

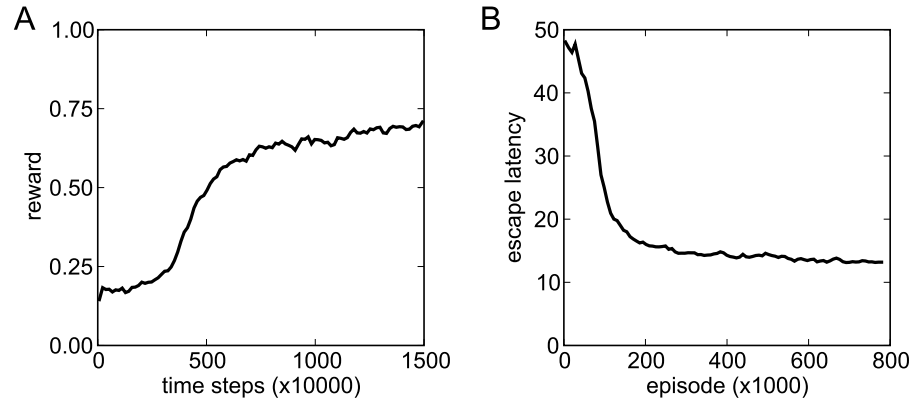


Figure S2: **Rewards and escape latencies during training of a feed-forward network of simple neurons on the control task with target and distractor.** A) Evolution of reward during training. A simulation step for all 100 parallel traces corresponds to 100 time-steps at the x-axis. The plotted values are averages over consecutive 150,000 time steps. B) Evolution of escape latencies (measured in time steps) during training. The number of episodes on the x-axis is the number of completed traces. The plotted values are averages over 8,000 consecutive episodes.