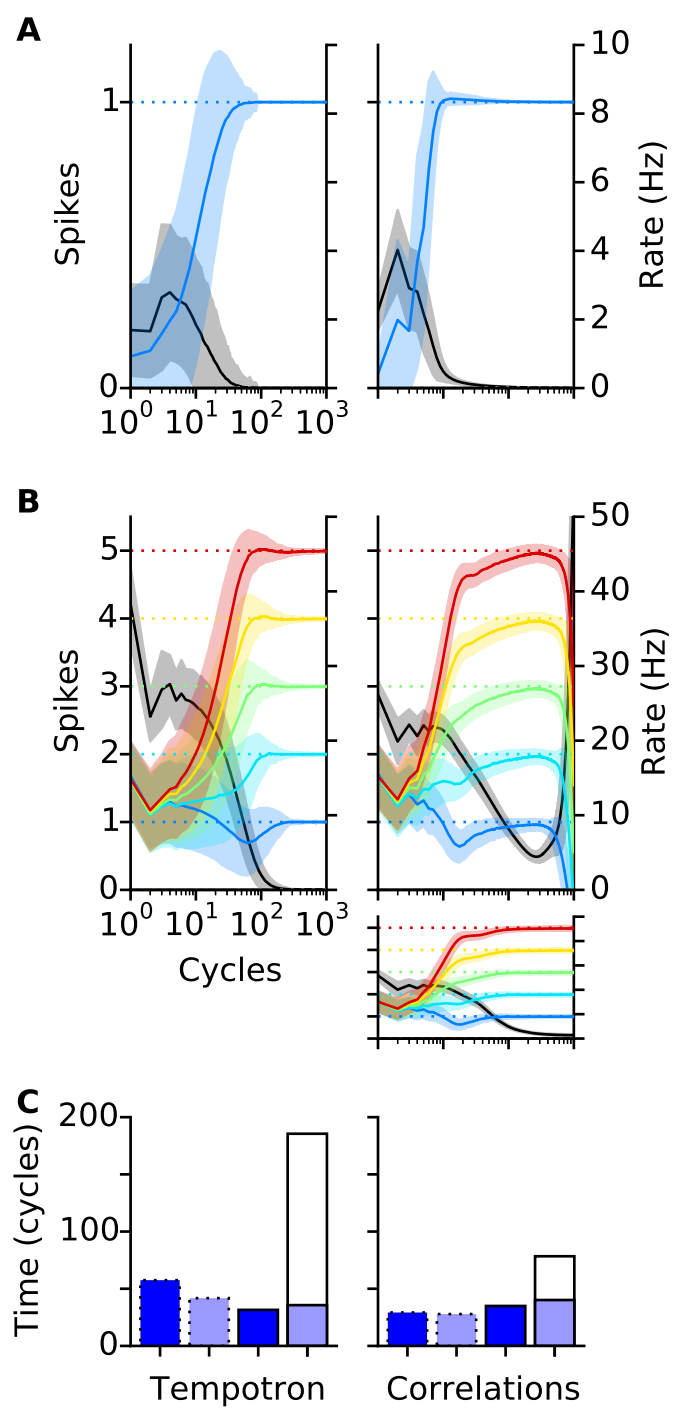


Fig. S3



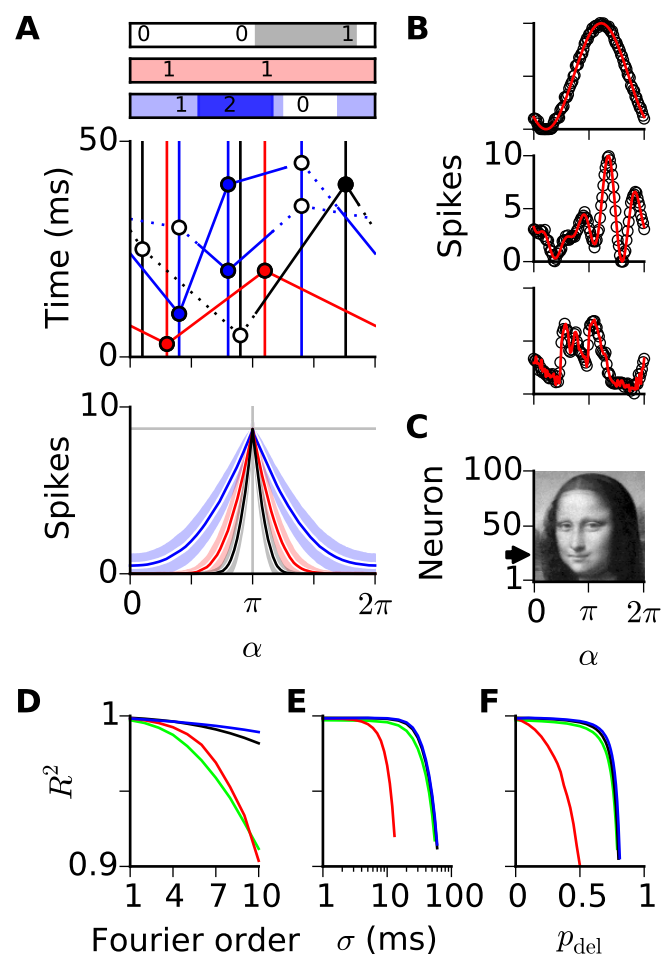


Fig. S4

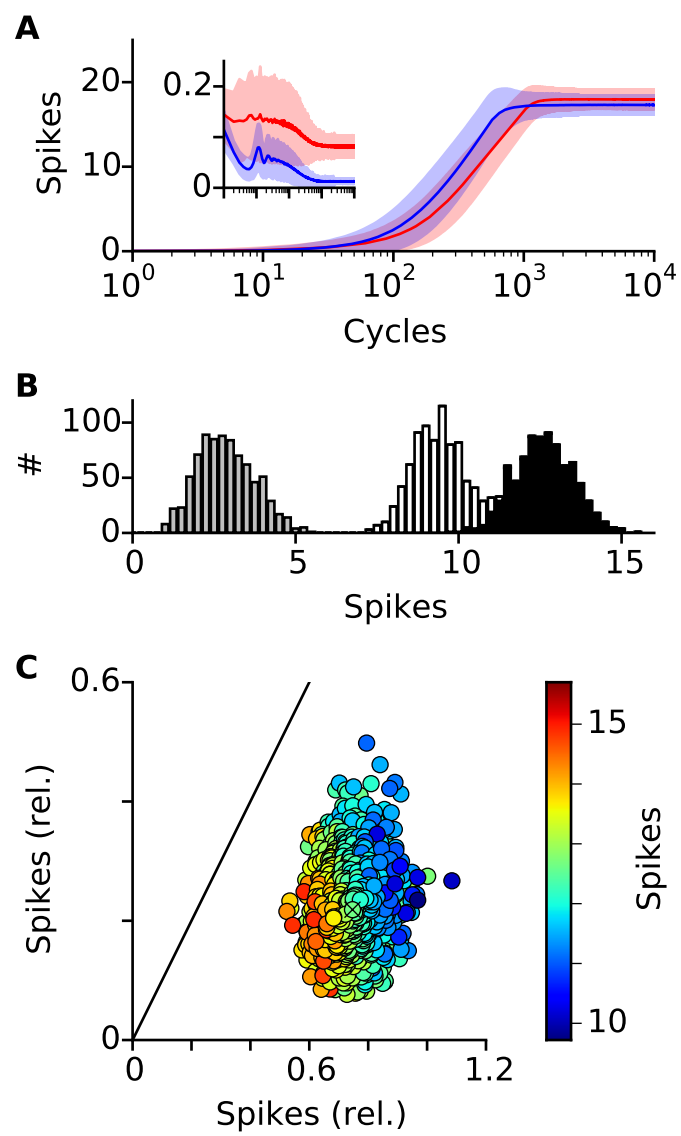


Fig. S5

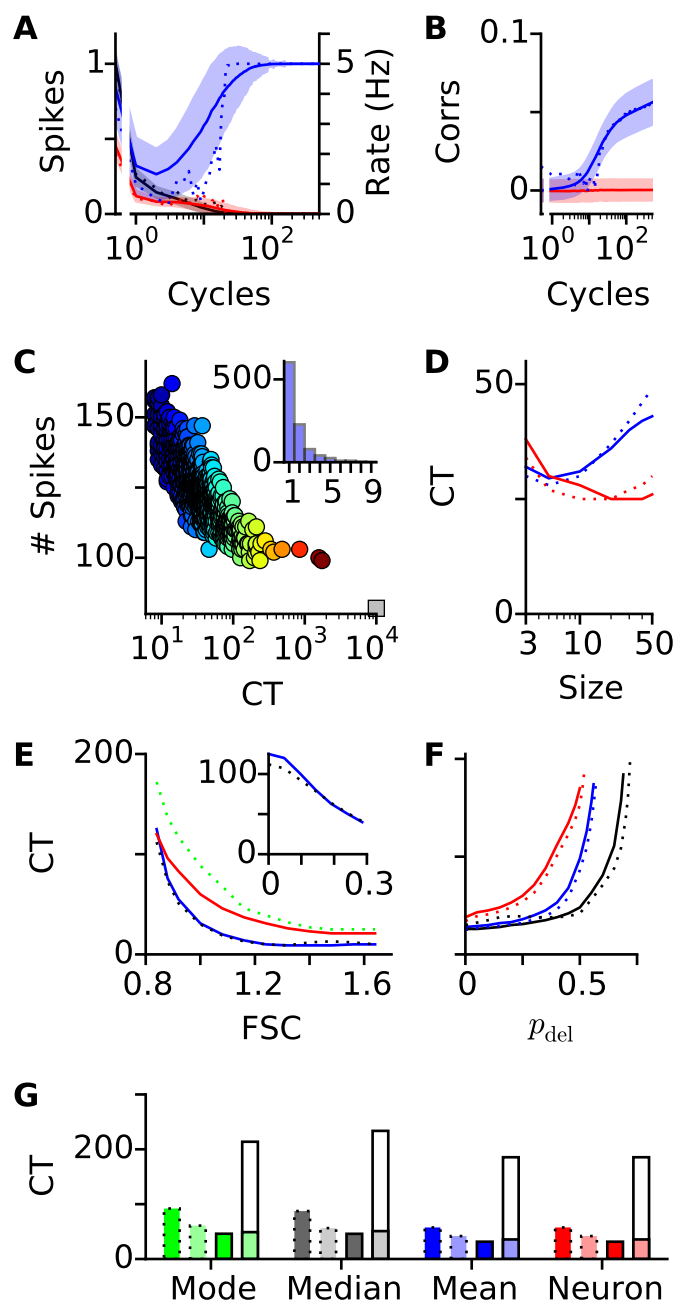


Fig. S6

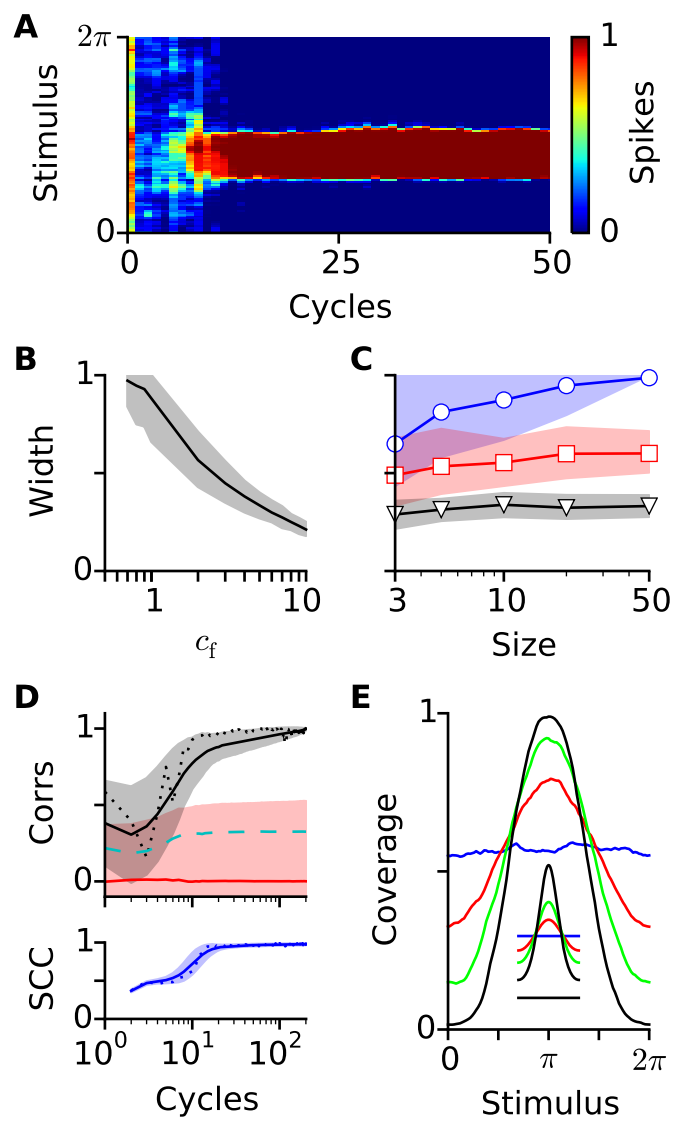


Fig. S7

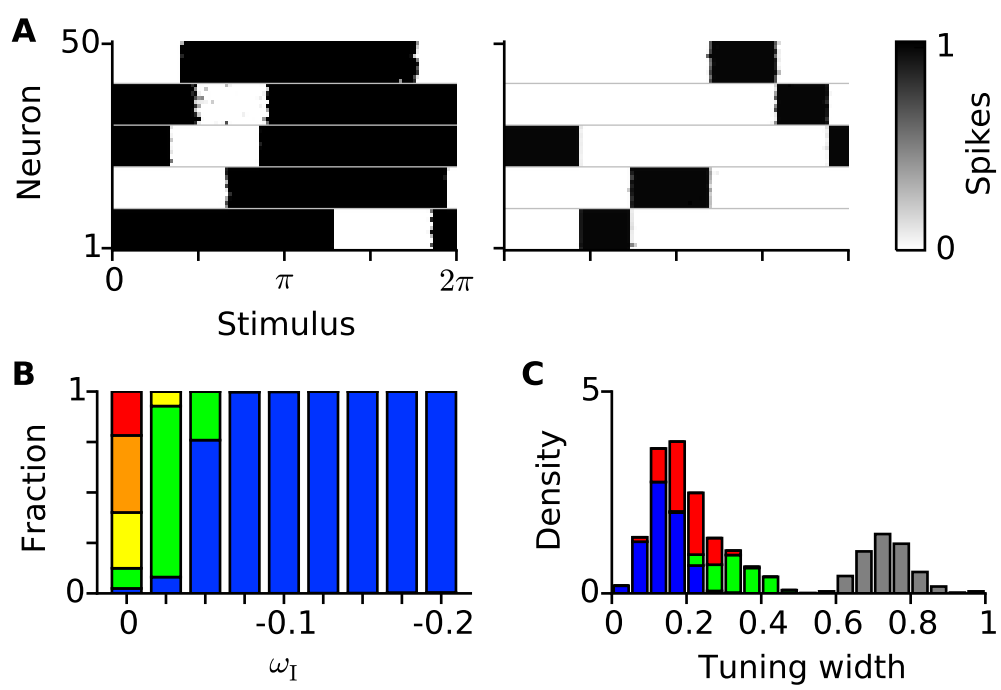


Fig. S8

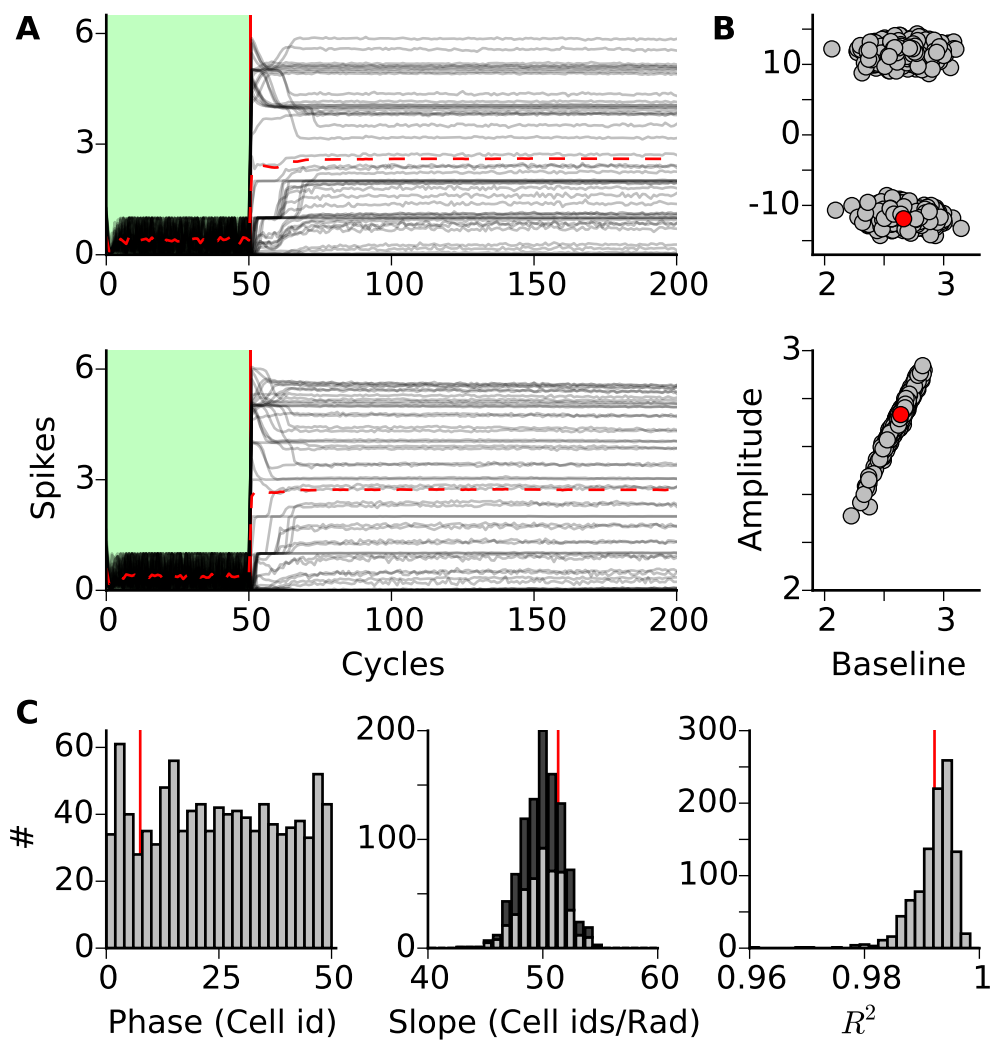


Fig. S9

Task	Label	Mean		STD	
1 Target	1	1.0000	(0.0001)	0.0006	(0.0044)
9 Distractors		0.0000	(0.0000)	0.0000	(0.0001)
Background (Hz)		0.0000	(0.0001)	0.0008	(0.0036)
1 Target	5	4.9996	(0.0014)	0.0045	(0.0046)
9 Distractors		0.0000	(0.0000)	0.0001	(0.0015)
Background (Hz)		0.0001	(0.0002)	0.0029	(0.0065)
5 Targets	1	1.0000	(0.0003)	0.0017	(0.0073)
5 Distractors		0.0000	(0.0000)	0.0000	(0.0010)
Background (Hz)		0.0000	(0.0001)	0.0009	(0.0038)
1 Target	1	1.0001	(0.0047)	0.0361	(0.0391)
1 Target	2	1.9988	(0.0087)	0.0246	(0.0204)
1 Target	3	2.9974	(0.0096)	0.0230	(0.0141)
1 Target	4	3.9951	(0.0150)	0.0231	(0.0123)
1 Target	5	4.9924	(0.0167)	0.0233	(0.0109)
5 Distractors		0.0001	(0.0005)	0.0043	(0.0116)
Background (Hz)		0.0098	(0.0093)	0.0641	(0.0281)

Table S1

Mean responses (Methods: Neural responses),  $R_y^{\text{mean}}$  and  $R_\emptyset^{\text{mean}}$  (in Hz) and their normalized standard deviations,  $R_y^{\text{std}}/R_y^{\text{mean}}$  and  $R_\emptyset^{\text{std}}/R_\emptyset^{\text{mean}}$  corresponding to the four tasks of Fig. 1B. Values in parentheses give the standard deviations over 1000 independent simulations.

Multi-spike tempotron			
N	Mean	(STD)	learning rate
500	7.8	( 4.8)	0.005
80	9.9	( 6.0)	0.02
20	22.0	(23.8)	0.08
Correlation-based learning			
500	3.7	(1.6)	0.02
80	8.5	(5.0)	0.05

Table S2

Average number of trials required by a single integrate-and-fire neuron to solve the reinforcement learning comparison task (4) (Methods: Reinforcement learning comparison task). Learning rates  $\lambda$  (Eq. 7) and  $\lambda_V$  (Eq. 8) were optimized for each number of inputs  $N$ .



## References and Notes

1. R. S. Sutton, Learning to predict by the methods of temporal differences. *Mach. Learn.* **3**, 9 (1988).
2. W. Schultz, P. Dayan, P. R. Montague, A neural substrate of prediction and reward. *Science* **275**, 1593–1599 (1997).
3. E. M. Izhikevich, Solving the distal reward problem through linkage of STDP and dopamine signaling. *Cereb. Cortex* **17**, 2443–2452 (2007).
4. J. Friedrich, R. Urbanczik, W. Senn, Spatio-temporal credit assignment in neuronal population learning. *PLOS Comput. Biol.* **7**, e1002092 (2011).
5. F. Ponulak, A. Kasiński, Supervised learning in spiking neural networks with ReSuMe: Sequence learning, classification, and spike shifting. *Neural Comput.* **22**, 467–510 (2010).
6. R. V. Florian, The chronotron: A neuron that learns to fire temporally precise spike patterns. *PLOS ONE* **7**, e40233 (2012).
7. Y. Xu, X. Zeng, S. Zhong, A new supervised learning algorithm for spiking neurons. *Neural Comput.* **25**, 1472–1511 (2013).
8. R. M. Memmesheimer, R. Rubin, B. P. Olveczky, H. Sompolinsky, Learning precisely timed spikes. *Neuron* **82**, 925–938 (2014).
9. F. Rosenblatt, *Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms* (Spartan Books, 1962).
10. R. Gütiğ, To spike, or when to spike? *Curr. Opin. Neurobiol.* **25**, 134–139 (2014).
11. H. S. Seung, Learning in spiking neural networks by reinforcement of stochastic synaptic transmission. *Neuron* **40**, 1063–1073 (2003).
12. R. Gütiğ, H. Sompolinsky, The tempotron: A neuron that learns spike timing-based decisions. *Nat. Neurosci.* **9**, 420–428 (2006).
13. I. R. Fiete, M. S. Fee, H. S. Seung, Model of birdsong learning based on gradient estimation by dynamic perturbation of neural conductances. *J. Neurophysiol.* **98**, 2038–2057 (2007).
14. L. R. Rabiner, A tutorial on hidden Markov models and selected applications in speech recognition. *Proc. IEEE* **77**, 257–286 (1989).
15. At first sight, it seems that losing the clue timing eliminates almost all the necessary feedback information, and asking for a specific number of spikes seems irrelevant because we do not really care about the total number of spikes if we only get at least one spike for each clue. However, a varying spike count per clue would indicate that the detector is operating in a regime in which noise could suppress or add spikes, which implies that the detector would not function reliably when the clues are embedded in background activity. This suggests that the number of spikes could serve as an objective function to stabilize an effective synaptic configuration and might also suffice for learning such a configuration from random initial conditions.

16. Because a neuron's memory for previous inputs decays rapidly (neural membrane time constants rarely exceed a few tens of milliseconds), it has to fire rather promptly in response to a clue, if at all.
17. B. Widrow, M. A. Lehr, 30 years of adaptive neural networks: Perceptron, Madaline, and backpropagation. *Proc. IEEE* **78**, 1415–1442 (1990).
18. D. Barber, "Learning in spiking neural assemblies" in *Advances in Neural Information Processing Systems*, S. Becker, S. Thrun, K. Obermayer, Eds. (MIT Press, 2003), vol. 15, pp. 149–156.
19. J. P. Pfister, T. Toyoizumi, D. Barber, W. Gerstner, Optimal spike-timing-dependent plasticity for precise action potential firing in supervised learning. *Neural Comput.* **18**, 1318–1348 (2006).
20. Although a population of 135 neurons (with 80 synapses each) required more than 3000 trials of reinforcement learning (4), the single integrate-and-fire neuron underlying the present study required an average of eight or four trials, when learning was implemented with the multi-spike tempotron or the correlation-based approximation (discussion and materials and methods, correlation-based learning), respectively. In fact, based on aggregate labels the task could be readily solved by a single neuron with only 20 synapses (table S2). Similarly strong performance differences between other neural implementations of gradient-based and reinforcement learning (11) have been reported (12) in the context of a binary classification task.
21. T. Yang, M. N. Shadlen, Probabilistic reasoning by neurons. *Nature* **447**, 1075–1080 (2007).
22. L. F. Abbott, S. B. Nelson, Synaptic plasticity: Taming the beast. *Nat. Neurosci.* **3** (suppl.), 1178–1183 (2000).
23. R. Gütiğ, H. Sompolinsky, Time-warp-invariant neuronal processing. *PLOS Biol.* **7**, e1000141 (2009).
24. R. G. Leonard, G. Doddington, *TIDIGITS*, vol. LDC93S10 (Linguistic Data Consortium, 1993).
25. T. Masquelier, R. Guyonneau, S. J. Thorpe, Spike timing dependent plasticity finds the start of repeating patterns in continuous spike trains. *PLOS ONE* **3**, e1377 (2008).
26. B. Nessler, M. Pfeiffer, L. Buesing, W. Maass, Bayesian computation emerges in generic cortical microcircuits through spike-timing-dependent plasticity. *PLOS Comput. Biol.* **9**, e1003037 (2013).
27. R. Linsker, Perceptual neural organization: Some approaches based on network models and information theory. *Annu. Rev. Neurosci.* **13**, 257–281 (1990).
28. R. Gütiğ, R. Aharonov, S. Rotter, H. Sompolinsky, Learning input correlations through nonlinear temporally asymmetric Hebbian plasticity. *J. Neurosci.* **23**, 3697–3714 (2003).
29. W. Schultz, Updating dopamine reward signals. *Curr. Opin. Neurobiol.* **23**, 229–238 (2013).
30. V. Pawlak, J. R. Wickens, A. Kirkwood, J. N. Kerr, Timing is not everything: Neuromodulation opens the STDP gate. *Front. Synaptic Neurosci.* **2**, 146 (2010).

31. A. Artola, S. Bröcher, W. Singer, Different voltage-dependent thresholds for inducing long-term depression and long-term potentiation in slices of rat visual cortex. *Nature* **347**, 69–72 (1990).
32. D. E. Feldman, The spike-timing dependence of plasticity. *Neuron* **75**, 556–571 (2012).
33. V. Pawlak, D. S. Greenberg, H. Sprekeler, W. Gerstner, J. N. Kerr, Changing the responses of cortical neurons from sub- to suprathreshold using single spikes in vivo. *eLife* **2**, e00012 (2013).
34. J. Amores, Multiple instance classification: Review, taxonomy and comparative study. *Artif. Intell.* **201**, 81–105 (2013).
35. H. Hotelling, Relations between two sets of variates. *Biometrika* **28**, 321–377 (1936).
36. S. Becker, G. E. Hinton, Self-organizing neural network that discovers surfaces in random-dot stereograms. *Nature* **355**, 161–163 (1992).
37. S. Becker, Mutual information maximization: Models of cortical self-organization. *Network* **7**, 7–31 (1996).
38. Y. Li, S. D. Van Hooser, M. Mazurek, L. E. White, D. Fitzpatrick, Experience with moving visual stimuli drives the early development of cortical direction selectivity. *Nature* **456**, 952–956 (2008).
39. S. D. Van Hooser, Y. Li, M. Christensson, G. B. Smith, L. E. White, D. Fitzpatrick, Initial neighborhood biases and the quality of motion stimulation jointly influence the rapid emergence of direction preference in visual cortex. *J. Neurosci.* **32**, 7258–7266 (2012).
40. L. Wiskott, T. J. Sejnowski, Slow feature analysis: Unsupervised learning of invariances. *Neural Comput.* **14**, 715–770 (2002).
41. N. Li, J. J. DiCarlo, Unsupervised natural experience rapidly alters invariant object representation in visual cortex. *Science* **321**, 1502–1507 (2008).
42. R. Min, M. Santello, T. Nevian, The computational power of astrocyte mediated synaptic plasticity. *Front. Comput. Neurosci.* **6**, 93 (2012).
43. P. Poirazi, T. Brannon, B. W. Mel, Pyramidal neuron as two-layer neural network. *Neuron* **37**, 989–999 (2003).
44. M. Larkum, A cellular mechanism for cortical associations: An organizing principle for the cerebral cortex. *Trends Neurosci.* **36**, 141–151 (2013).
45. J. A. Cummings, R. M. Mulkey, R. A. Nicoll, R. C. Malenka,  $\text{Ca}^{2+}$  signaling requirements for long-term depression in the hippocampus. *Neuron* **16**, 825–833 (1996).
46. R. C. Malenka, R. A. Nicoll, Long-term potentiation—A decade of progress? *Science* **285**, 1870–1874 (1999).
47. J. Hertz, A. Krough, R. G. Palmer, *Introduction to the Theory of Neural Computation* (Westview Press, 1991).
48. M. Wehr, G. Laurent, Odour encoding by temporal sequences of firing in oscillating neural assemblies. *Nature* **384**, 162–166 (1996).

49. T. Gollisch, M. Meister, Rapid neural coding in the retina with relative spike latencies. *Science* **319**, 1108–1111 (2008).
50. D. H. O'Connor, S. A. Hires, Z. V. Guo, N. Li, J. Yu, Q. Q. Sun, D. Huber, K. Svoboda, Neural coding during active somatosensation revealed using illusory touch. *Nat. Neurosci.* **16**, 958–965 (2013).
51. R. Gütig, T. Gollisch, H. Sompolinsky, M. Meister, Computing complex visual features with retinal spike times. *PLOS ONE* **8**, e53063 (2013).
52. D. J. Tolhurst, J. A. Movshon, A. F. Dean, The statistical reliability of signals in single neurons in cat and monkey visual cortex. *Vision Res.* **23**, 775–785 (1983).
53. M. N. Shadlen, W. T. Newsome, The variable discharge of cortical neurons: Implications for connectivity, computation, and information coding. *J. Neurosci.* **18**, 3870–3896 (1998).
54. M. Abeles, Role of the cortical neuron: Integrator or coincidence detector? *Isr. J. Med. Sci.* **18**, 83–92 (1982).
55. W. Singer, Neuronal synchrony: A versatile code for the definition of relations? *Neuron* **24**, 49–65, 111–125 (1999).
56. J. Biederlack, M. Castelo-Branco, S. Neuenschwander, D. W. Wheeler, W. Singer, D. Nikolić, Brightness induction: Rate enhancement and neuronal synchronization as complementary codes. *Neuron* **52**, 1073–1083 (2006).