Correction: biological learning

Camille Gontier (camille.gontier@unibe.ch)
19/12/22

Exercise 1: Rescorla-Wagner (or delta) rule

1. **Pavlovian conditioning.** In Pavlov's animal's experiment, we denote $\bar{r}$ the reward expected by the animal, r the reward actually received, x the stimulus (e.g. the bell), w its weight, and η the learning rate.

   (a) What are the values of $\bar{r}$ and w at the beginning of the experiment for an untrained animal ?

   Correction. $\bar{r} = 0$ and $w = 0$ (the animal expects nothing).

   (b) At each trial, the animal is presented a stimulus ($x = 1$) and a reward ($r = 1$). Assume $\eta = 0.3$. What are the values of w and $\bar{r}$ after the first, second, third, and fourth trials ? Plot w as a function of the trial number.

   Hint: use the following steps:

   - Start from the current values of w and $\bar{r}$, from the previous trial;
   - Apply the learning rule $\Delta w = \eta(r - \bar{r})x$;
   - Update[1] the value of the weight: $w \leftarrow w + \Delta w$.
   - Compute the expected reward $\bar{r} = wx$; you now have the values of w and $\bar{r}$ for the current trial;
   - Repeat for the next trial.

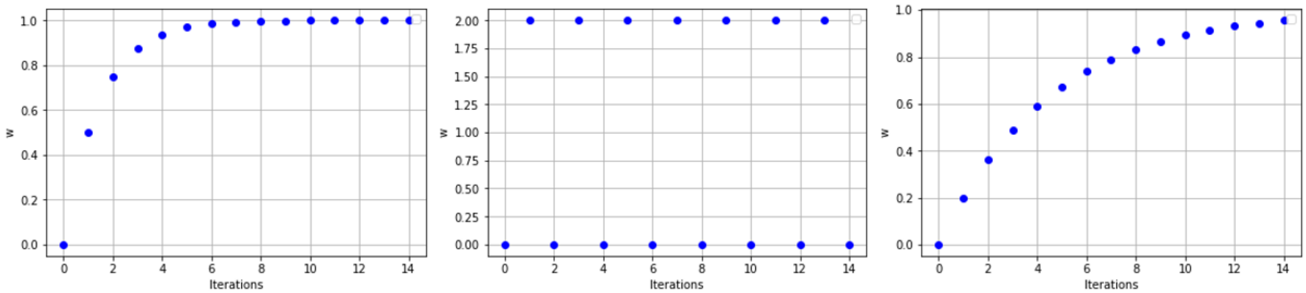
   Correction. At the first trial we have
   $\Delta w = \eta(r - \bar{r})x = 0.3(1 - 0) \times 1 = 0.3$
   $w = 0 + 0.3 = 0.3$
   $\bar{r} = wx = 0.3 \times 1 = 0.3$
   So after the first trial $\bar{r} = w = 0.3$
   Similarly, at the second trial:
   $\Delta w = \eta(r - \bar{r})x = 0.3(1 - 0.3) \times 1 = 0.21$
   $w = 0.3 + 0.21 = 0.51$
   $\bar{r} = wx = 0.51 \times 1 = 0.51$
   So after the second trial: $\bar{r} = w = 0.51$
   Third trial: $\bar{r} = w = 0.657$
   Fourth trial: $\bar{r} = w = 0.7599$

   (c) Repeat for $\eta = 0.8$. What happens after a sufficiently large number of trials ?

   Correction. First trial: $\bar{r} = w = 0.8$
   Second trial: $\bar{r} = w = 0.96$
   Third trial: $\bar{r} = w = 0.992$
   Fourth trial: $\bar{r} = w = 0.9984$
   After a sufficiently large number of trials, w converges towards 1. In this case, the expected reward $\bar{r} = wx$ corresponds to the actual reward (which is the aim of the RW learning rule). As the difference between the actual and the predicted rewards is 0, w stops increasing.

   (d) Which of the following curves correspond to $\eta = 0.2$? $\eta = 0.5$? $\eta = 2$?

[1] Δ is the classical symbol to indicate the evolution of a quantity. Intuitively, $\Delta w =$ "new value of w" - "old value of w".

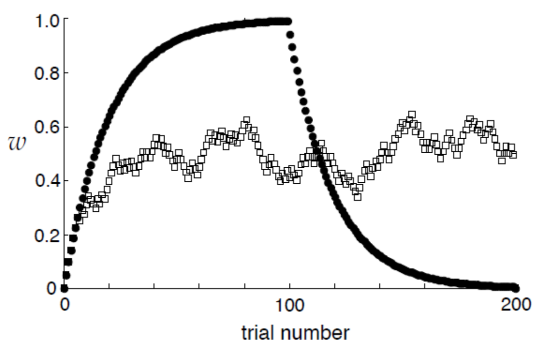


Correction. $\eta = 0.2 \rightarrow$ third picture.
$\eta = 0.5 \rightarrow$ first picture (learning is faster as the learning rate increases).
$\eta = 2 \rightarrow$ second picture (upon a certain threshold, the learning scheme becomes unstable).

2. **Partial blocking.** Besides Pavlovian conditioning, the RW rule can also be applied to other paradigms (partial conditioning, blocking, overshadowing...) We now assume that, at each trial, the animal is indeed presented a stimulus ($x = 1$), but that the reward is randomly and uniformly distributed between 0 and 1 ($r \sim \mathcal{U}([0, 1])$).

   (a) What is the average value of the reward $\langle r \rangle$?
   Correction. $\langle r \rangle = 0.5$

   (b) If the animal expects a reward $\bar{r} = \langle r \rangle$, what is the average value $\langle \delta \rangle$ of the reward prediction error ?
   Correction. $\langle \delta \rangle = \langle r - \bar{r} \rangle = \langle r \rangle - \bar{r} = 0$



Acquisition and extinction curves for Pavlovian conditioning and partial reinforcement as predicted by the Rescorla-Wagner model. The filled circles show the time evolution of the weight w over 200 trials. In the first 100 trials, a reward of $r = 1$ was paired with the stimulus, while in trials 100-200 no reward was paired ($r = 0$). Open squares show the evolution of the weights when a reward of $r = 1$ was paired with the stimulus randomly on 50% of the trials. From Dayan and Abbott 2001, p.334

3. **Overshadowing.** In this paradigm, the animal is presented 2 different stimuli: e.g. a bell (x_1) and a light (x_2). Each of them is associated with respective weights w_1 and w_2. The expected reward is then computed as the scalar product of the vectors $\mathbf{w} = (w_1, w_2)$ and $\mathbf{x} = (x_1, x_2)$:

$$\bar{r} = \mathbf{w} \cdot \mathbf{x} = w_1 x_1 + w_2 x_2$$

Similarly, weights are updated as

$$\mathbf{w} \to \mathbf{w} + \eta(r - \bar{r})\mathbf{x}$$

(the same learning rate η is used for both stimuli).

(a) At each trial, the animal is presented both stimuli ($x_1 = 1$ and $x_2 = 1$) and a reward ($r = 1$). Assume $\eta = 0.4$. What are the values of w_1 and w_2 after the first, second, third, and fourth trials ?

Correction. First trial: $w_1 = w_2 = 0.4$

Second trial: $w_1 = w_2 = 0.48$

Third trial: $w_1 = w_2 = 0.496$

Fourth trial: $w_1 = w_2 = 0.4992$

(b) What is the value of $w_1 + w_2$ after application of the Rescorla-Wagner (or delta) rule ? How do you interpret this result ?

Correction. $w_1 + w_2 = 1$. The combination of the 2 stimuli x_1 and x_2 is interpreted as a single conditioned stimulus of total weight $w = 1$.

(c) After learning, what will be the expected reward if only the first stimulus is presented ($x_1 = 1$ and $x_2 = 0$)?

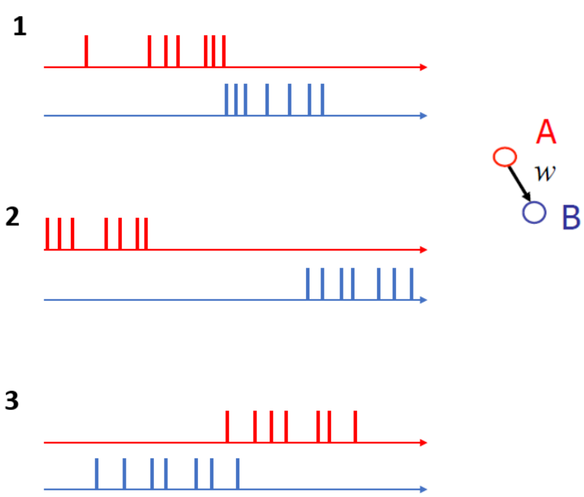
Correction. $\bar{r} = 0.5$ (only half of the reward is expected as only half of the conditioned stimulus is presented).

(d) Assume the animal has learned the weights $w_1 = 0.8$ and $w_2 = 0.2$. What will be the expected reward if only the first stimulus is presented ($x_1 = 1$ and $x_2 = 0$)? Justify that in this case, contrary to the previous question, x_1 *overshadows* x_2.

Correction. $\bar{r} = 0.8$. In this case, the sum of the weights is still equal to 1, but $w_1 > w_2$ means that x_1 is more salient (and thus overshadows) x_2.

Exercise 2: Spike-Timing Dependent Plasticity (STDP)

1. **Introduction.** Herebelow are presented 3 sets of synaptic pair-wised trainings. Pre-synaptic spikes are presented above in red, post-synaptic spikes are presented below in blue. On average, for which of these training sets is the synaptic weight w going to increase ? Decrease ? Remain unchanged ?



Correction. First plot: w is going to increase (on average, pre-synaptic spikes take place shortly before post-synaptic spikes).

Second plot: w stays unchanged as the Δt between pre and post-synaptic spikes is too high.

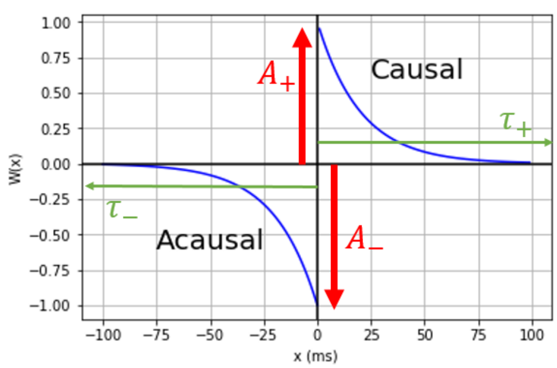
Third plot: w is going to decrease (on average, pre-synaptic spikes take place shortly after post-synaptic spikes).

2. **Forced training.** In the pair-based STDP model with all-to-all interactions, the variation of synaptic weight Δw depends on the relative timing of the post-synaptic spikes $(t_i^{post})_i$ and pre-synaptic spikes $(t_j^{pre})_j$. For n post-synaptic and m pre-synaptic spikes, Δw can be computed as

$$\Delta w = \sum_{i=1}^{n} \sum_{j=1}^{m} W(t_i^{post} - t_j^{pre})$$

where $W(x)$ is a STDP function, for which a popular choice is

$$\begin{cases} x > 0 \implies W(x) = A_+ \exp(-x/\tau_+) \\ x < 0 \implies W(x) = -A_- \exp(x/\tau_-) \end{cases}$$



Physical meaning of the parameters A_+,A_-,τ_+, and τ_-

(a) Assume $A_+ = A_- = 1$ and $\tau_+ = \tau_- = 10$ms. The pre-synaptic neuron is activated 2 times, at $t_1^{pre} = 2$ms and $t_2^{pre} = 5$ms. The post-synaptic neuron is also activated 2 times, at $t_1^{post} = 7$ms and $t_2^{post} = 10$ms. Is w going to increase or decrease ?

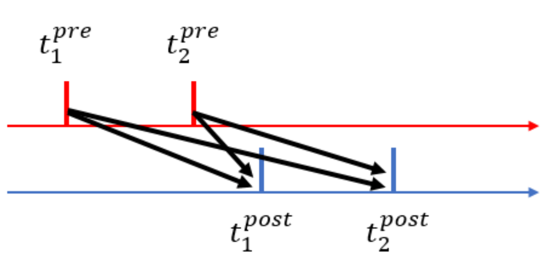


Illustration of all-to-all interactions. With 2 pre-synaptic spikes, and 2 post-synaptic spikes, there are 4 interactions to consider: t_1^{pre} to t_1^{post}, t_1^{pre} to t_2^{post}, t_2^{pre} to t_1^{post}, and t_2^{pre} to t_2^{post}.

Correction. w is going to increase (as all pre-synaptic spikes take place shortly before post-synaptic spikes).

(b) Compute Δw for the spike trains described in (a).
Correction. $\Delta w = W(t_1^{post} - t_1^{pre}) + W(t_1^{post} - t_2^{pre}) + W(t_2^{post} - t_1^{pre}) + W(t_2^{post} - t_2^{pre}) \approx 2.48$

(c) What will happen to w if we keep on stimulating the neuron by repeating the same spike pattern? Is that realistic ?

Correction. The positive increments Δw will keep on adding up and w will eventually diverge to ∞ (while an infinite synaptic weight would be biologically unrealistic).

(d) To solve this problem, a possible solution[2] is to set $A_+(w) = w_{max} - w$, where w_{max} is a fixed parameter and w is the current value of the synaptic weight ($w \to w + \Delta w$).

i. Compute w after the spike trains from (a) (i.e. $t_1^{pre} = 2$ms and $t_2^{pre} = 5$ms; $t_1^{post} = 7$ms and $t_2^{post} = 10$ms) has been delivered. Assume $w_{max} = 1$ and an initial weight $w = 0.5$.
Correction. $w \approx 0.9923$

ii. Repeat for $w_{max} = 100$.
Correction. $w \approx 98.462$

iii. How can we set the maximum weight that the synapse will reach?
Correction. w will not exceed w_{max} and will converge towards this value if the spike trains pattern is repeated.

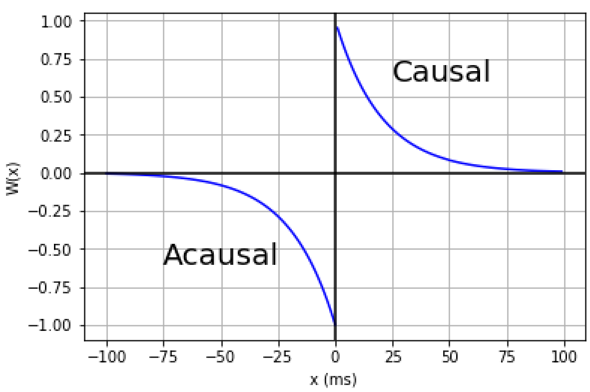
[2]Such a solution is called soft-bounds weight dependence. Interested readers can have a look at this article: `http://www.scholarpedia.org/article/Spike-timing_dependent_plasticity`

3. **Random spikes.** We now study how w varies when neurons are freely spiking at random times.

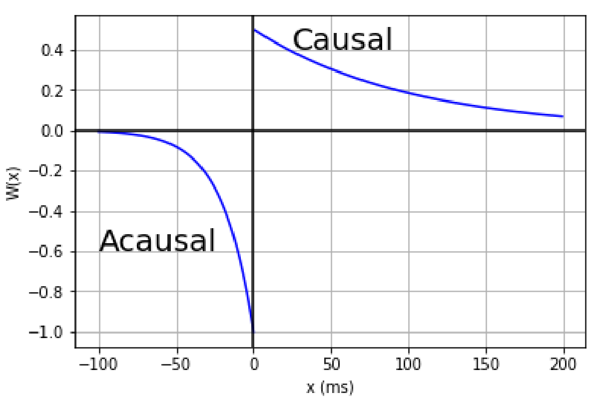
   (a) Assume the following STDP function:

$$\begin{aligned} x > 0 &\implies W(x) = A_+ \exp(-x/\tau_+) \\ x < 0 &\implies W(x) = -A_- \exp(x/\tau_-) \end{aligned} \quad (1)$$

   where τ_+ and τ_- are positive parameters.
   For $t^{post} - t^{pre} = \Delta t$, the corresponding modification of synaptic weight is $\Delta w = W(\Delta t)$.
   Assuming $\tau_+ = \tau_- = 20ms$, $A_+ = 1$, and $A_- = 1$, show that $\forall x \in \mathbb{R}$, $W(-x) = -W(x)$ and plot W.
   Correction.
   $x > 0 \implies -x < 0 \implies W(-x) = -\exp(-x/\tau_-) = -W(x)$
   $x < 0 \implies -x > 0 \implies W(-x) = \exp(x/\tau_+) = -W(x)$



   (b) Plot again the STDP function W, assuming this time $A_+ = 0.5$, $A_- = 1$, $\tau_+ = 100ms$ and $\tau_- = 20ms$.
   Correction.



   (c) 3 different stimulation protocols are applied to a synapse:

      i. The presynaptic neuron fires exactly once every second, and a postsynaptic spike occurs exactly 50ms after each presynaptic spike;

ii. The presynaptic neuron fires exactly once every second, and the postsynaptic neuron fires also once per second but with a delay δt randomly chosen between -10ms and 10ms;

iii. Presynaptic and postsynaptic neurons are freely spiking at random times.

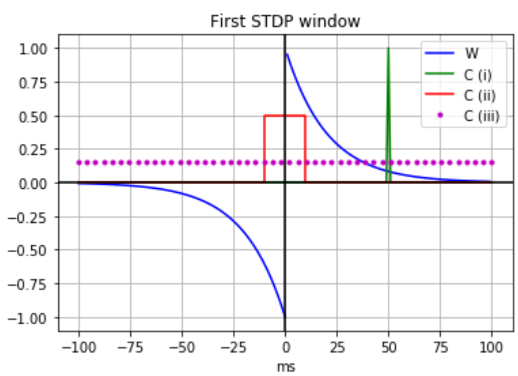
How would you expect each of these protocols to modify the synaptic weight w, given each of the STDP windows described in (a) and (b) ? Provide a graphical justification. Are your results for the first protocol coherent with Hebb's rule ?

Hint: The expected weight change Δw is given by

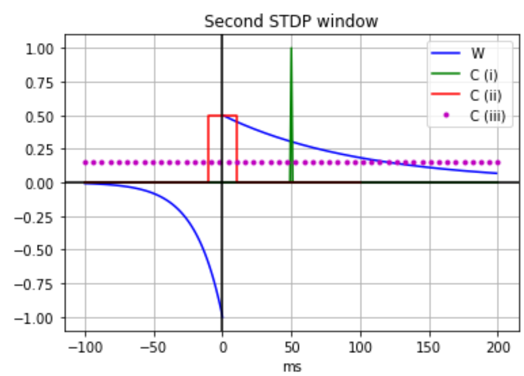
$$\Delta w = \int_{-\infty}^{\infty} W(s)C(s)ds$$

where s represents the values of $t^{post} - t^{pre}$ and $C(s)$ is the cross-correlation of the pre and post-synaptic spike trains. Intuitively, $C(s)$ measures the distribution of $t^{post} - t^{pre}$.

Correction. A sketch of $C(s)$ is plotted on the figures below, along with the different STDP windows $W(s)$, with which it can be compared to assess how w will vary for each protocol.



(a)



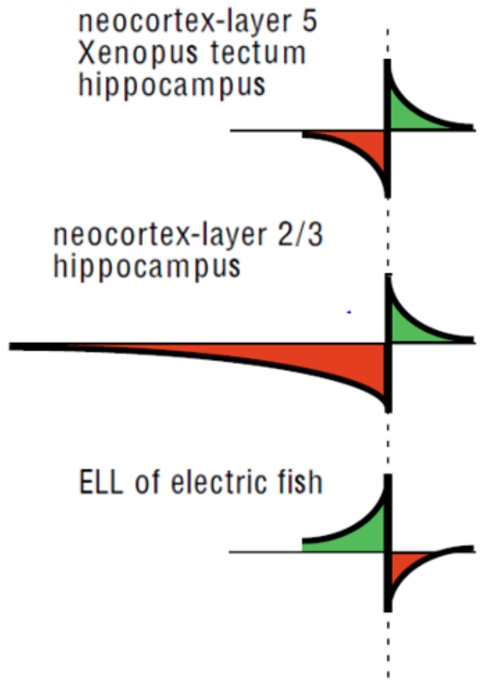
(b)

i. For the first protocol, the delay $t^{post} - t^{pre}$ is always equal to 50ms (we only consider pairs of spikes, as 50ms $\ll$ 1s). As the causal parts of both windows are positive, we expect w to increase. This is coherent with Hebb's rule: as the postsynaptic neuron spikes slightly after ($t^{post} - t^{pre} > 0$) the presynaptic neuron (i.e. they fire together), we expect the weight to increase (i.e. they wire together). Note however that Hebb's rule is only about synaptic potentiation (i.e. $\Delta w > 0$) and does not predict the acausal part of the learning rule (i.e. depression, $\Delta w < 0$).

ii. As the first STDP window is antisymmetric, $\int_{-10}^{10} W(s)ds = 0$, so we expect w to remain unchanged.
For the second STDP window, since $A_- > A_+$, and since $10ms \ll \tau_+, \tau_-$: $\int_{-10}^{10} W(s)ds < 0$, so w will decrease.

iii. As the first STDP window is antisymmetric, $\int_{-\infty}^{\infty} W(s)ds = 0$, so we expect w to remain unchanged.
For the second STDP window, since $A_+\tau_+ > A_-\tau_-$,$\int_{-\infty}^{\infty} W(s)ds > 0$, so w will increase.

(d) Herebelow are represented the STDP functions for 3 different kinds of synapses. For each of them, how will w vary after a sufficiently high number of randomly distributed pre- and post-synaptic spikes? Which of them are causal learning rules ?



Correction. First plot: $\int_{-\infty}^{\infty} W(s)ds \approx 0 \implies \Delta w \approx 0$

Second plot: $\int_{-\infty}^{\infty} W(s)ds < 0 \implies \Delta w < 0$

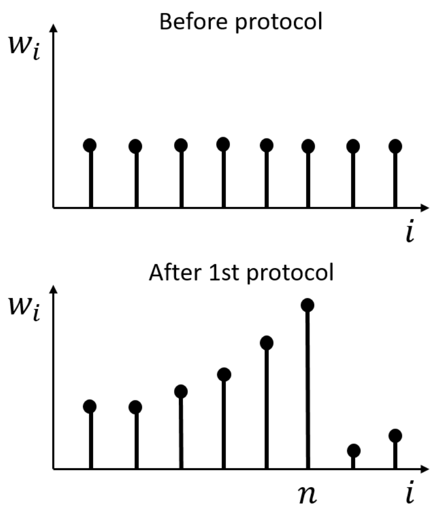
Third plot: $\int_{-\infty}^{\infty} W(s)ds > 0 \implies \Delta w > 0$

The first 2 rules are causal learning rules, since $\Delta w > 0$ for $t_{post} - t_{pre} > 0$.

4. **Earliest predictor.** We now assume that the pre-synaptic spike times are forced, while post-synaptic firing is free. The post-synaptic is linked to N pre-synaptic neurons which all fire successively: the first neuron at $t_1 = 1$ms, the second at $t_2 = 2$ms, etc. These inputs will build up and the post-synaptic neuron will elicit a spike once its membrane current reaches a certain threshold. We assume this occurs after the n^{th} pre-synaptic spike. We note w_i the synaptic weight between the i^{th} pre-synaptic unit and the post-synaptic unit.

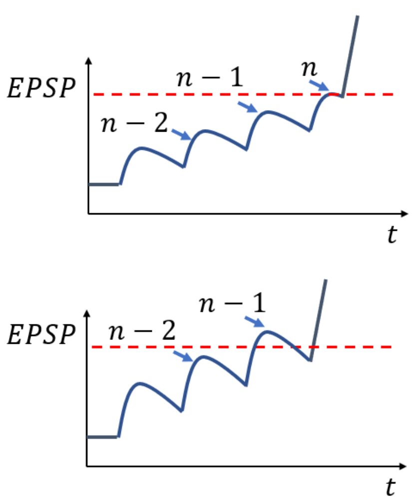
   (a) After the post-synaptic spike, how does w_n vary? w_{n-1}? w_{n-2}?

   Correction. As the post-synaptic spike happens immediately after the n^{th} pre-synaptic spike, w_n is going to increase. So is w_{n-1}, but slightly less than w_n because the delay between t_n and the post-synaptic spike is shorter. The same applies to all the previous pre-synaptic spikes. See below before STDP (upper figure) and after STDP (lower figure).



(b) We repeat the same experiment a second time. How is the firing time of the post-synaptic unit going to evolve? Remember that a higher synaptic weight w means that, upon arrival of a pre-synaptic spike, the increase in post-synaptic current will also be higher.

Correction. As w_{n-1} is increased, the increment of post-synaptic current elicited by the $(n-1)^{th}$ spike is also increased, making it more likely to elicit a post-synaptic spike at the next trial. Trial after trials, the post-synaptic spike is thus likely to be elicited earlier. See below before STDP (upper figure) and after STDP (lower figure).



(c) We repeat the same experiment several times. Explain why the first pre-synaptic spike at $t_1 = 1$ms is eventually called the *earliest predictor.*

Correction. The first pre-synaptic spike at $t_1 = 1$ms is the beginning of a pattern which, if repeated, eventually leads to the post-synaptic spike being elicited earlier and earlier, up to t_1.