The Neuroscience of Learning

Dr. Olav E. Krigolson
Associate Director
Centre for Biomedical Research
University of Victoria

University of Victoria | Theoretical and Applied Neuroscience Laboratory

How do we learn?

FEEDBACK

Definition: Learning

Changes in internal processes that are reflected by relatively stable changes in performance.

Inherent versus Augmented Feedback

TABLE 12.1. Dimensions of Augmented Feedback

Concurrent: Presented during the movement	**Terminal:** Presented after the movement
Immediate: Presented immediately after the relevant action	**Delayed:** Delayed in time after the relevant action
Verbal: Presented in a form that is spoken or capable of being spoken	**Nonverbal:** Presented in a form that is not capable of being spoken
Accumulated: Feedback that represents an accumulation of past performance	**Distinct:** Feedback that represents each performance separately

Knowledge of Results versus Knowledge of Performance

Comparison of KR and KP

Knowledge of Results	Knowledge of Performance

Similarities

Verbal
Extrinsic
Post-response

Differences

Outcome	Information about outcome pattern / quality
Redundant with inherent feedback	Distinct from inherent feedback

Feedback Frequency

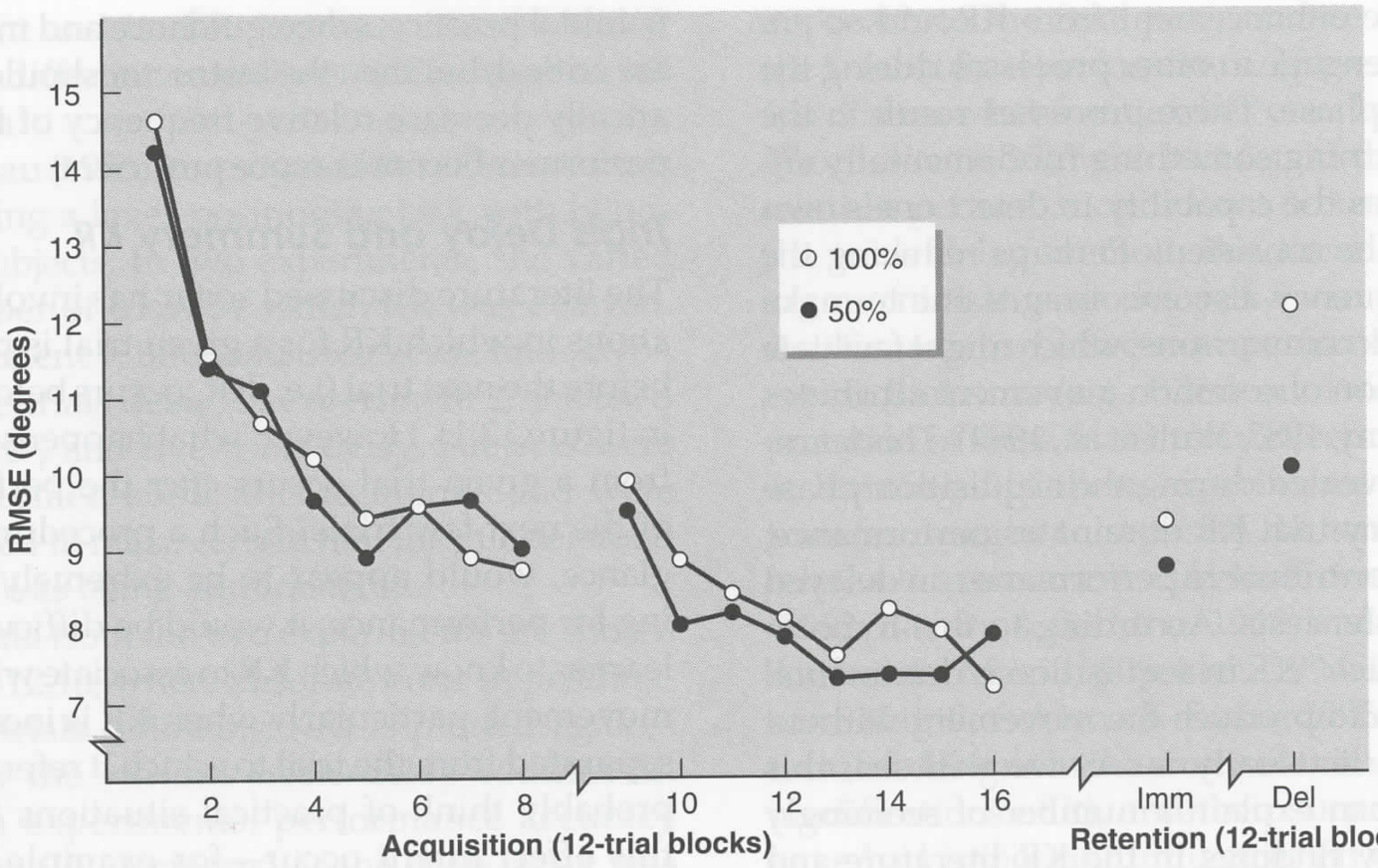


FIGURE 12.8 Effects of 100% versus 50% relative frequency of knowledge of results (KR) in acquisition and retention.

Reprinted, by permission, from C.J. Winstein and R.A. Schmidt, 1990, "Reduced frequency of knowledge of results enhances motor skill learning," *Journal of Experimental Psychology: Learning, Memory, and Cognition* 16; 910. Copyright © 1990 by the American Psychological Association.

Feedback Timing

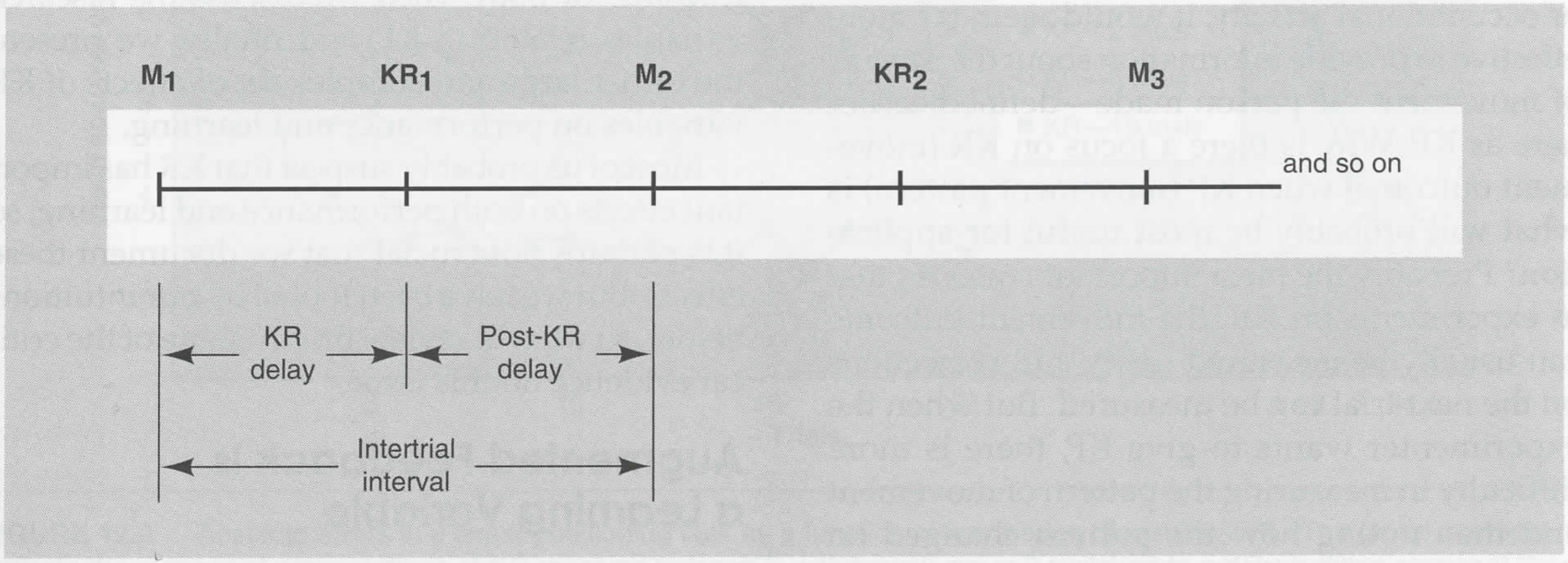


FIGURE 12.1 Temporal placement of events in the knowledge-of-results (KR) paradigm. M_1 refers to movement trial 1. KR_1 refers to the augmented feedback provided about results of movement trial 1.

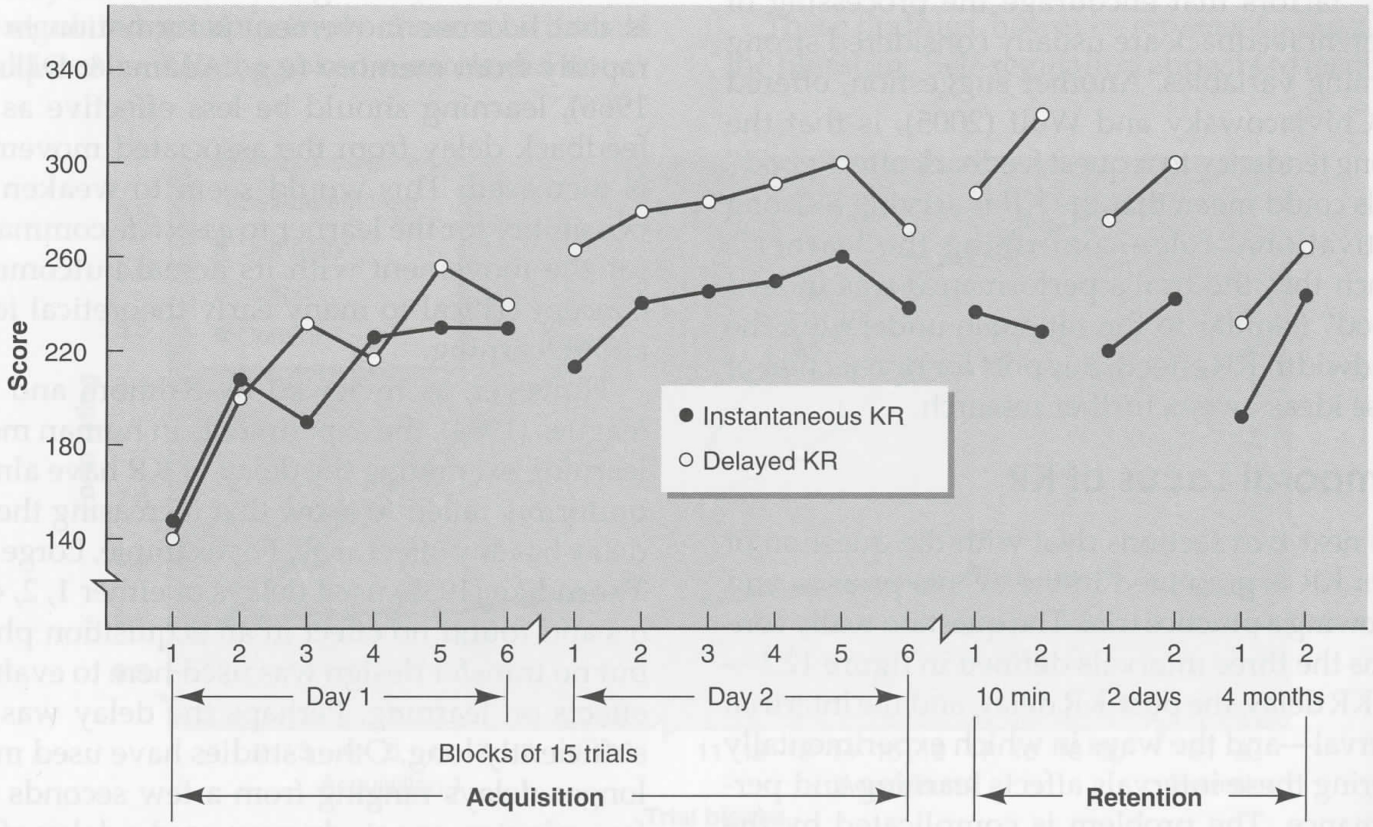


FIGURE 12.15 Performance scores of instantaneous- and delayed-KR (knowledge of results) conditions in acquisition and retention.

Reprinted, by permission, from S.P. Swinnen et al., 1990, "Information feedback for skill acquisition: instantaneous knowledge of results degrades learning," *Journal of Experimental Psychology: Learning, Memory, and Cognition* 16: 712. Copyright © 1990 by the American Psychological Association.

Post-KR Delay Interval

Learner is actively modifying and creating new movement in this interval

1. Shortening post-KR interval may impair learning

2. Shea & Upton (1976)
   Activity during KR/KP interferes with learning of the task

What sort of feedback can you provide?

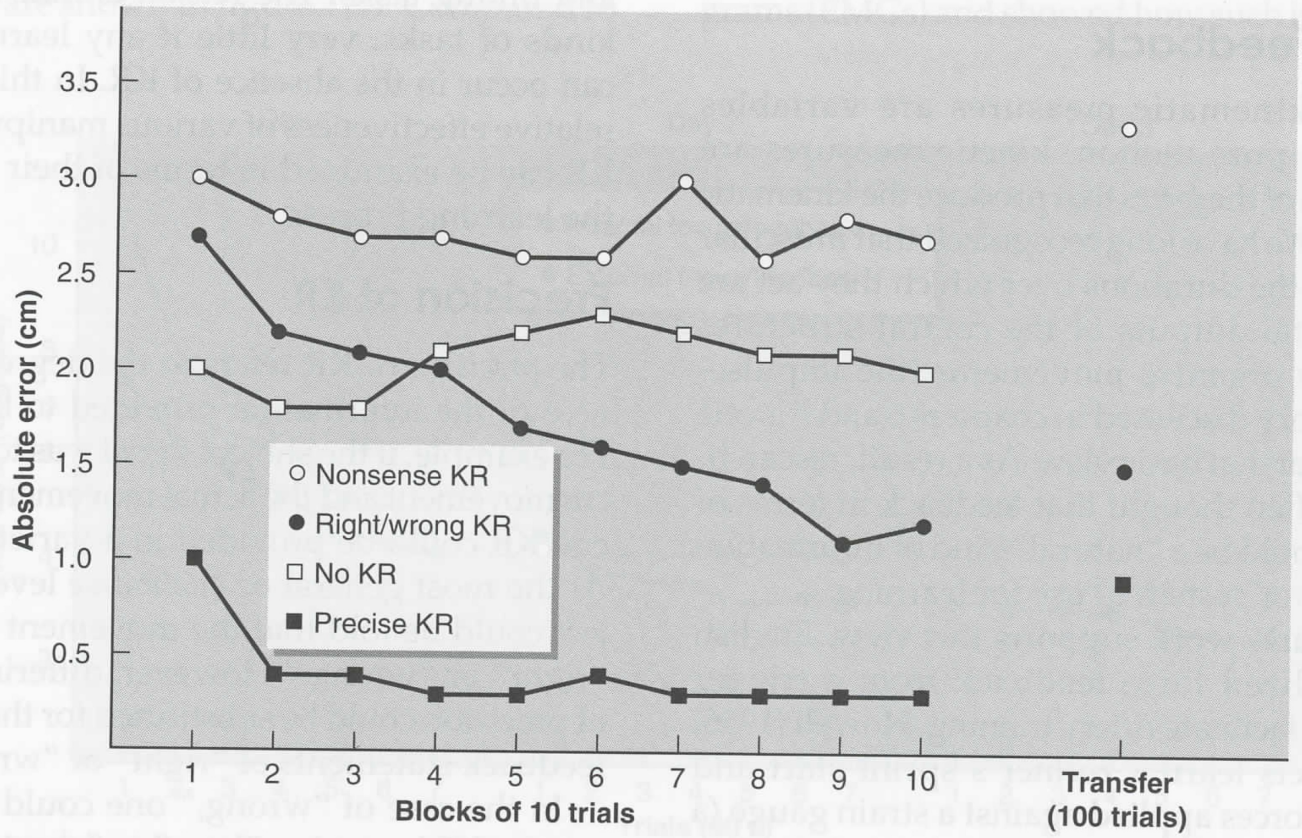


FIGURE 12.5 Qualitative and quantitative knowledge of results (KR) effects in acquisition and transfer. The No-KR group did not perform the transfer rest.

Data from Trowbridge and Cason 1932.

Techniques to Reduce Feedback Frequency

1. Bandwidth Feedback
2. Self-Selected Frequency
3. Summary Feedback

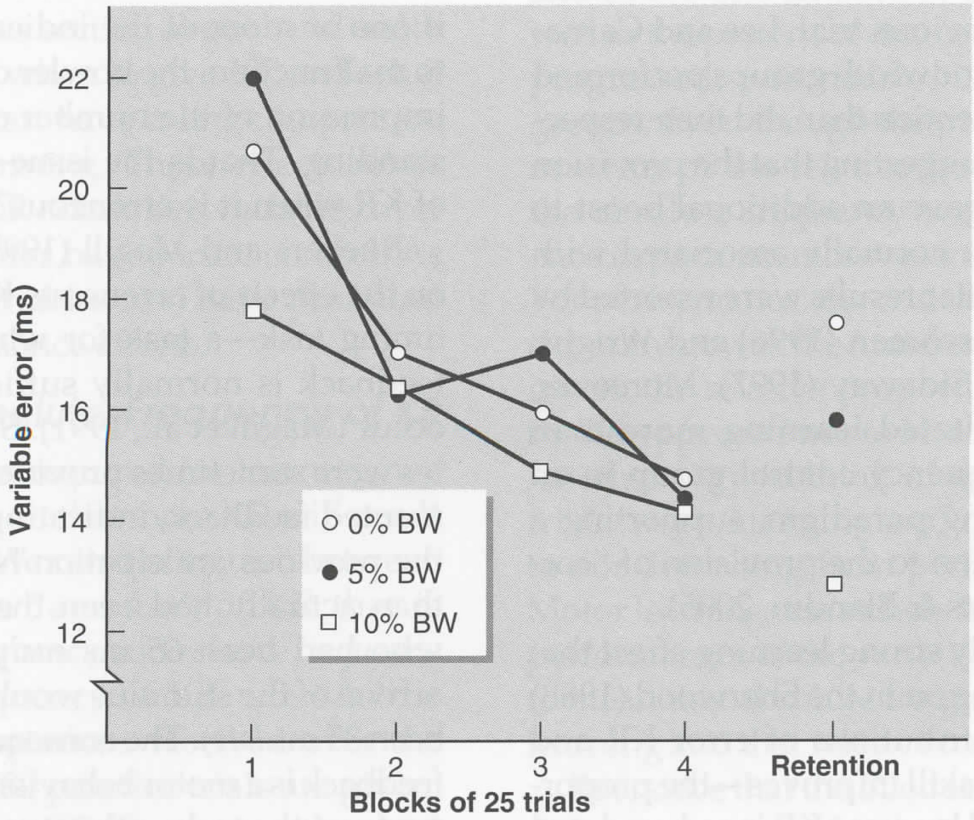


FIGURE 12.6 Bandwidth knowledge-of-results (KR) effects in acquisition and retention.

Reproduced and adapted from Table 1 and 2 with permission of author and publisher from: Sherwood, D.E. Effect of bandwidth knowledge of results on movement consistency. *Perceptual and Motor Skills*, 1988, 66, 535-542. © Perceptual and Motor Skills 1988.

Video Feedback

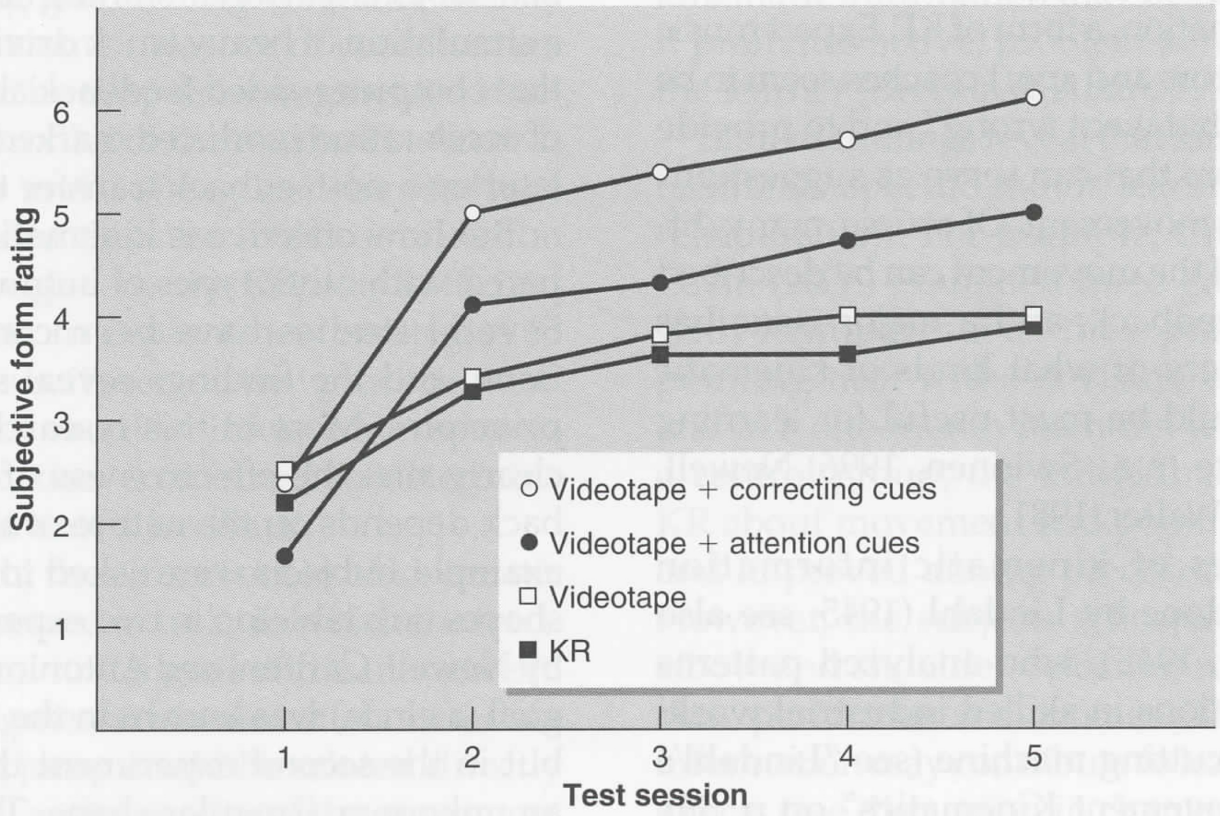


FIGURE 12.3 Throwing performance under various conditions of videotape replays.

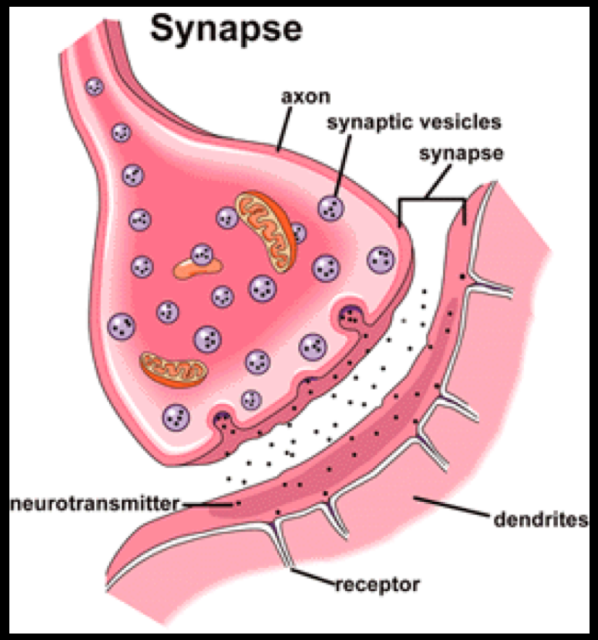
Data from Kernodle and Carlton 1992.

How does feedback work?

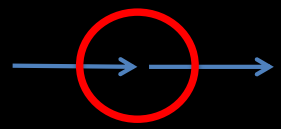
You do not learn from your mistakes…

You learn when expectancies deviate from outcomes

We learn from…


Synapse
axon
synaptic vesicles
synapse
neurotransmitter
dendrites
receptor

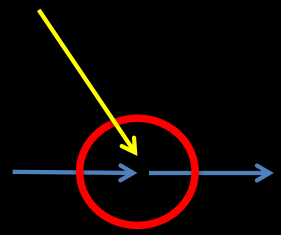
REPETITION

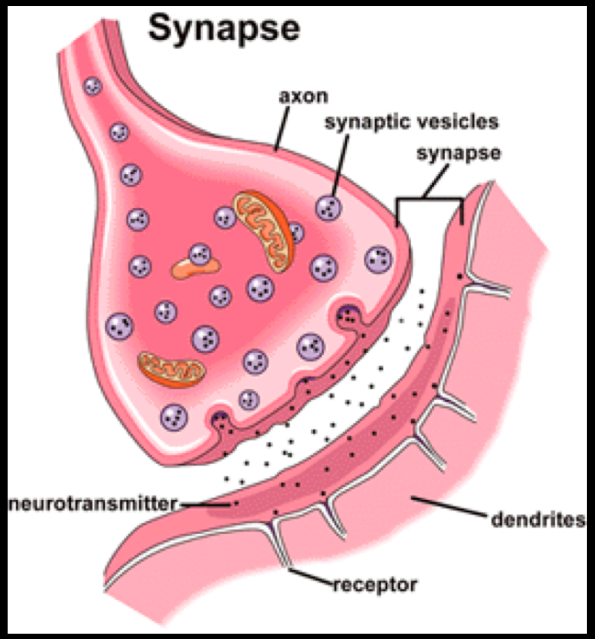


In practice, we talk about the strength of a connection in terms of a “weight” or a “value”.

But we also learn from…

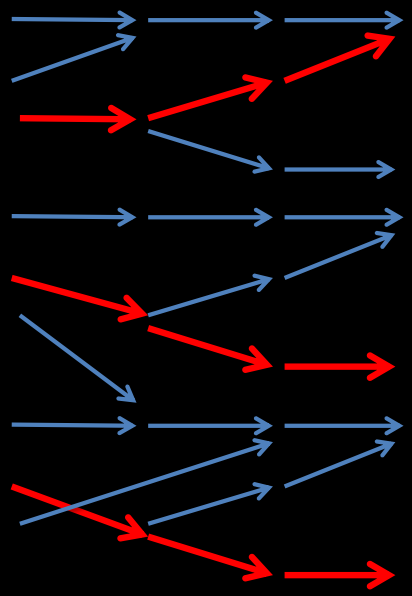

44



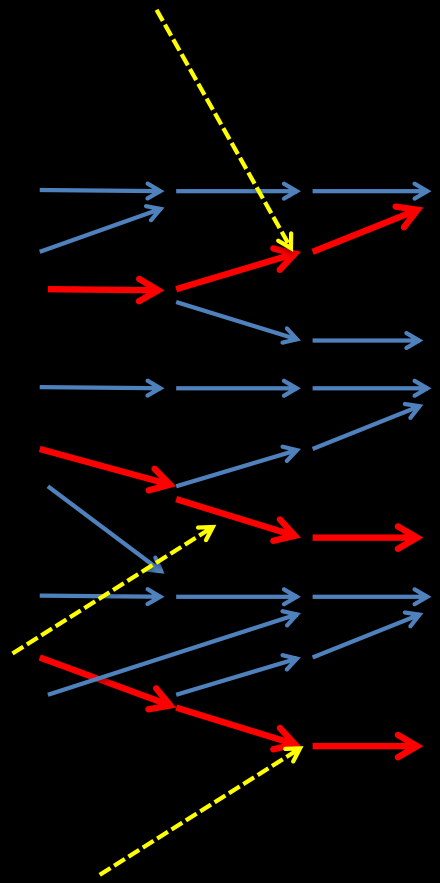


REPETITION

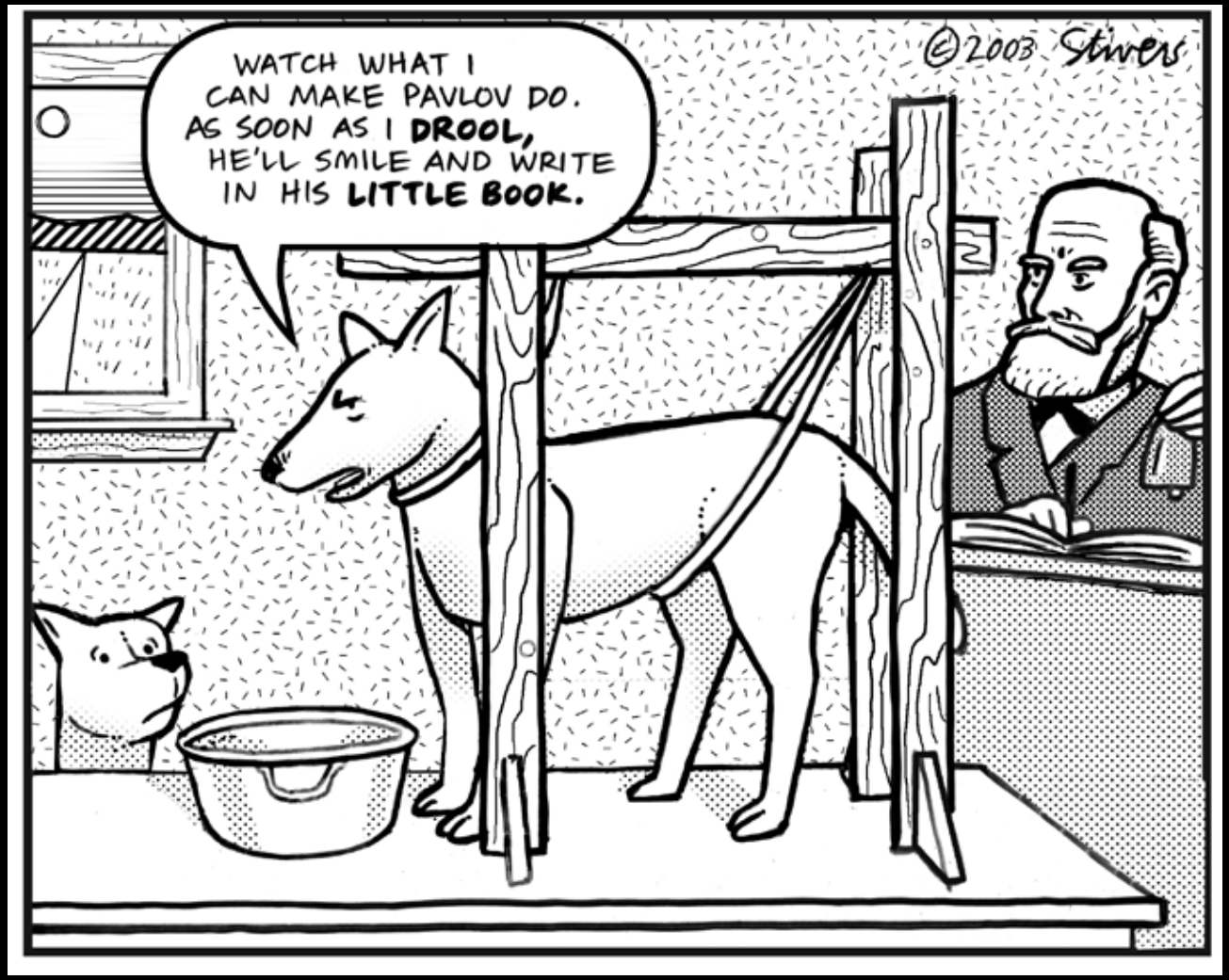
But what does feedback do?

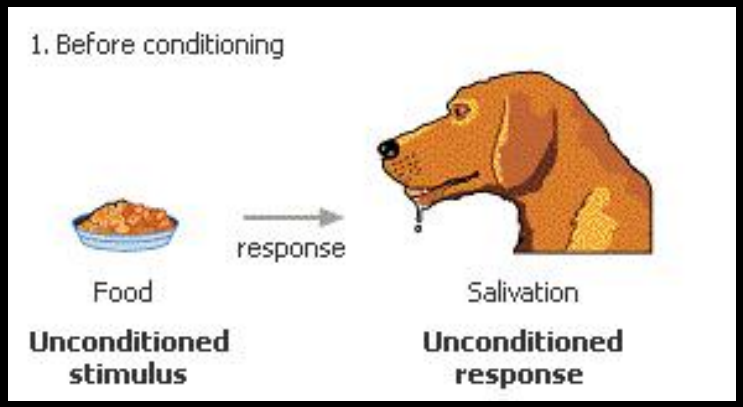


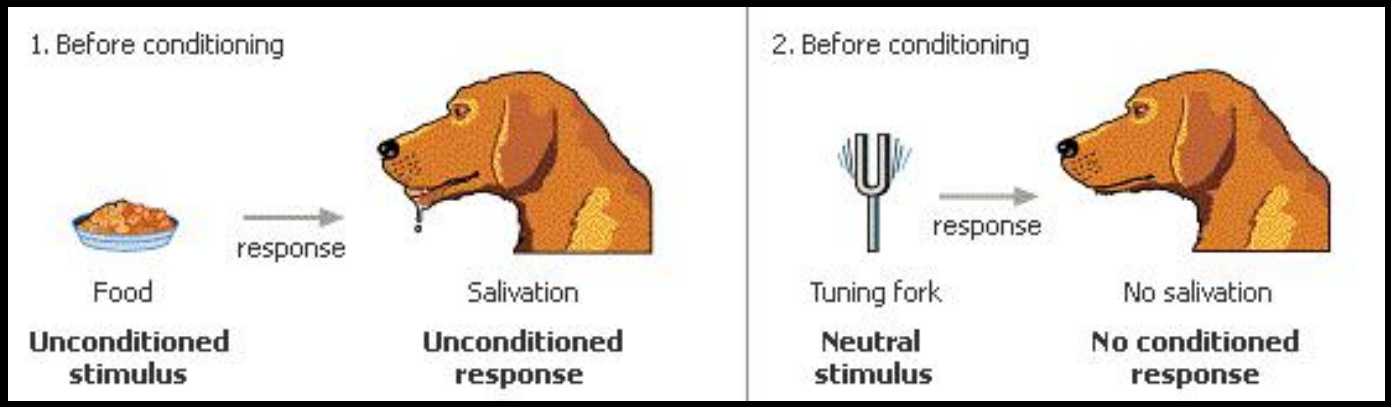
Hebbian Learning

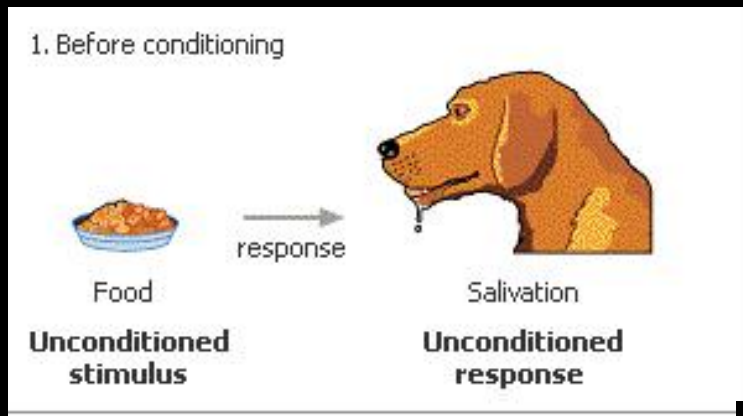


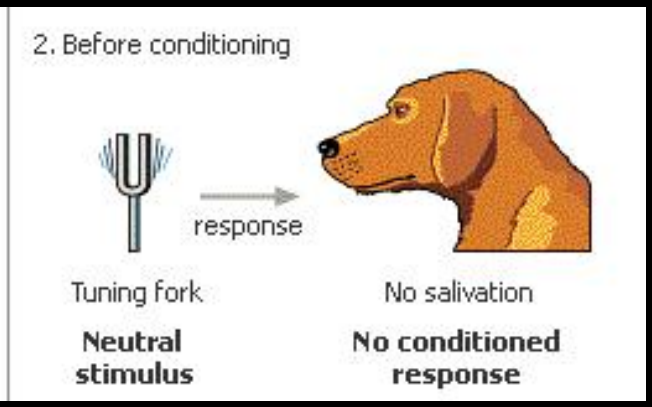
Hebbian Learning &
Reinforcement Learning

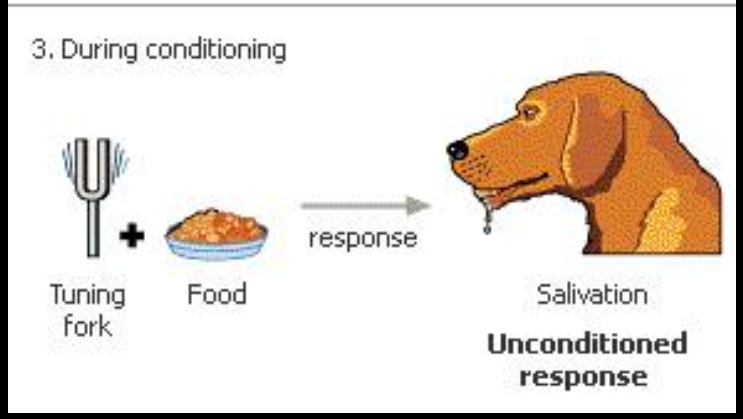

WATCH WHAT I CAN MAKE PAVLOV DO. AS SOON AS I DROOL, HE'LL SMILE AND WRITE IN HIS LITTLE BOOK.
©2003 Stivers

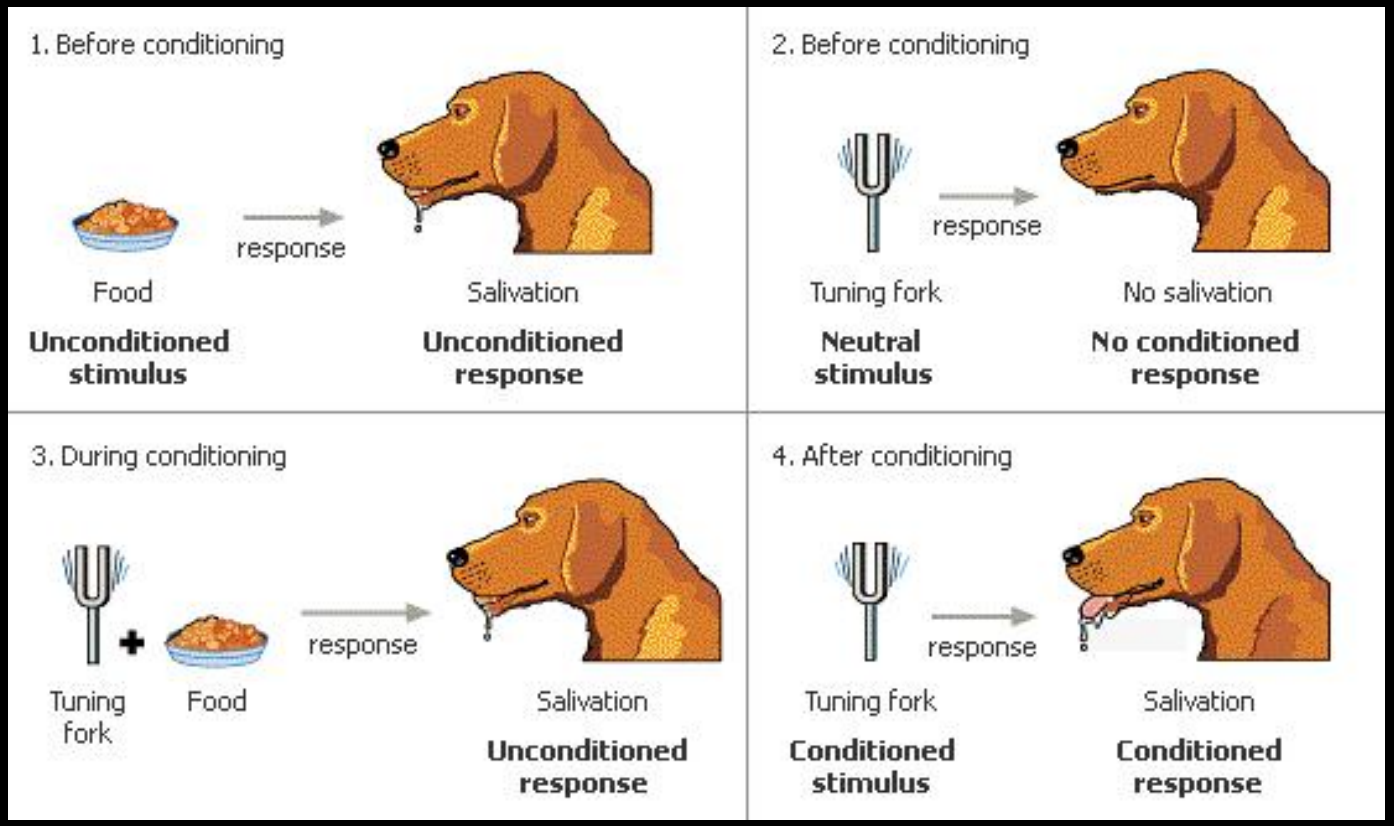

1. Before conditioning
response
Food
Salivation
Unconditioned stimulus
Unconditioned response


1. Before conditioning
response
Food
Salivation
Unconditioned stimulus
Unconditioned response
2. Before conditioning
response
Tuning fork
No salivation
Neutral stimulus
No conditioned response


1. Before conditioning
response
Food
Salivation
Unconditioned stimulus
Unconditioned response

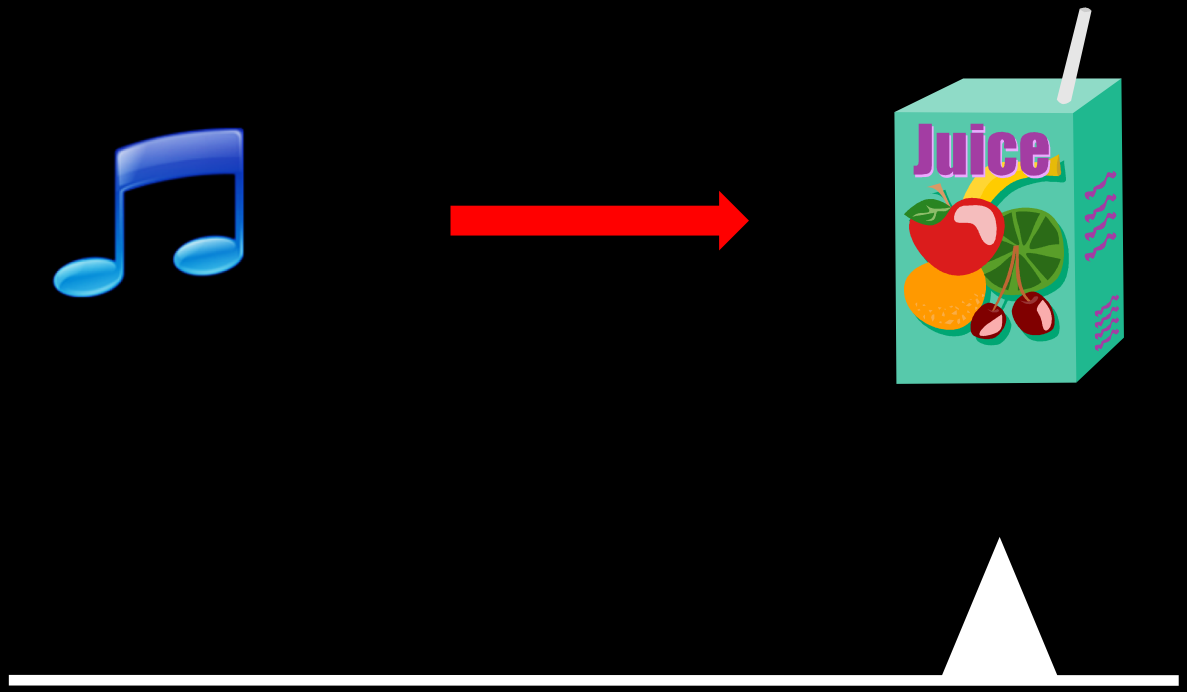

2. Before conditioning
response
Tuning fork
No salivation
Neutral stimulus
No conditioned response

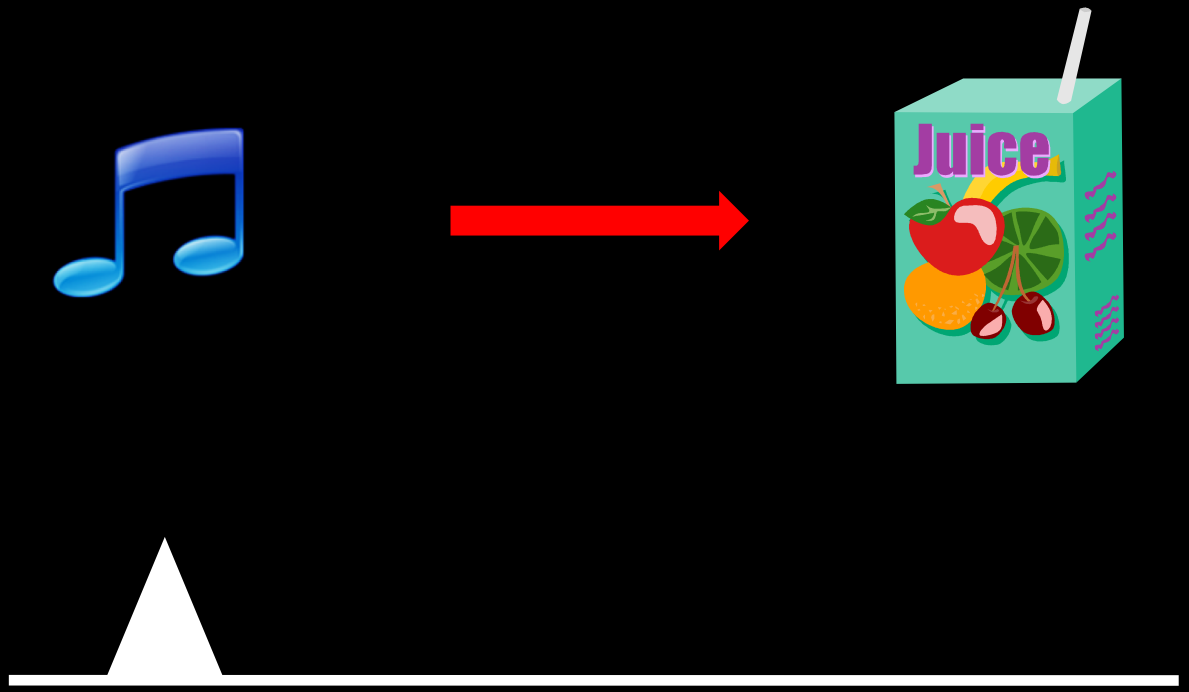

3. During conditioning
response
Tuning fork
Food
Salivation
Unconditioned response

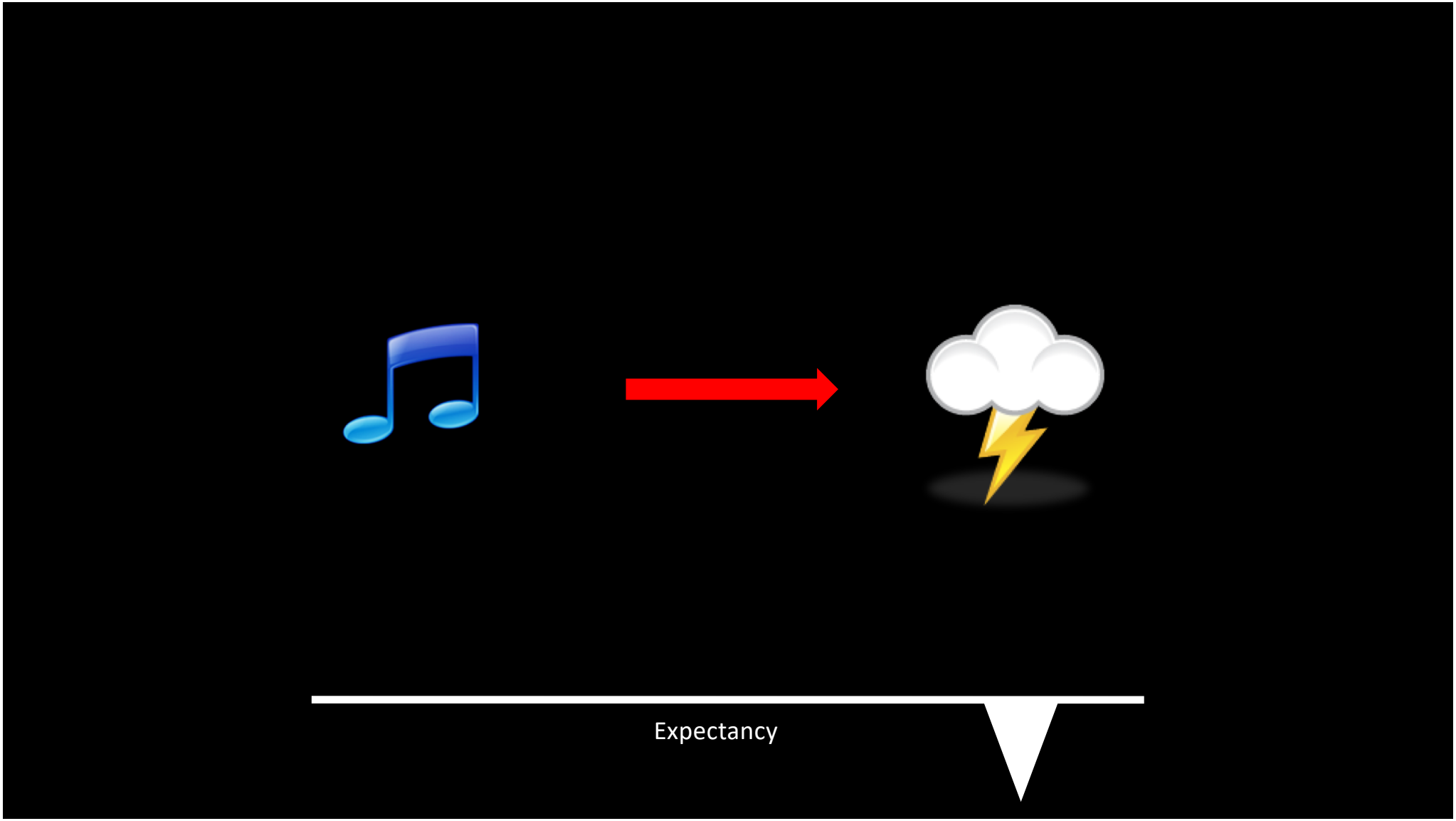

1. Before conditioning
response
Food
Salivation
Unconditioned stimulus
Unconditioned response
2. Before conditioning
response
Tuning fork
No salivation
Neutral stimulus
No conditioned response
3. During conditioning
response
Tuning fork
Food
Salivation
Unconditioned response
4. After conditioning
response
Tuning fork
Salivation
Conditioned stimulus
Conditioned response

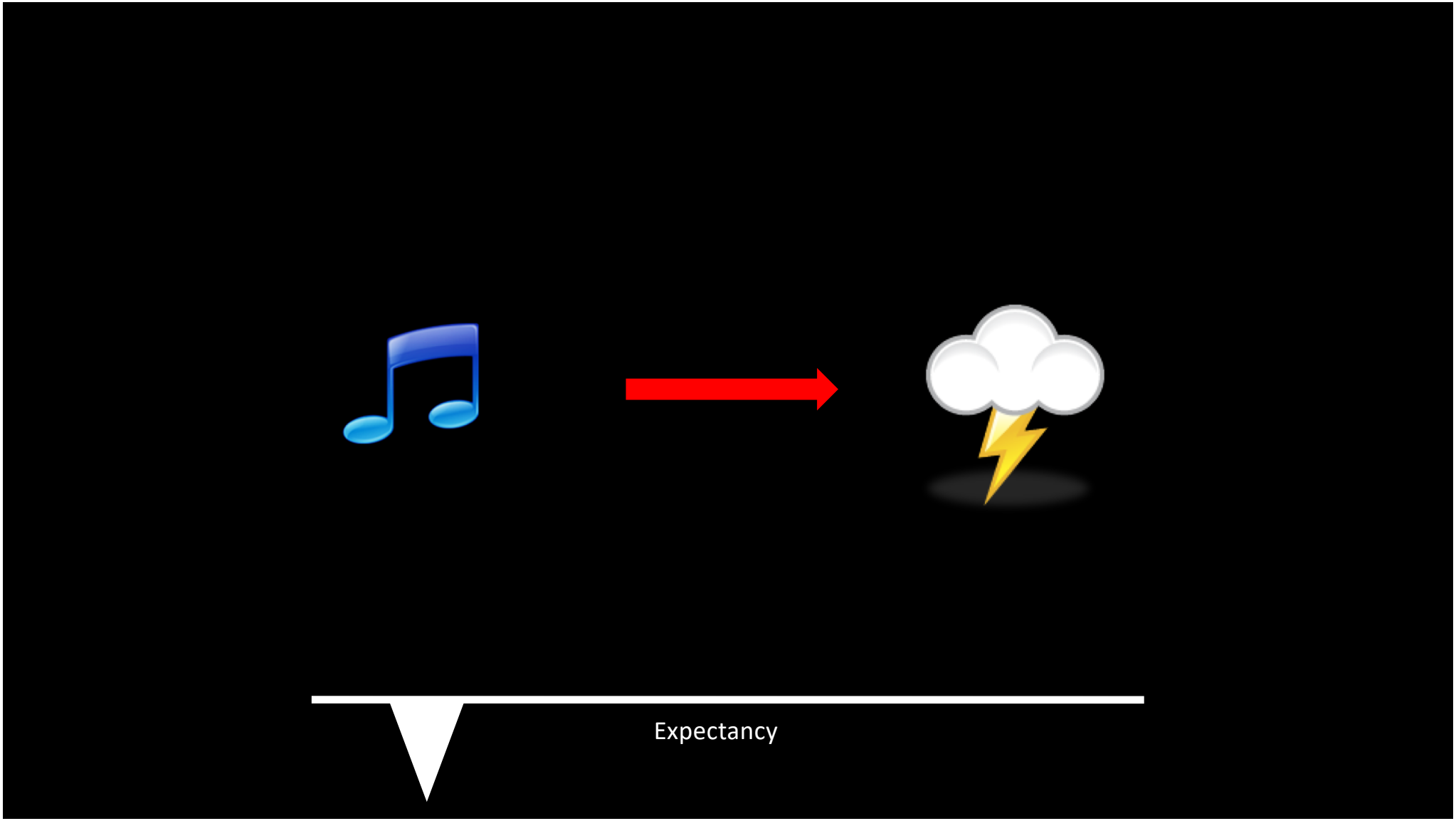
Rescorla - Wagner

Expectancy


Juice
Expectancy


Juice
Expectancy


Expectancy


Expectancy

Prediction Error

When the an outcome is different
from an expectation

Value

V_{reward}

$V_{punishment}$

V_{tone}

Prediction Error

The difference in VALUE between the expectation and the outcome

$$PE = (V_{reward} - V_{cue})$$

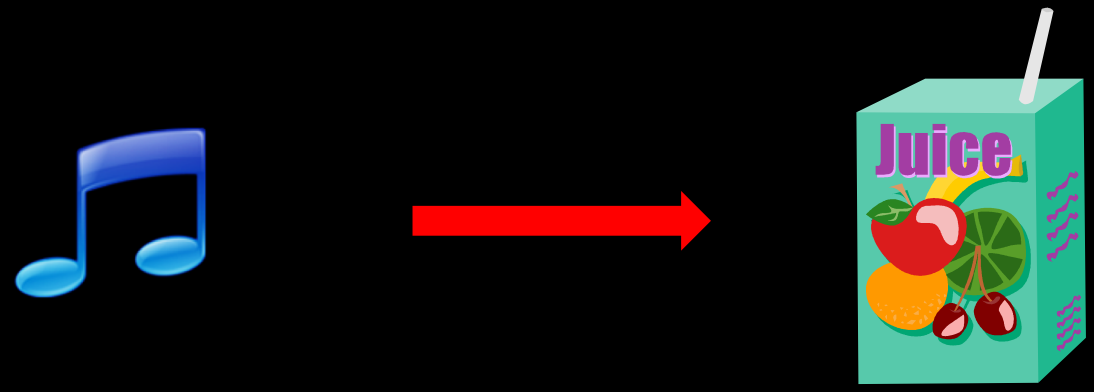
$$PE = (V_{outcome} - V_{expectation})$$

Learning IS ALWAYS a two step process.

At each point in time we:

1) Calculate a prediction error
2) Update the previous value

$$V_{cue\ new} = V_{cue\ old} + PE$$



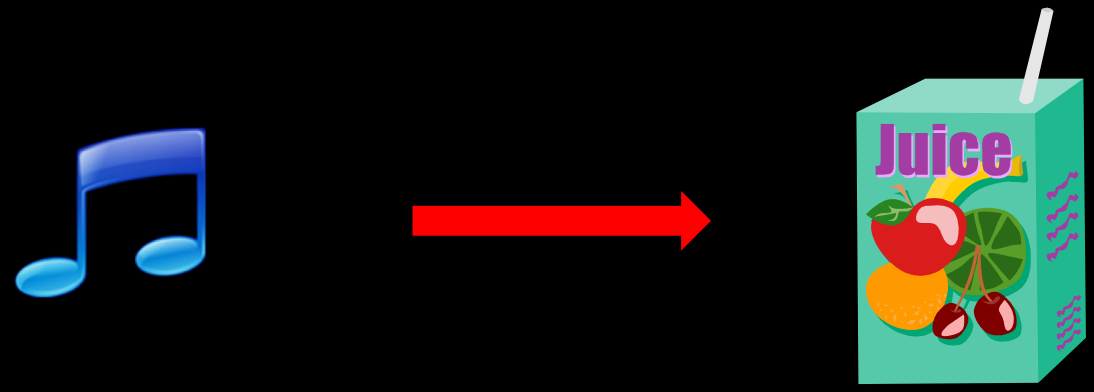
$V_{tone} = 0$ $V_{reward} = 100$

$PE = (100 - 0)$

$PE = 100$

$$V_{cue\ new} = V_{cue\ old} + PE$$

$$V_{cue\ new} = 0 + 100$$



$V_{tone} = 100$ $V_{reward} = 100$

$PE = (100 - 100)$

$PE = 0$

$$V_{cue\ new} = V_{cue\ old} + PE$$

$$V_{cue\ new} = 100 + 0$$

Learning Rates

$$V_{cue\ new} = V_{cue\ old} + PE * \mathbf{LR}$$

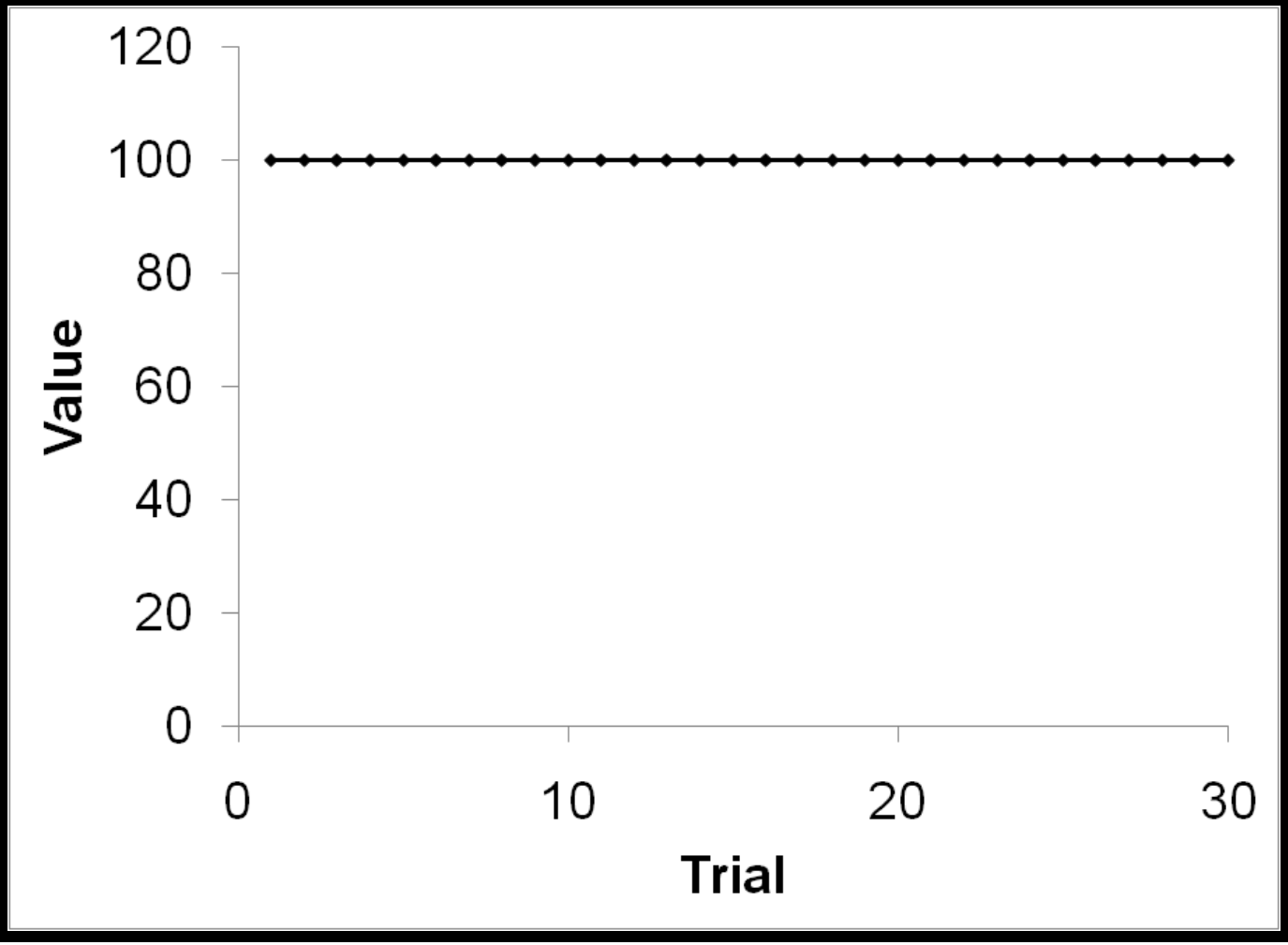
$$V_{cue\ new} = 0 + 100 * \mathbf{0.2}$$

$$V_{cue\ new} = 0 + 20$$

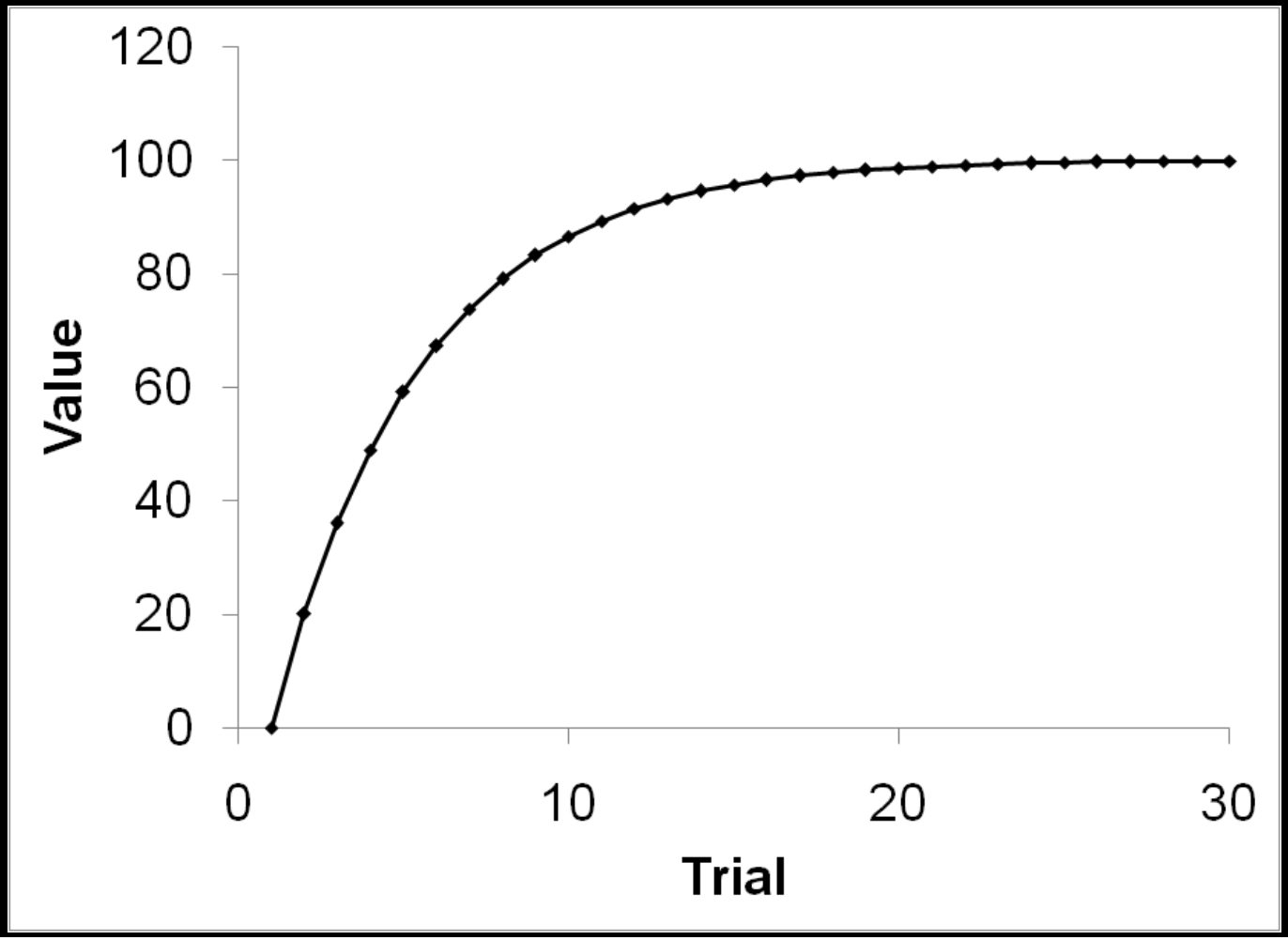
LR = 0.2

Trial	V_{cue}	V_{reward}	PE	PE x LR
1	0	100	100	20
2	20	100	80	16
3	36	100	64	12.8
4	48.8	100	51.2	10.24
5	59.04	100	40.96	8.192

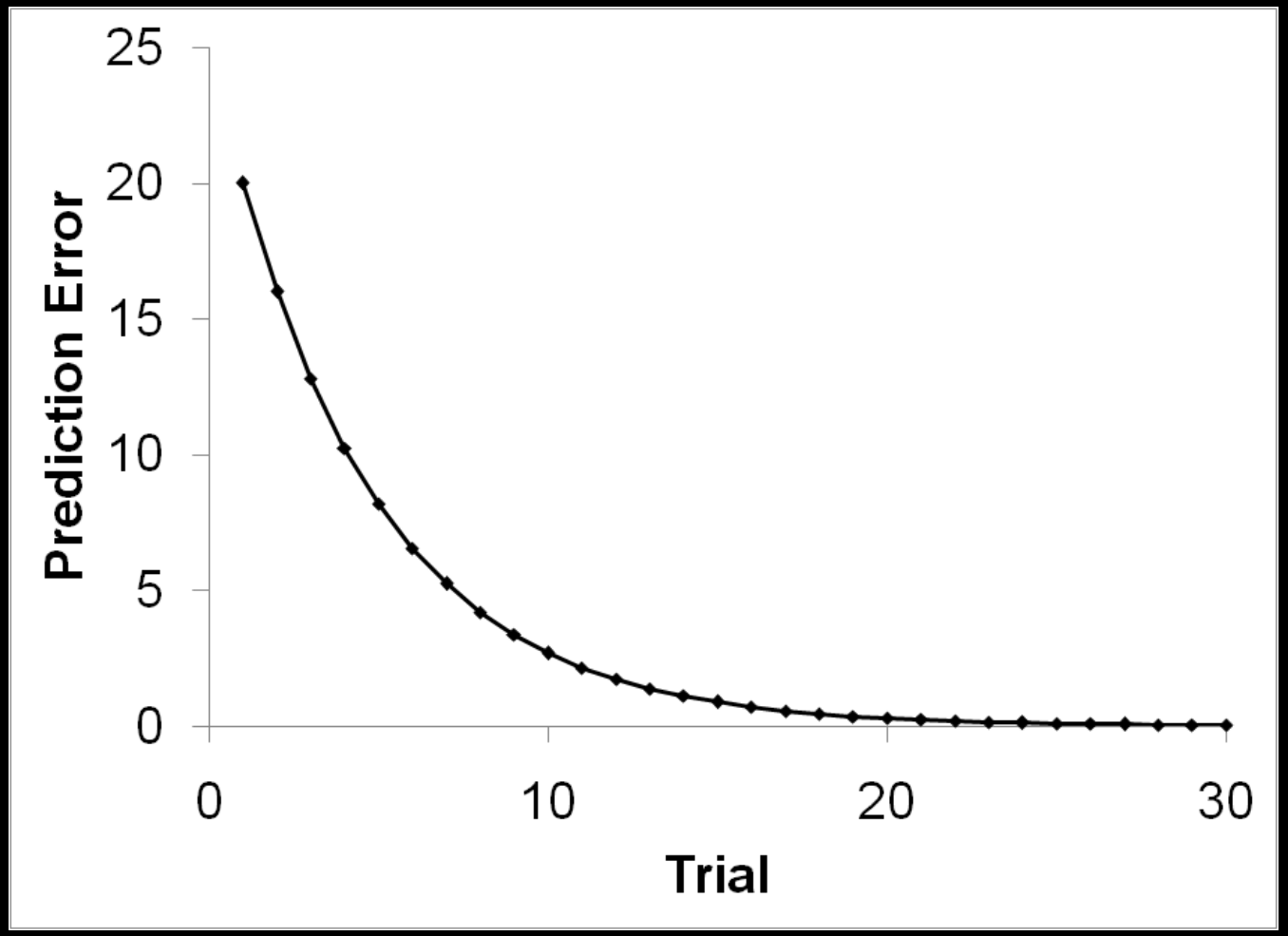
V_{reward}



V_{cue}



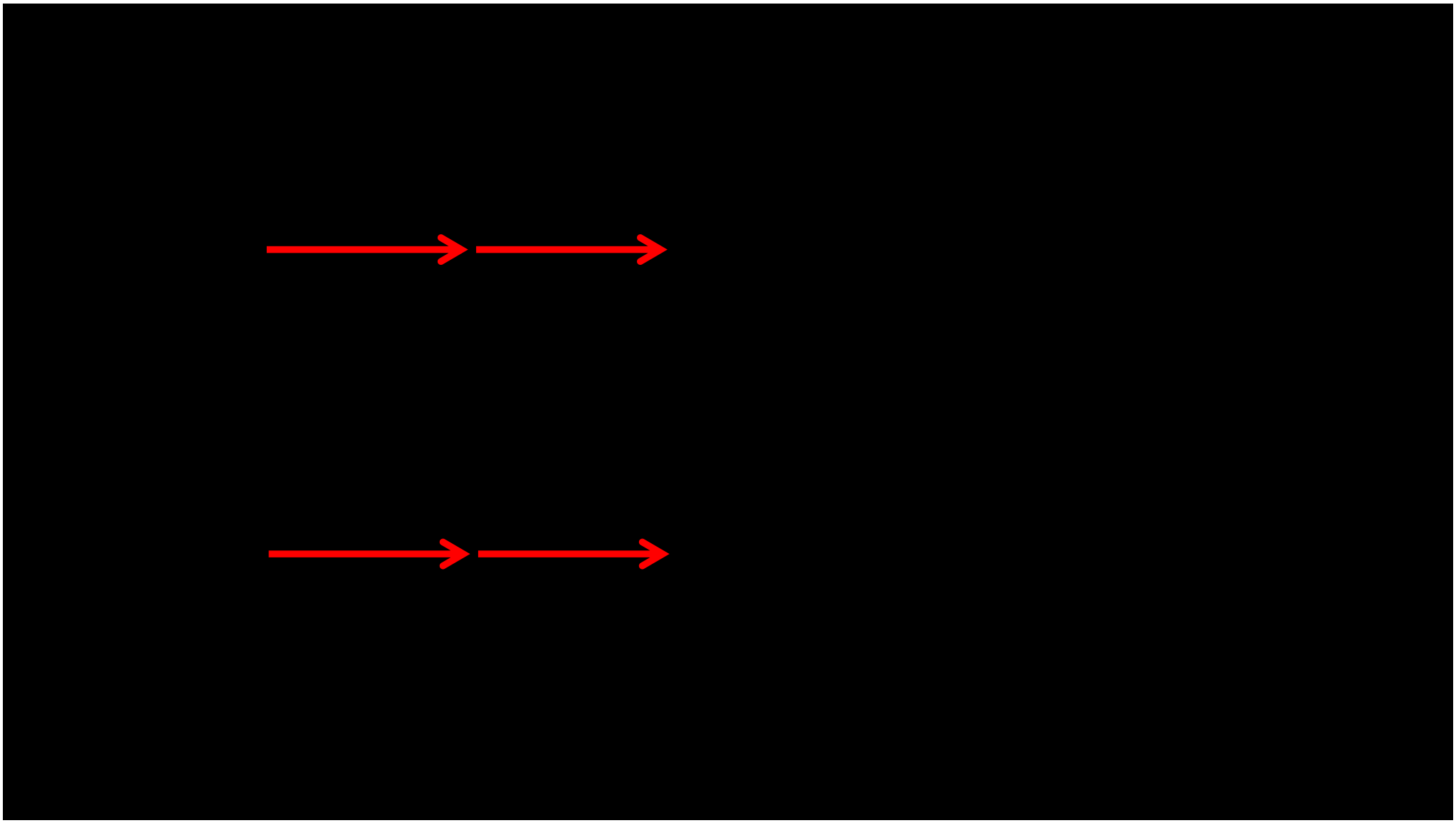
Prediction Error x Learning Rate

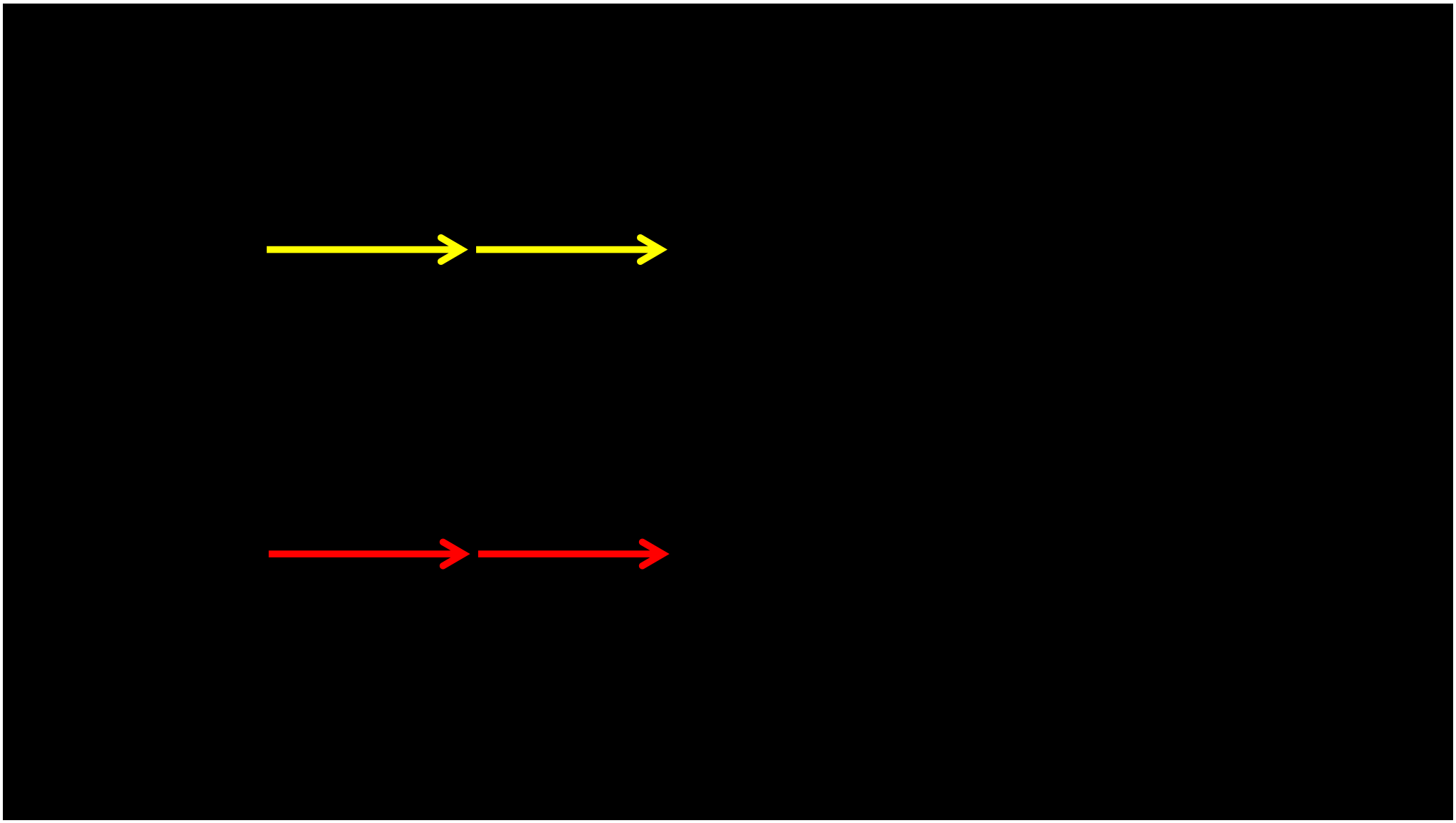


This applies to what is learned!

Recall that a memory is a collection of neurons being activated.

PE's can be used in principle to strengthen the connections between these neurons to "learn" the correct pattern.





$$PE = V_{outcome} - V_{action}$$

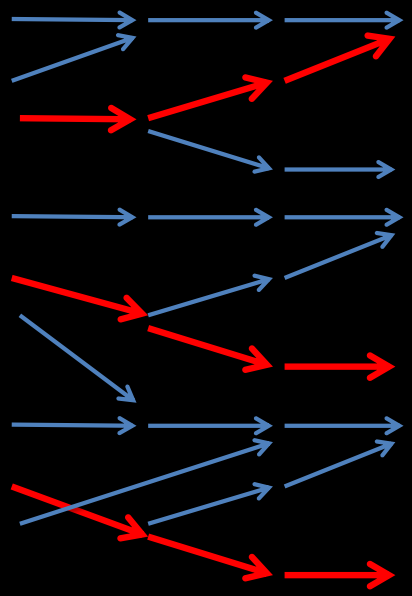
$$PE = +$$

$$V_{action} = V_{action} + PE * LR$$

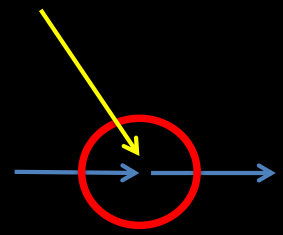
$$V_{action} \uparrow$$

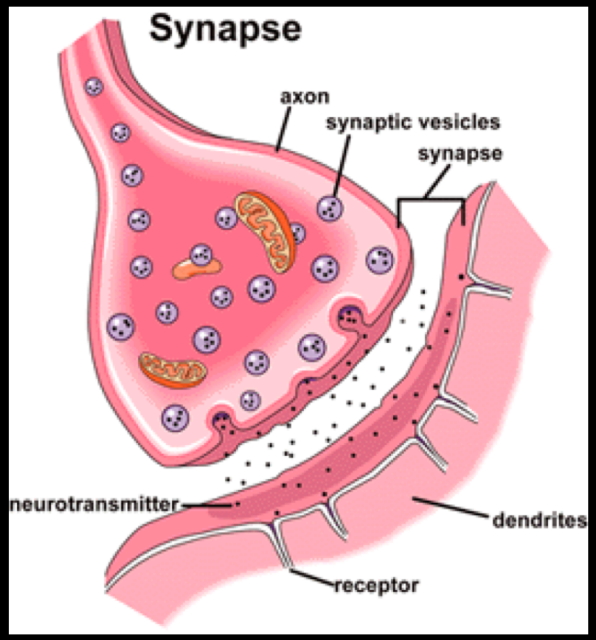
Thus, we choose/remember the top set of neurons because this choice has a higher VALUE

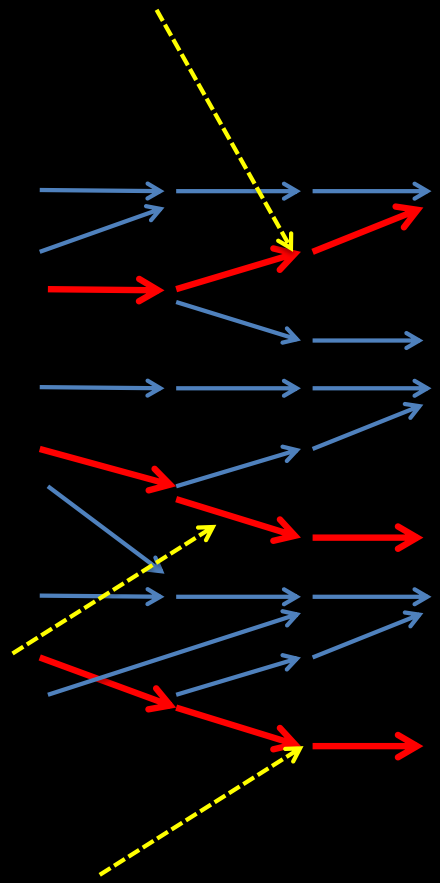
This is a basic principle of decision making – always choose the highest value option



Hebbian Learning




Synapse
axon
synaptic vesicles
synapse
neurotransmitter
dendrites
receptor



Hebbian Learning &
Reinforcement Learning

Other Types of Learning

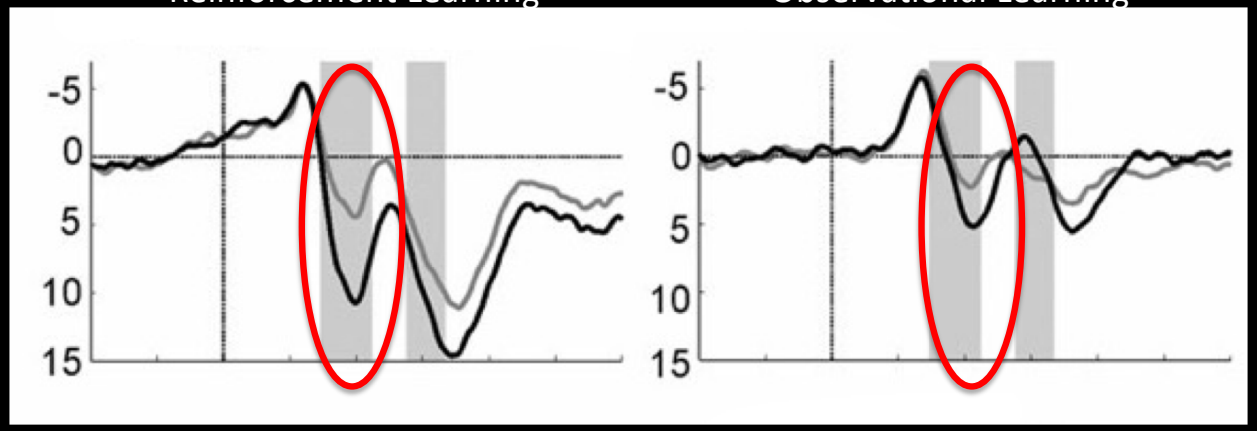
Observational Learning

What is observational learning?

- learning that occurs when we watch someone else perform a skill

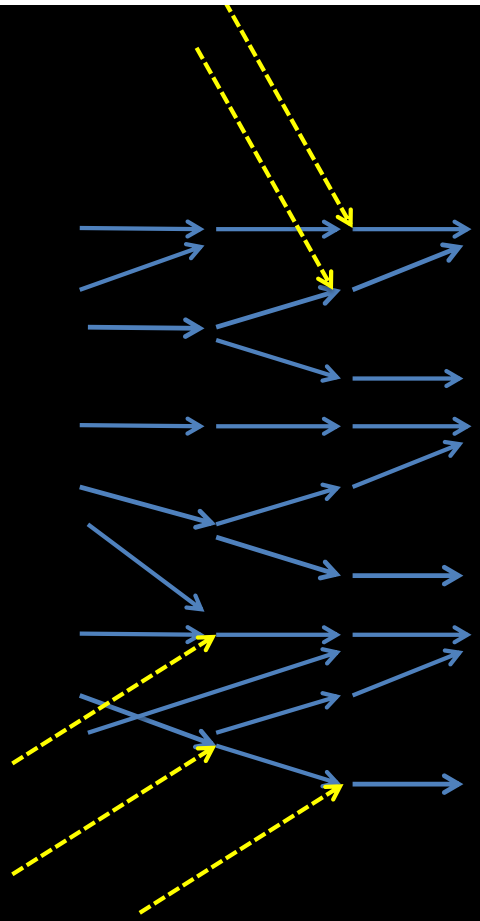
Reinforcement Learning

Observational Learning



Observational Learning

- same system as reinforcement learning
- not as active as for third person learning



Observational Learning

Supervised Learning

What is supervised learning?

- in RL feedback is binary

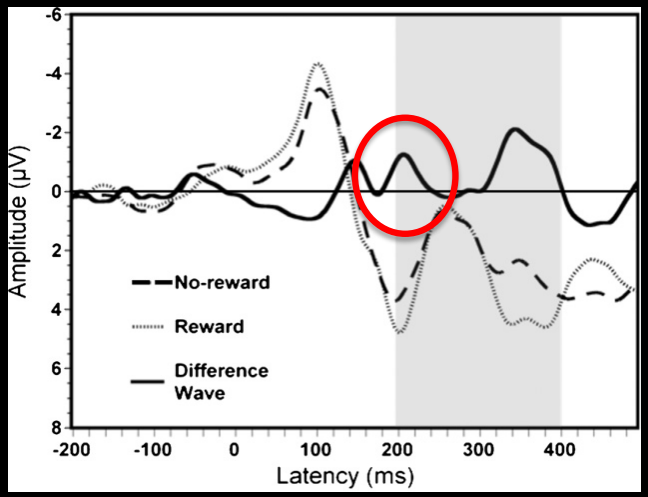
  "right", "wrong"

- In SL feedback is a vector

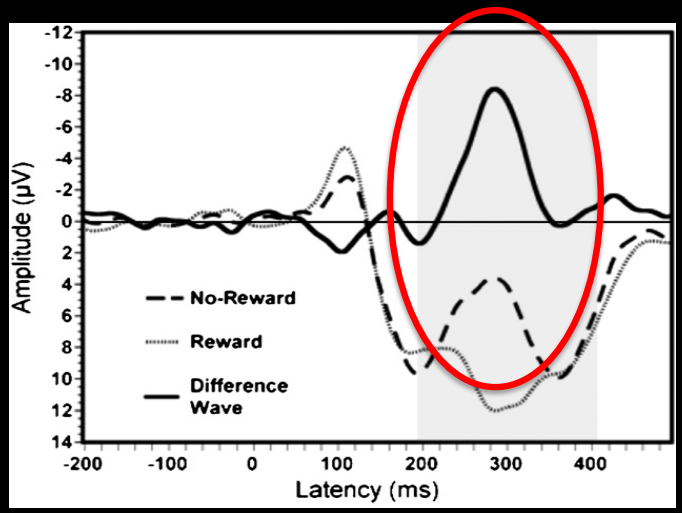
  "right, wrong" + what you did wrong

e.g., "Just tell me the answer / how to do it"

Supervised Learning

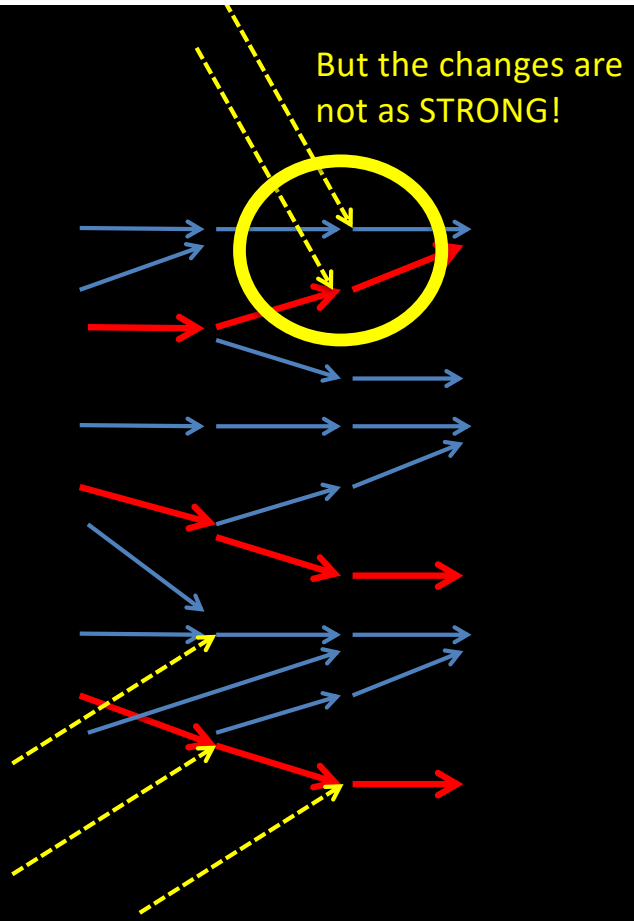


Reinforcement Learning



Supervised Learning

- the data suggests SL does not generate as large a dopaminergic response as RL
- i.e., less learning occurs on a trial to trial basis
- but, RL is a VERY slow process. in a learning environment would you have the time to use it by itself?



Hebbian Learning &
Supervised Learning

So what is the answer?

- obviously, a combination of HL, RL, OL, and SL is needed.
- before you think that is an obvious statement, how often have you experienced true RL?
- how often has impatience led to more SL than RL?
- what about SOL vs OL – do you tell people what is wrong or prompt them to figure it out?