ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA
CAMPUS DI CESENA

Contiguity, Contingency and Surprise as drivers of Reinforcement Learning

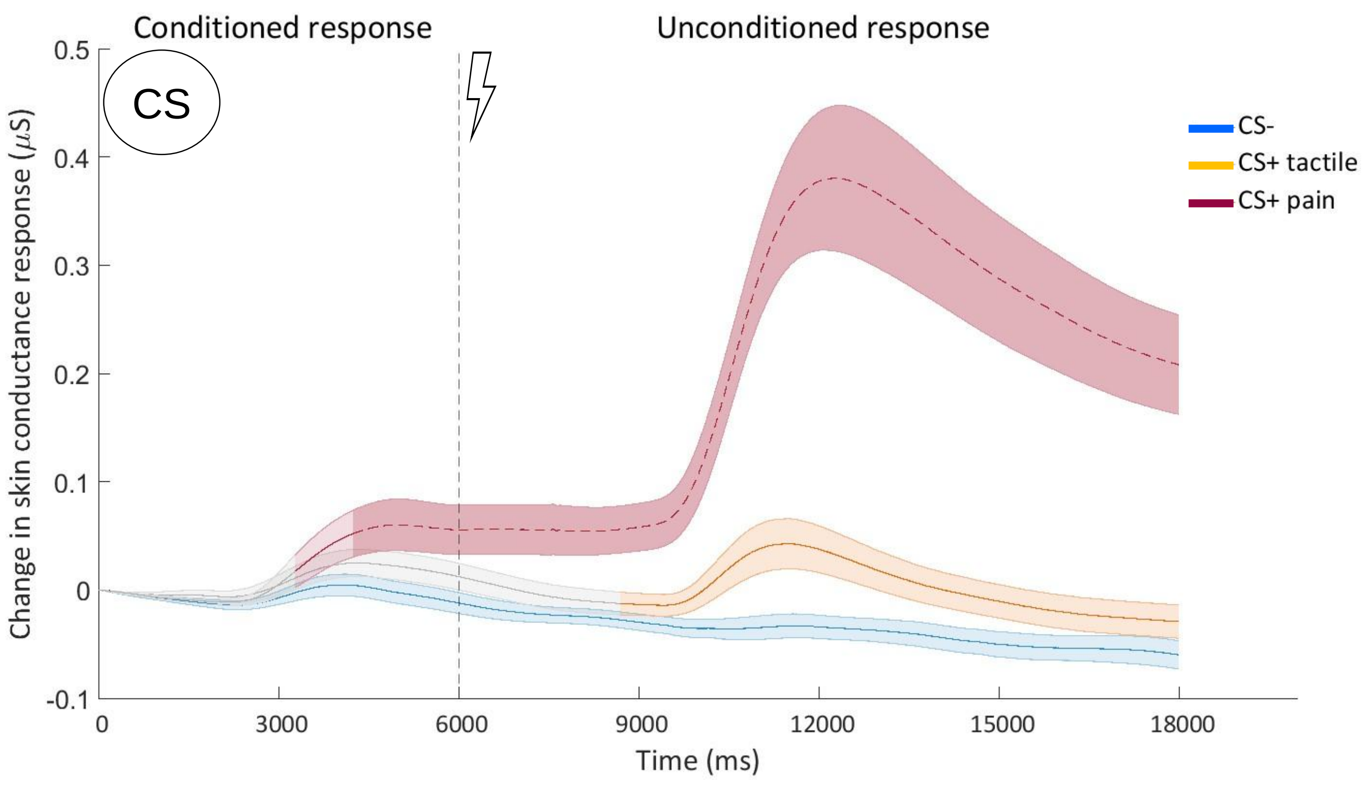
Cognition and Neuroscience
Academic year 2024/2025

Francesca Starita

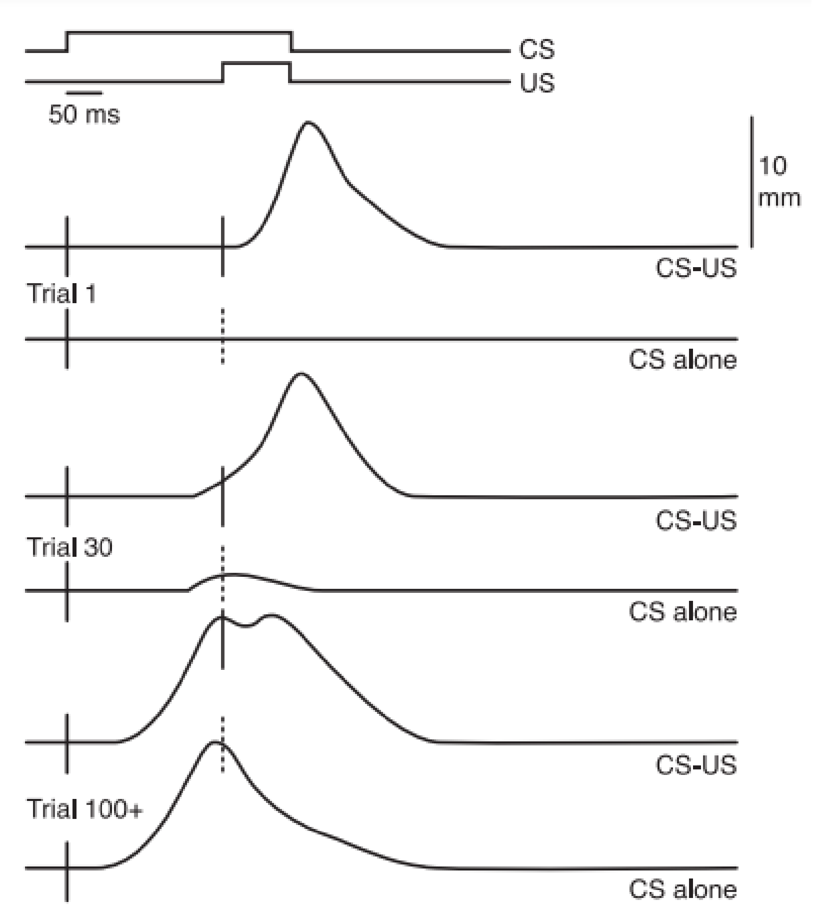
francesca.starita2@unibo.it

Learning is a predictive process!!

Conditioned responses (CRs), being initiated in anticipation of the unconditioned stimulus (US), offer better response for survival than simply initiating CRs as a reaction to the US.



development of the conditioned eyelid response



Sutton, R. S., & Barto, A. G. (2018). Reinforcement learning: An introduction. MIT press.

BUT…

What relation must exist between conditioned stimuli and unconditioned stimuli (outcomes) for reinforcement learning to take place and conditioned responses to occur?

Learning objectives

- Explain the concept of contiguity
- Distinguish between delay and trace conditioning and describe the experimental evidence on the effect of the two on learning
- Define contingency and explain its importance in reinforcement learning
- Describe Rescorla's truly random control procedure and how it demonstrated that contiguity alone is not sufficient for conditioning.
- Understand the concept of surprise in learning and that we primarily learn from surprising or unpredicted events.
- Explain the blocking effect and its implications for understanding that learning is driven by surprise
- Define prediction error and explain how prediction errors drive learning

Learning objectives

- Describe the Rescorla-Wagner model as a computational model that formalizes learning through error-driven updating of the expected value of a CS.
- Identify the limitations of the Rescorla-Wagner model, such as its trial-level nature and its failure to explain second-order conditioning.
- Describe the Temporal Difference (TD) model as a real-time computational model that considers time steps within a trial.
- Explain how the TD model generates predictions and prediction errors by comparing the predicted value at consecutive time steps.
- Understand how the TD model addresses some of the limitations of the Rescorla-Wagner model by incorporating temporal information.
- Relate the principles of predictive learning to sensory perception, understanding that our brains actively construct our sensory experience based on predictions.

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA
CAMPUS DI CESENA

Early theories of Pavlovian conditioning took for granted that CS–US contiguity was the critical determinant of reinforcement learning

Contiguity: closeness in time between stimulus/behavior and outcome.

- Stimuli that are close to one another in time become associated.

Early theories of Pavlovian conditioning took for granted that CS–US contiguity was the critical determinant of reinforcement learning

Contiguity: closeness in time between stimulus/behavior and outcome.

- Stimuli that are close to one another in time become associated.

How would you test this?

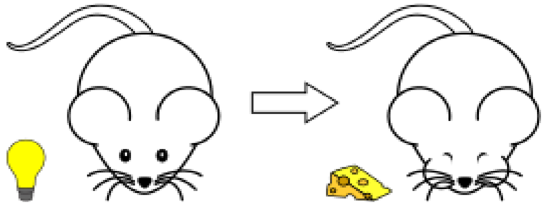
7

Is CS-US contiguity necessary for learning to occur?

- Do events that are close to one another in time become associated?

How would you test this?

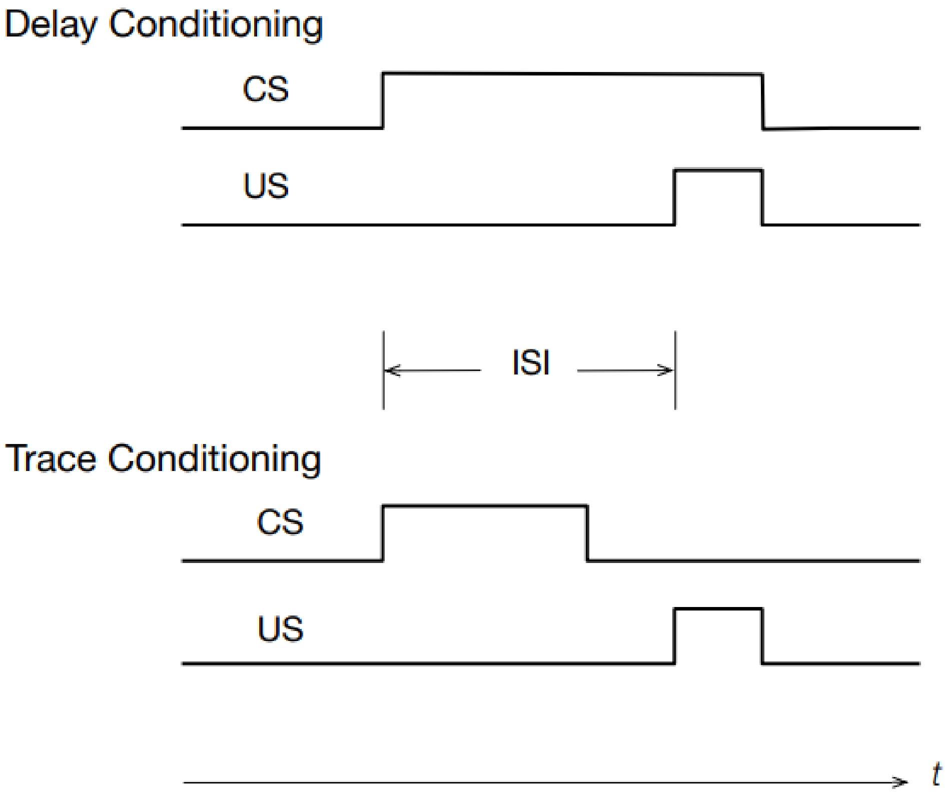
Is CS-US contiguity necessary for learning to occur?



We can compare learning under:

Delay conditioning

Trace conditioning



ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA
CAMPUS DI CESENA

Is CS-US contiguity necessary for learning to occur?

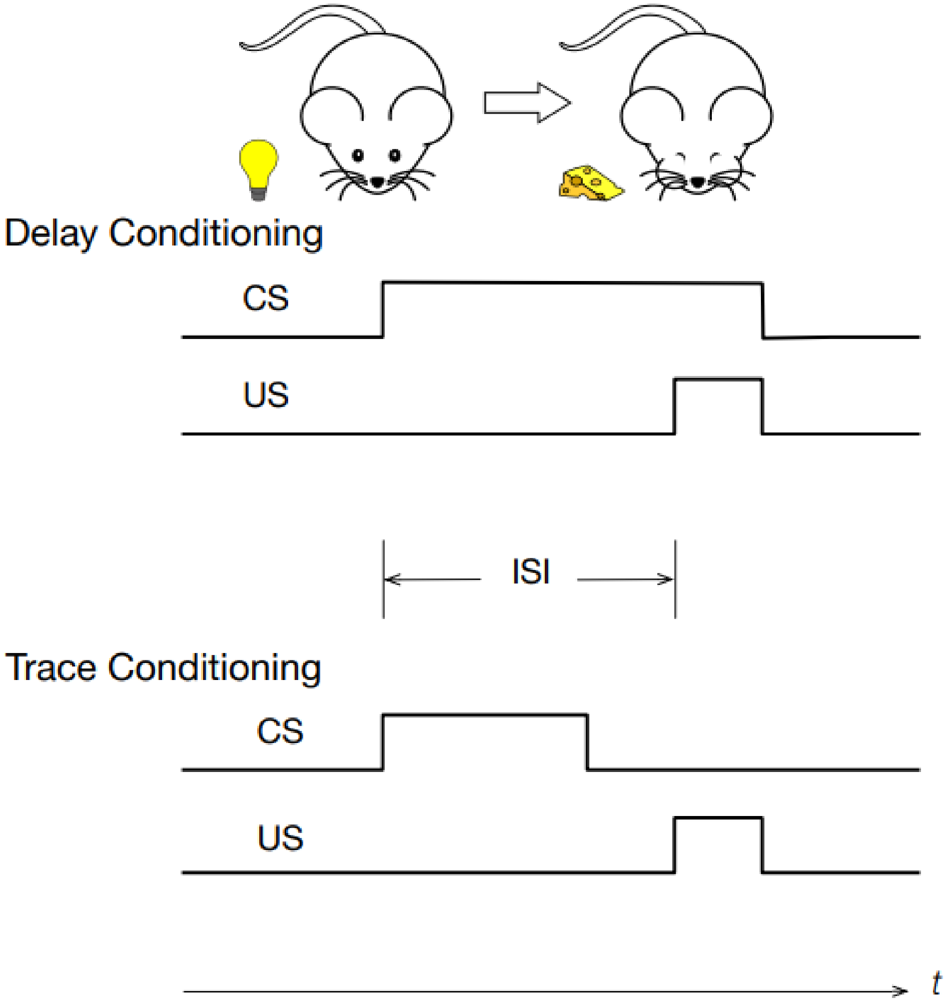
We can compare learning under:

Delay conditioning

- the CS extends throughout the interstimulus interval (ISI), which is the time interval between the CS onset and the US onset (often with the CS and US ending at the same time as shown here).

Trace conditioning

- there is a time interval, called the trace interval, between CS offset and US onset.



ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA
CAMPUS DI CESENA

Is CS-US contiguity necessary for learning to occur?

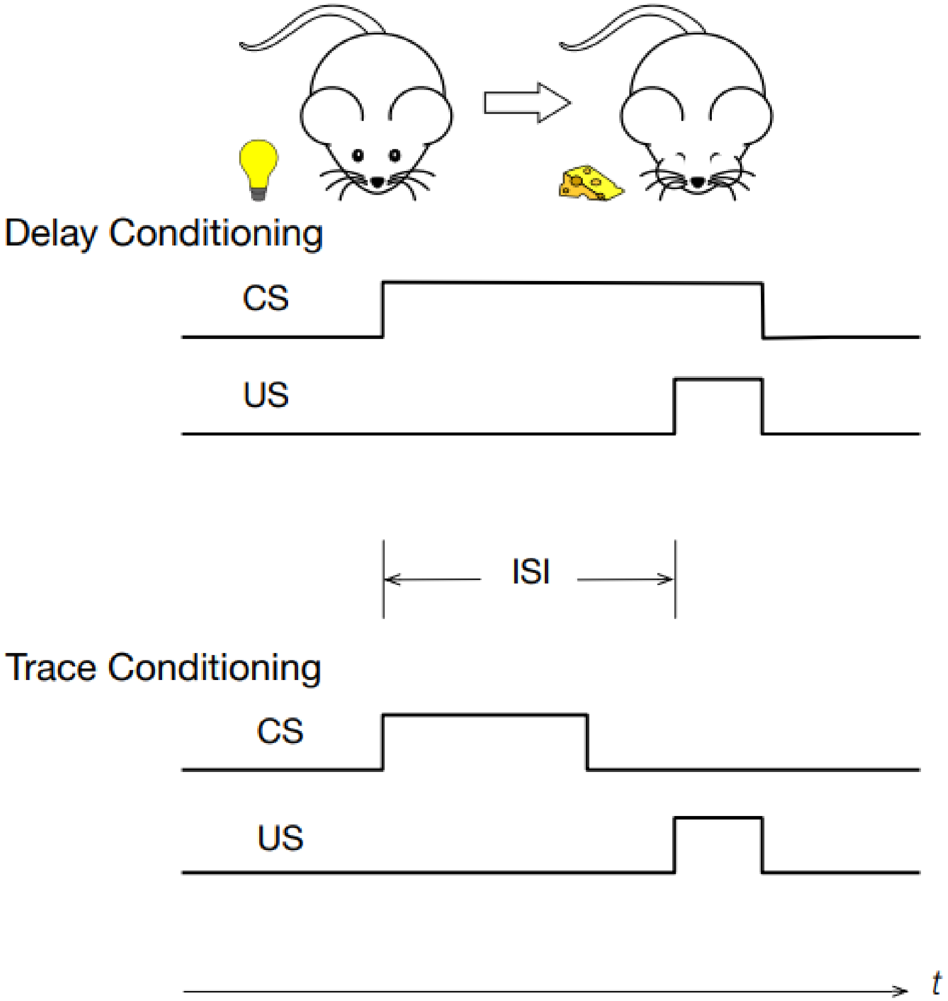
We can compare learning under:

Delay conditioning

- the CS extends throughout the interstimulus interval (ISI), which is the time interval between the CS onset and the US onset (often with the CS and US ending at the same time as shown here).

Trace conditioning

- there is a time interval, called the trace interval, between CS offset and US onset.



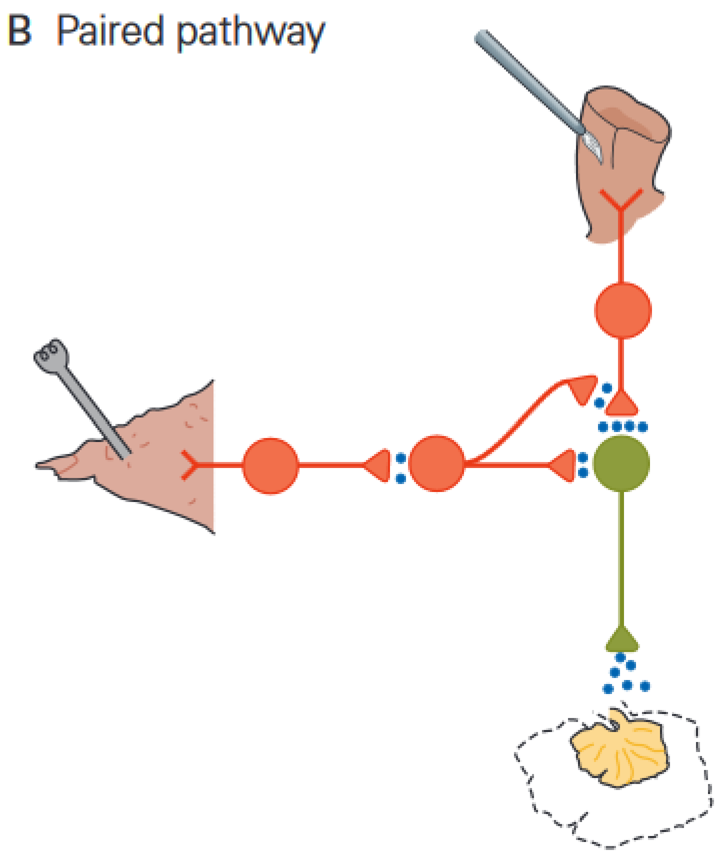
Experiments show that CR is more readily acquired under delay rather than trace conditioning.

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA
CAMPUS DI CESENA

Hebbian plasticity is an example of how contiguity enables learning in the nervous system

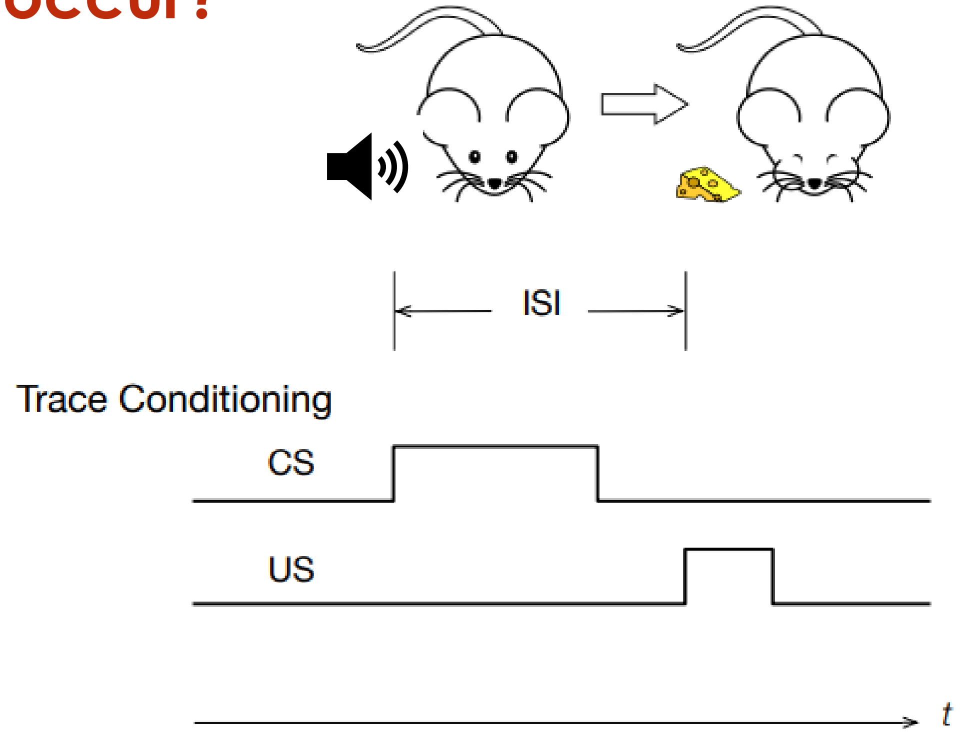
Hebb's Law: **Neurons that fire together, wire together**

All that is needed to condition the gill withdrawal reflex is the simultaneous activation of the syphon and tail neuronal pathways



ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA
CAMPUS DI CESENA

Is CS-US contiguity necessary for learning to occur?

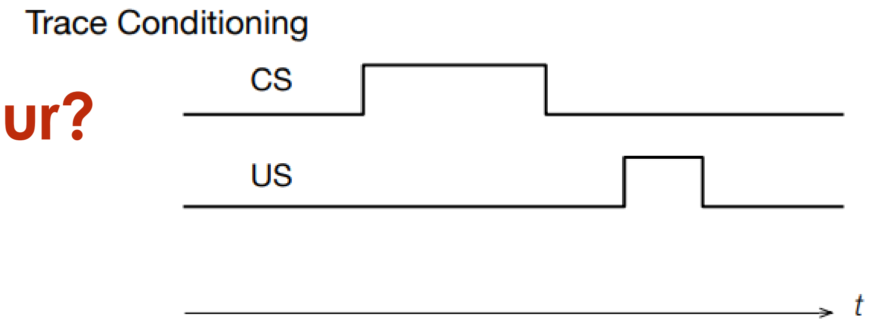


Trace conditioning

- there is a time interval, called the trace interval, between CS offset and US onset.
- **Learning does not occur across long trace intervals BUT learning can still occur for "no too long" intervals**
- Assumes that <u>a trace of the CS remains</u> when the US arrives. So, learning occurs through the simultaneous presence of the trace and the US

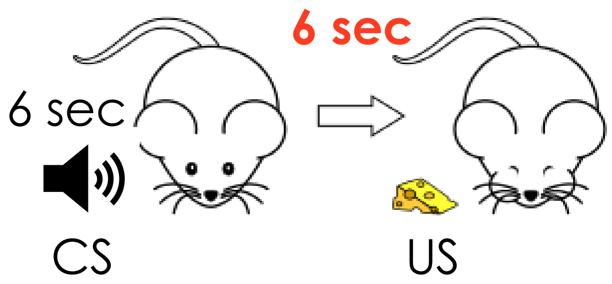
How would you test this?

Is CS-US contiguity necessary for learning to occur?



Hypothesis: When this CS-US interval is lengthened, a decrement in conditioning is observed. Meaning that it takes more trials for the conditioned response (CR) to appear, and/or CR strength is reduced

Group 1



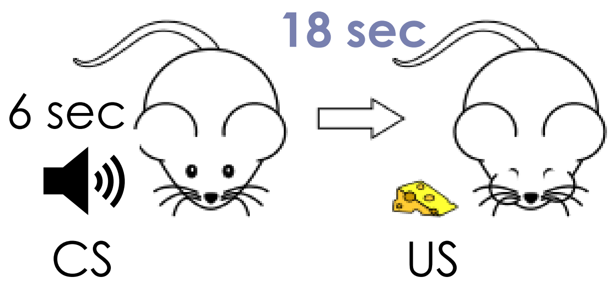
Experiment

2 groups of rats were exposed to a 6 s tone CS, followed by pellet delivery after:

- **Group 1:** 6s (red data)
- **Group 2:** 18s (purple data)

Anticipatory head entries into the feeding hopper were recorded

Group 2

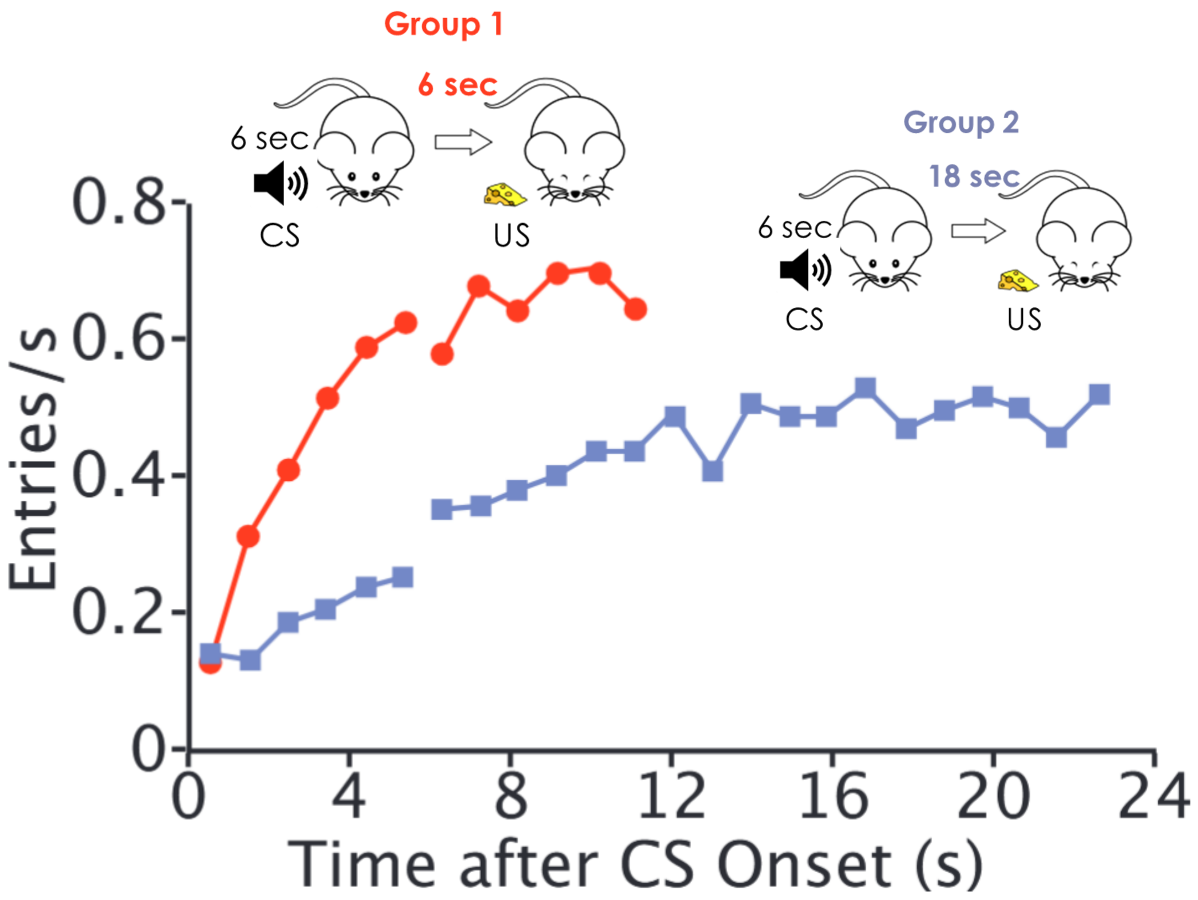


Balsam, P. D., Drew, M. R., & Gallistel, C. R. (2010). Time and Associative Learning. *Comparative cognition & behavior reviews*, *5*, 1–22. https://doi.org/10.3819/ccbr.2010.50001

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA
CAMPUS DI CESENA

Is CS-US contiguity necessary for learning to occur?

Results



Balsam, P. D., Drew, M. R., & Gallistel, C. R. (2010). Time and Associative Learning. *Comparative cognition & behavior reviews*, *5*, 1–22. https://doi.org/10.3819/ccbr.2010.50001

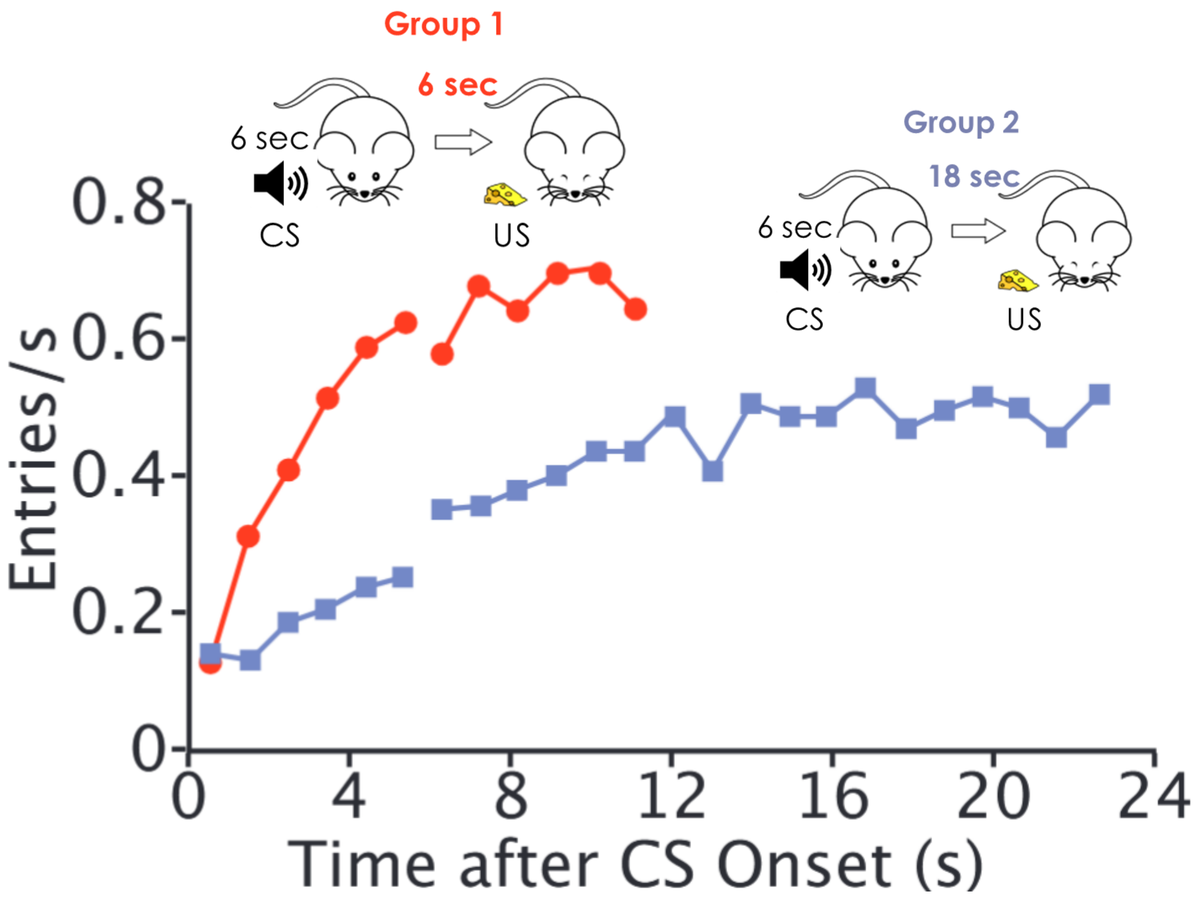
Is CS-US contiguity necessary for learning to occur?

Results

- Both groups increase their rate of entries as time of US delivery approaches.

- BUT:
  - **Group 1** shows greater rate of entries during the entire course of CS presentation and trace interval
  - **Group 2** never reaches the rate of entries of Group 1, even at the end of its trace interval

- Longer trace interval results in weaker CR



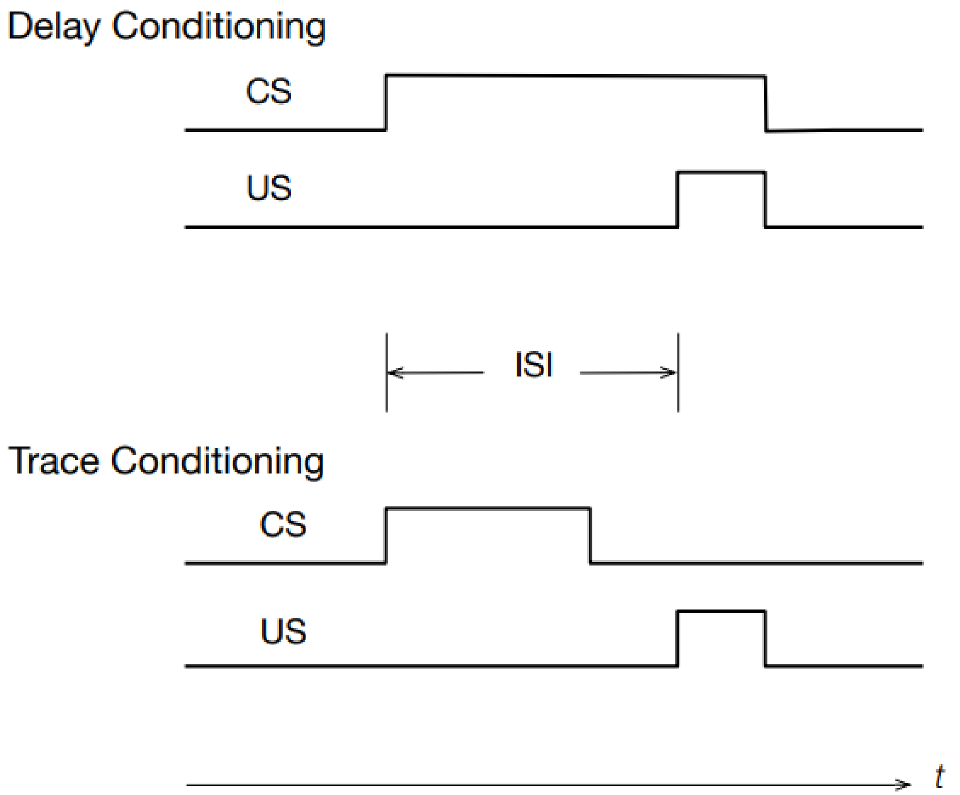
Balsam, P. D., Drew, M. R., & Gallistel, C. R. (2010). Time and Associative Learning. *Comparative cognition & behavior reviews*, *5*, 1–22. https://doi.org/10.3819/ccbr.2010.50001

Early theories of Pavlovian conditioning took for granted that CS–US contiguity was the critical determinant of learning

Contiguity is indeed important for learning:

only close presentations in time of the CS and US allow for the acquisition of the conditioned response

If a long trace interval is introduced, then the CS representation will have decayed by the time the US is presented and thus the CS will not be learned about. (McNally & Westbrook, 2006)



ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA
CAMPUS DI CESENA

18

Early theories of Pavlovian conditioning took for granted that CS–US contiguity was the critical determinant of learning

Contiguity: closeness in time between stimulus/behavior and outcome.

- Stimuli that are close to one another in time become associated.

19

Early theories of Pavlovian conditioning took for granted that CS–US contiguity was the critical determinant of learning

Contiguity: closeness in time between stimulus/behavior and outcome.

- Stimuli that are close to one another in time become associated.

BUT: Is contiguity sufficient to learn?

Challenging the contiguity assumption: The truly random control procedure

In the 1960's and 70's, evidence began to accumulate that posed a challenge to the simple contiguity assumption.

The truly random control procedure (Rescorla, 1968) showed that **contiguity** between a conditioned stimulus (CS) and an unconditioned stimulus (US) **is not sufficient to lead to learning**

Journal of Comparative and Physiological Psychology
1968, Vol. 66, No. 1, 1–5

PROBABILITY OF SHOCK IN THE PRESENCE AND ABSENCE OF CS IN FEAR CONDITIONING[1]

ROBERT A. RESCORLA

Yale University

(1940 - 2020)

2 experiments indicate that CS-US contingency is an important determinant of fear conditioning and that presentation of US in the absence of CS interferes with fear conditioning. In Experiment 1, equal probability of a shock US in the presence and absence of a tone CS produced no CER suppression to CS; the same probability of US given only during CS produced substantial conditioning. In Experiment 2, which explored 4 different probabilities of US in the presence and absence of CS, amount of conditioning was higher the greater the probability of US during CS and was lower the greater the probability of US in the absence of CS; when the 2 probabilities were equal, no conditioning resulted.

Two conceptions of Pavlovian conditioning have been distinguished by Rescorla (1967). The first, and more traditional, notion emphasizes the role of the number of pairings of CS and US in the formation of a CR. The second notion suggests that it is the contingency between CS and US which is important. The notion of contingency differs from that of pairing in that it includes not only what events are paired but also what events are not paired. As used here, contingency refers to the relative probability of occurrence of US in the presence of CS as contrasted with its probability in the absence of CS. The contingency notion suggests that, in fact, conditioning only occurs when these probabilities differ; when the probability of US is higher during CS than at other times, excitatory conditioning occurs; when the probability is lower, inhibitory conditioning results. Notice that the probability of a US can be the same in the absence and presence of CS and yet there can be a fair number of CS-US pairings. It is this that makes it possible to assess the relative importance of pairing and contingency in the development of a CR.

CONTIGUITY → CS-US pairing is crucial for conditioning

CONTINGENCY: p(US/CS) vs p(US/no-CS)

Conditioning occurs only if
p(US/CS) > p(US/no-CS)

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA
CAMPUS DI CESENA

Methods

Experimental phases

1. TRAINING
- Train rats to press lever to get food

2. PAVLOVIAN FEAR CONDITIONING

CS: tone

US: shock

3. TEST
- Measure the number of lever presses in presence vs absence of CS
- No shocks delivered --> the test is conducted under extinction
- General HP: if conditioning happened in phase 2, the presence of the CS should suppress lever pressing

Methods

Experimental phases

1. TRAINING
- Train rats to press lever to get food

2. PAVLOVIAN FEAR CONDITIONING

CS: tone

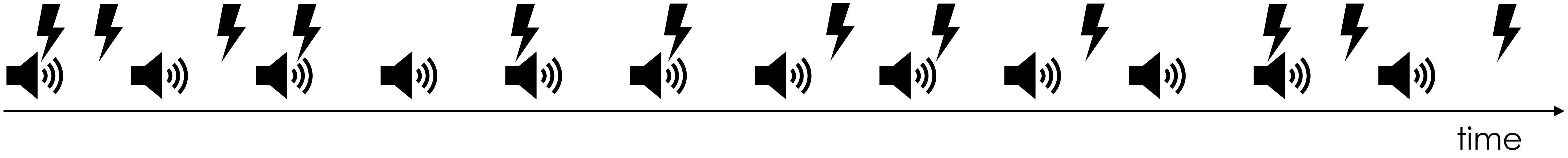
US: shock

3. TEST
- Measure the number of lever presses in presence vs absence of CS
- No shocks delivered --> the test is conducted under extinction
- General HP: if conditioning happened in phase 2, the presence of the CS should suppress lever pressing

Experimental groups

For eight *S*s in Group R-1 (random), 12 2-min. tone CSs were given in each session with a mean intertone interval of 8 min. with 12 .5-sec. .9-ma. electric shocks programmed randomly throughout the session; shocks were programmed independently of the tones in such a way that shock was equiprobable at any time within the session. The eight *S*s in Group G (gated) received a treatment identical to that of Group R-1 except that all shocks which would have occurred in the absence of the tone CS were simply omitted. This means that on the average *S*s in Group G received the same number of shocks during CS as *S*s in Group R-1, but *S*s in Group G received no shocks in the absence of CS and therefore fewer total shocks. The eight *S*s in Group R-2 received a treatment identical to those of Group R-1 except that the average number of shocks they received (2.4 per session) was the same as that for Group G. These shocks were programmed randomly throughout the session, independently of the tones.

Group R-1: 12 CS; 12 US randomly programmed, such that US equiprobalbe at anytime



Methods

Experimental phases

1. TRAINING
- Train rats to press lever to get food

2. PAVLOVIAN FEAR CONDITIONING

CS: tone

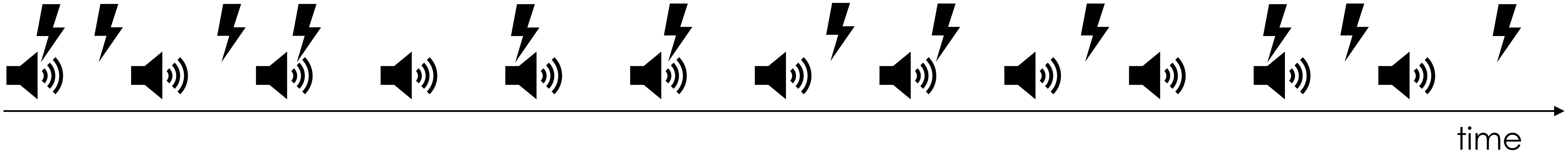
US: shock

3. TEST
- Measure the number of lever presses in presence vs absence of CS
- No shocks delivered --> the test is conducted under extinction
- General HP: if conditioning happened in phase 2, the presence of the CS should suppress lever pressing

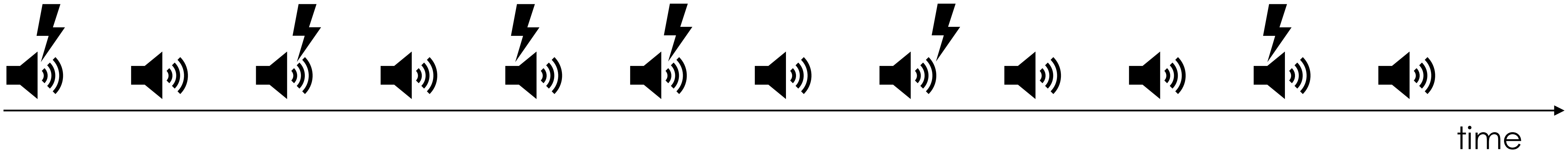
Experimental groups

For eight *Ss* in Group R-1 (random), 12 2-min. tone CSs were given in each session with a mean intertone interval of 8 min. with 12 .5-sec. .9-ma. electric shocks programmed randomly throughout the session; shocks were programmed independently of the tones in such a way that shock was equiprobable at any time within the session. The eight *Ss* in Group G (gated) received a treatment identical to that of Group R-1 except that all shocks which would have occurred in the absence of the tone CS were simply omitted. This means that on the average *Ss* in Group G received the same number of shocks during CS as *Ss* in Group R-1, but *Ss* in Group G received no shocks in the absence of CS and therefore fewer total shocks. The eight *Ss* in Group R-2 received a treatment identical to those of Group R-1 except that the average number of shocks they received (2.4 per session) was the same as that for Group G. These shocks were programmed randomly throughout the session, independently of the tones.

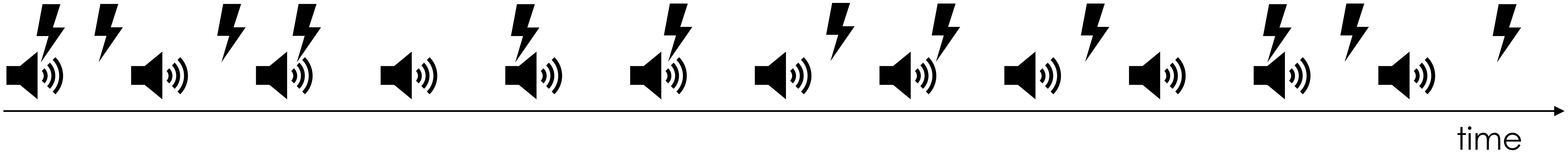
Group R-1: 12 CS; 12 US randomly programmed, such that US equiprobalbe at anytime



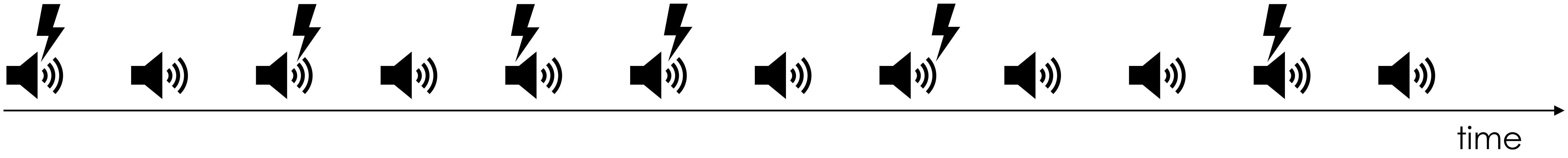
Group G: same number of CS-US pairings as Group R-1, BUT unpaired USs were removed. US occurrence is gated to CS occurrence. No US occurring in absence of CS.



Group R-1: 12 CS; 12 US randomly programmed, such that US equiprobalbe at anytime



Group G: same number of CS-US pairings as Group R-1, BUT unpaired USs were removed. US occurrence is gated to CS occurrence. No US occurring in absence of CS.



Thus, Group R-1 & Group G have the same CS-US contiguity!

Methods

Experimental phases

1. TRAINING
- Train rats to press lever to get food

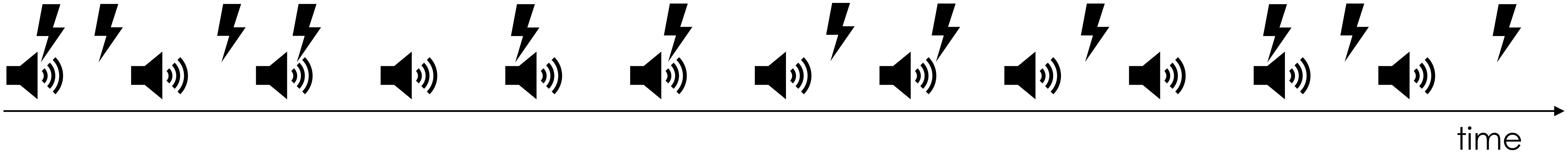
2. PAVLOVIAN FEAR CONDITIONING
- CS: tone
- US: shock

3. TEST
- Measure the number of lever presses in presence vs absence of CS
- No shocks delivered --> the test is conducted under extinction
- General HP: if conditioning happened in phase 2, the presence of the CS should suppress lever pressing

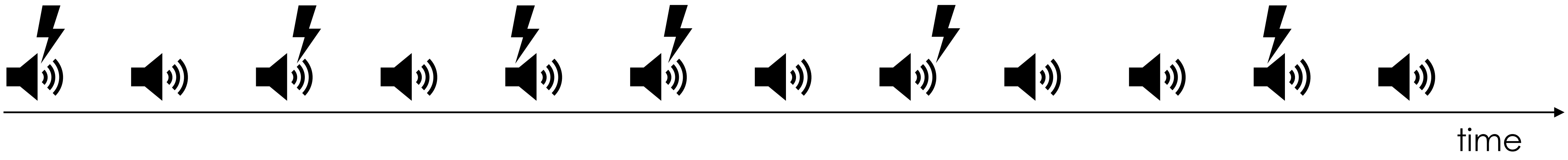
Experimental groups

For eight *Ss* in Group R-1 (random), 12 2-min. tone CSs were given in each session with a mean intertone interval of 8 min. with 12 .5-sec. .9-ma. electric shocks programmed randomly throughout the session; shocks were programmed independently of the tones in such a way that shock was equiprobable at any time within the session. The eight *Ss* in Group G (gated) received a treatment identical to that of Group R-1 except that all shocks which would have occurred in the absence of the tone CS were simply omitted. This means that on the average *Ss* in Group G received the same number of shocks during CS as *Ss* in Group R-1, but *Ss* in Group G received no shocks in the absence of CS and therefore fewer total shocks. The eight *Ss* in Group R-2 received a treatment identical to those of Group R-1 except that the average number of shocks they received (2.4 per session) was the same as that for Group G. These shocks were programmed randomly throughout the session, independently of the tones.

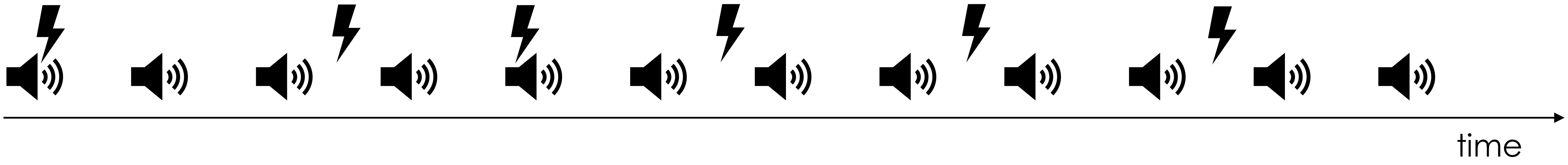
Group R-1: 12 CS; 12 US randomly programmed, such that US equiprobalbe at anytime



Group G: same number of CS-US pairings as Group R-1, BUT unpaired USs were removed. US occurrence is gated to CS occurrence. No US occurring in absence of CS.



Group R-2: random occurrence of US, as Group R-1, BUT total number of Uss same as Group G



Results

General HP: if conditioning happened in phase 2, the presence of the CS should suppress lever pressing at test

A median suppression ratio of:

- 0 indicates no responding in presence of CS
- .5 indicates similar rates of responding in presence vs absence of CS

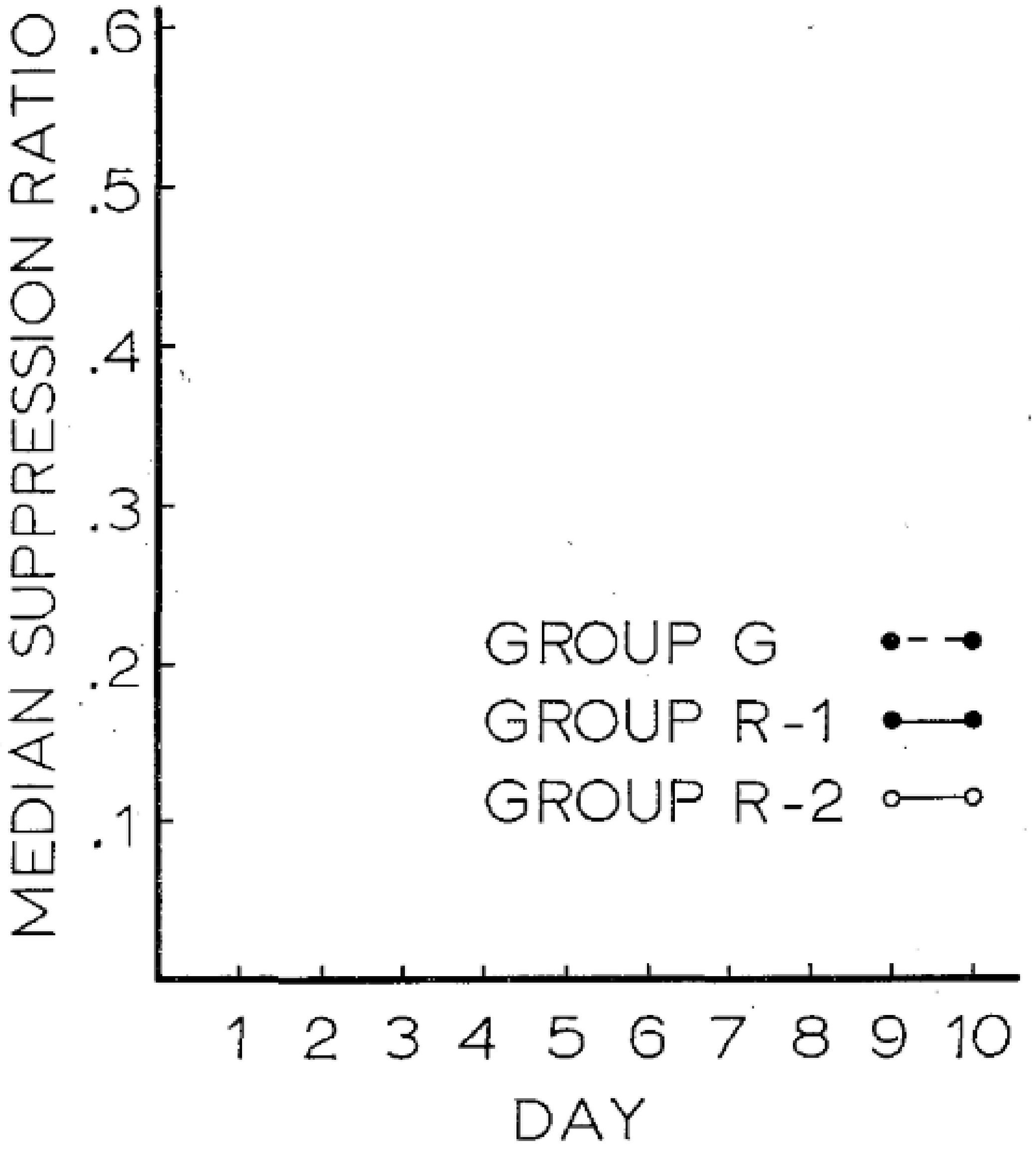


FIG. 1. Median suppression ratio for each group over the ten test sessions of Experiment 1.

Results

General HP: if conditioning happened in phase 2, the presence of the CS should suppress lever pressing at test

A median suppression ratio of:

- 0 indicates no responding in presence of CS
- .5 indicates similar rates of responding in presence vs absence of CS

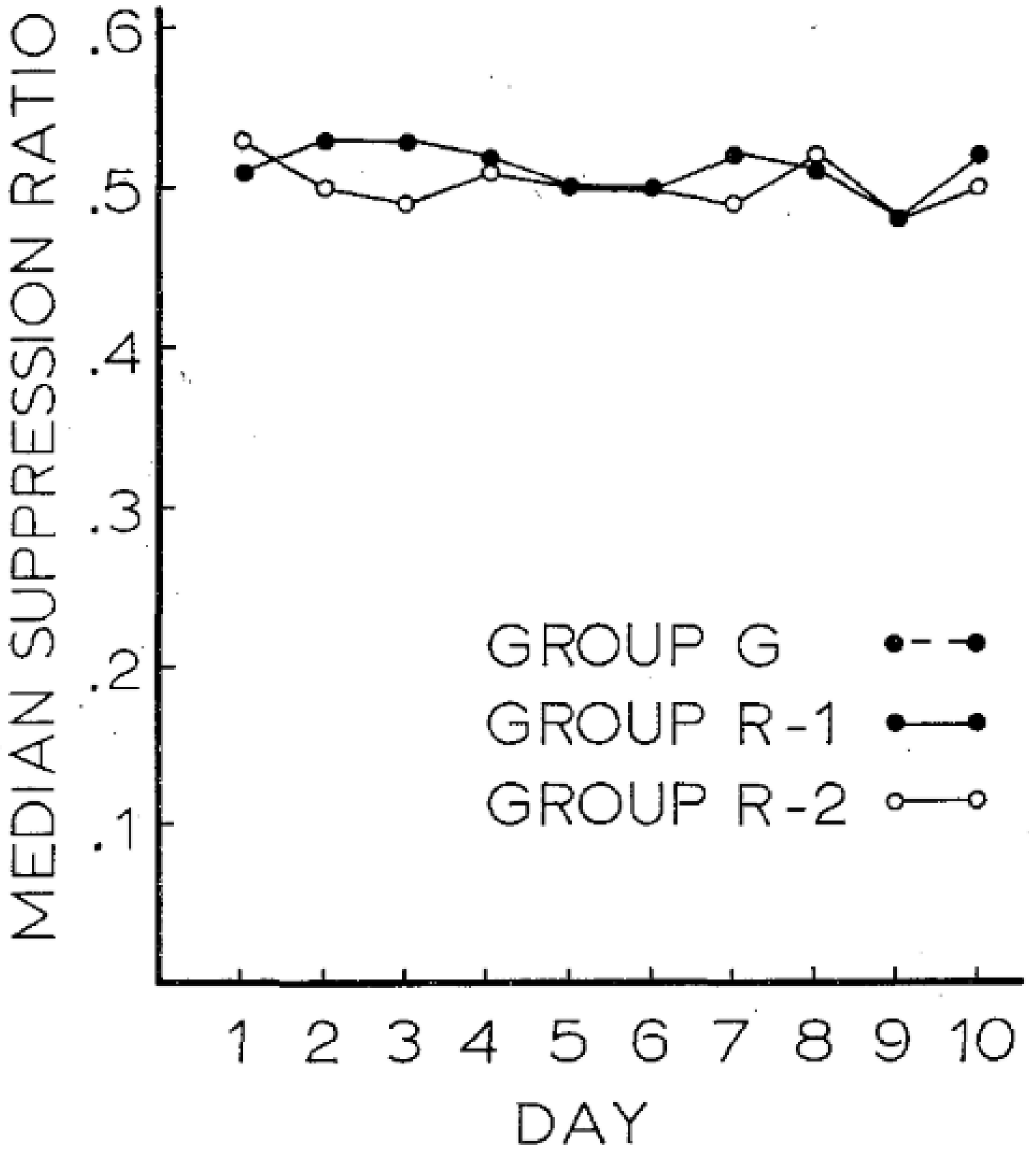


Fig. 1. Median suppression ratio for each group over the ten test sessions of Experiment 1.

Results

General HP: if conditioning happened in phase 2, the presence of the CS should suppress lever pressing at test

A median suppression ratio of:

- 0 indicates no responding in presence of CS
- .5 indicates similar rates of responding in presence vs absence of CS

Conditioned suppression occurs only in Group G where p(US/CS)>p(US/no-CS)

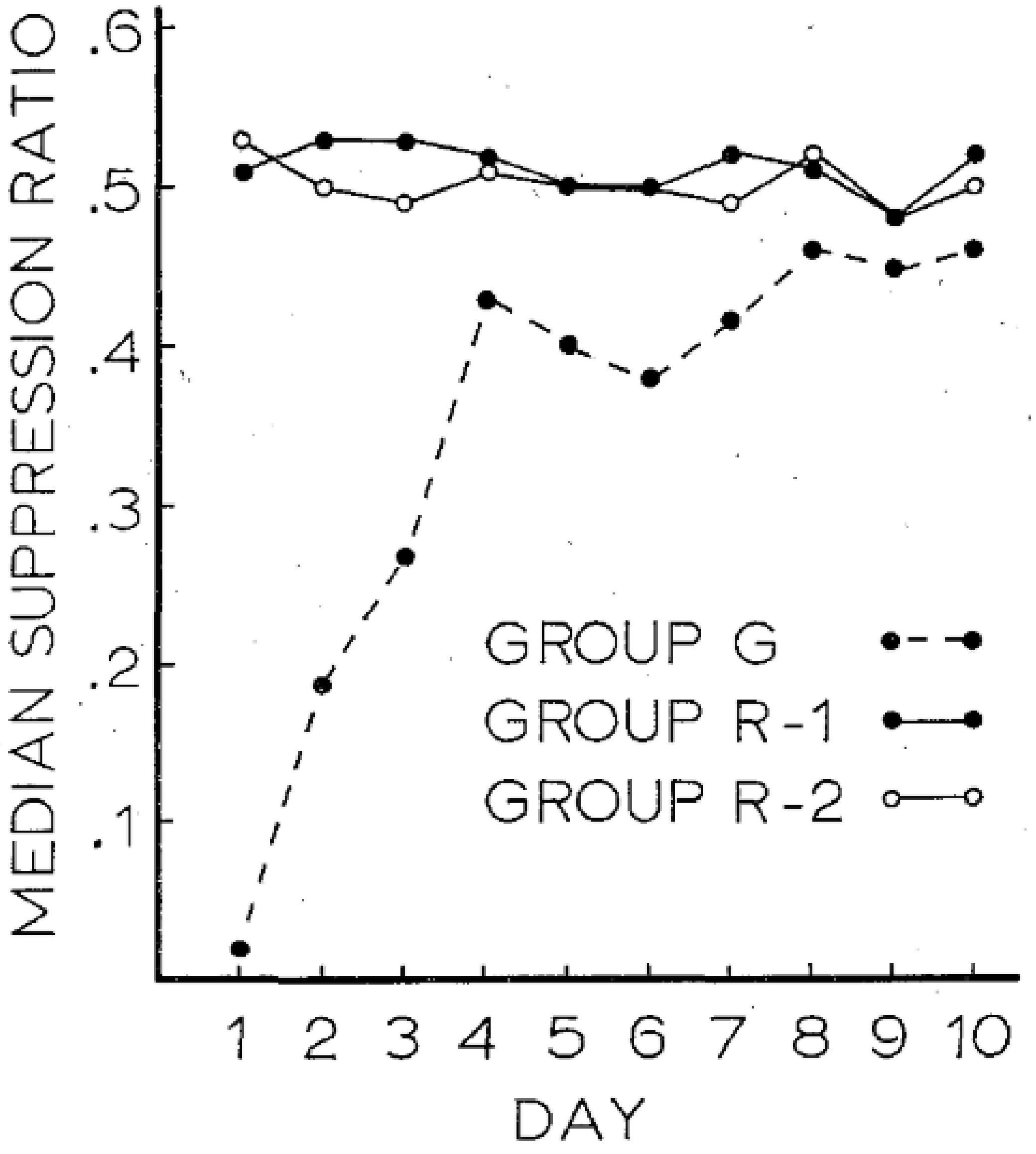


Fig. 1. Median suppression ratio for each group over the ten test sessions of Experiment 1.

MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA
CAMPUS DI CESENA

An additional experiment

Experimental phases

1. TRAINING
- Train rats to press lever to get food

2. PAVLOVIAN FEAR CONDITIONING
- CS: tone
- US: shock

3. TEST
- Measure the number of lever presses in presence vs absence of CS
- No shocks delivered --> the test is conducted under extinction
- General HP: if conditioning happened in phase 2, the presence of the CS should suppress lever pressing

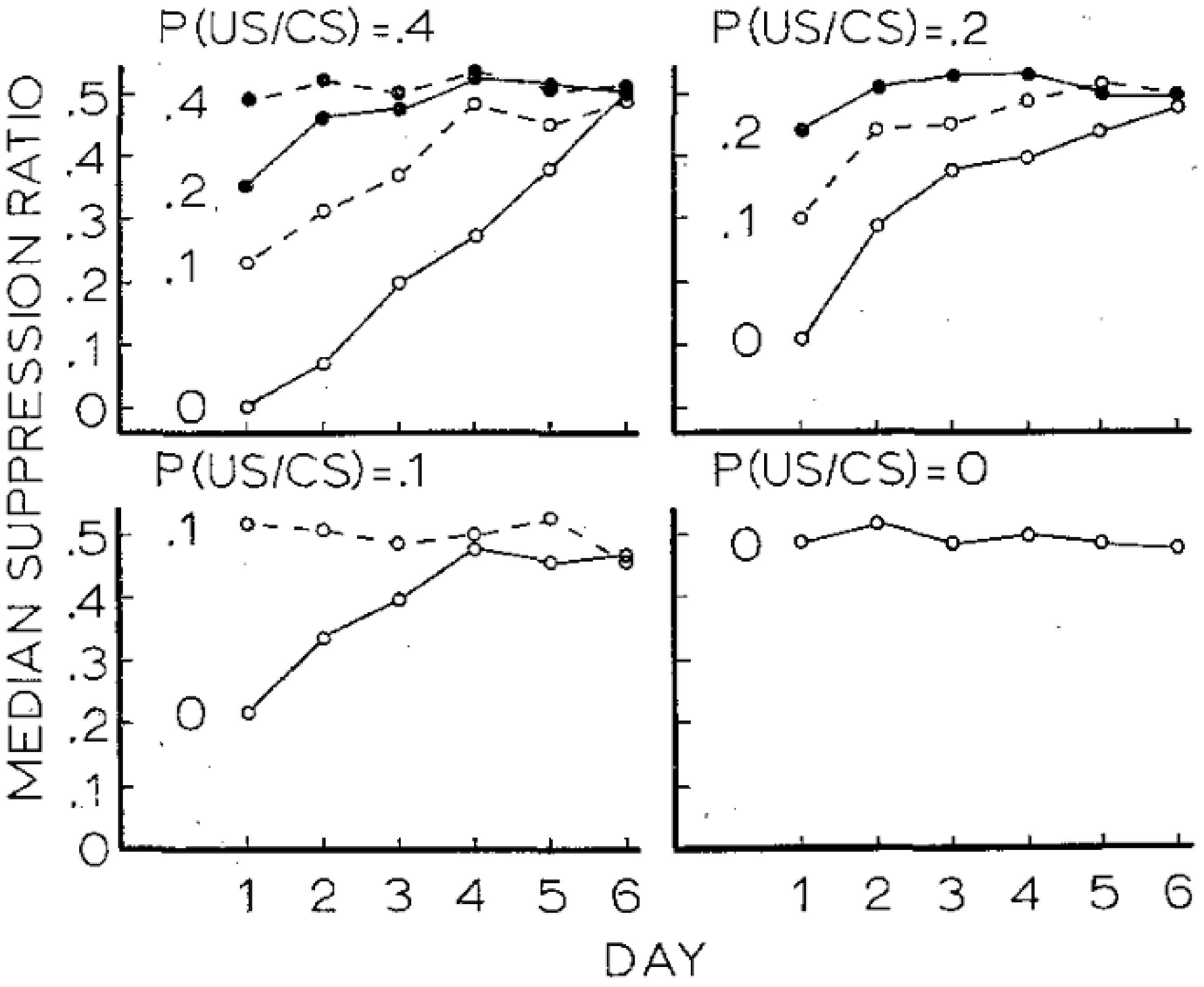


FIG. 3. Median suppression ratio for each group over the six test sessions of Experiment 2. (Within each panel, all groups have the same probability of US during CS; the parameter in each panel is the probability of US in the absence of CS.)

An additional experiment

Within each panel, all groups have the:

- **same p(US/CS)**

BUT

- **different p(US/no-CS)**

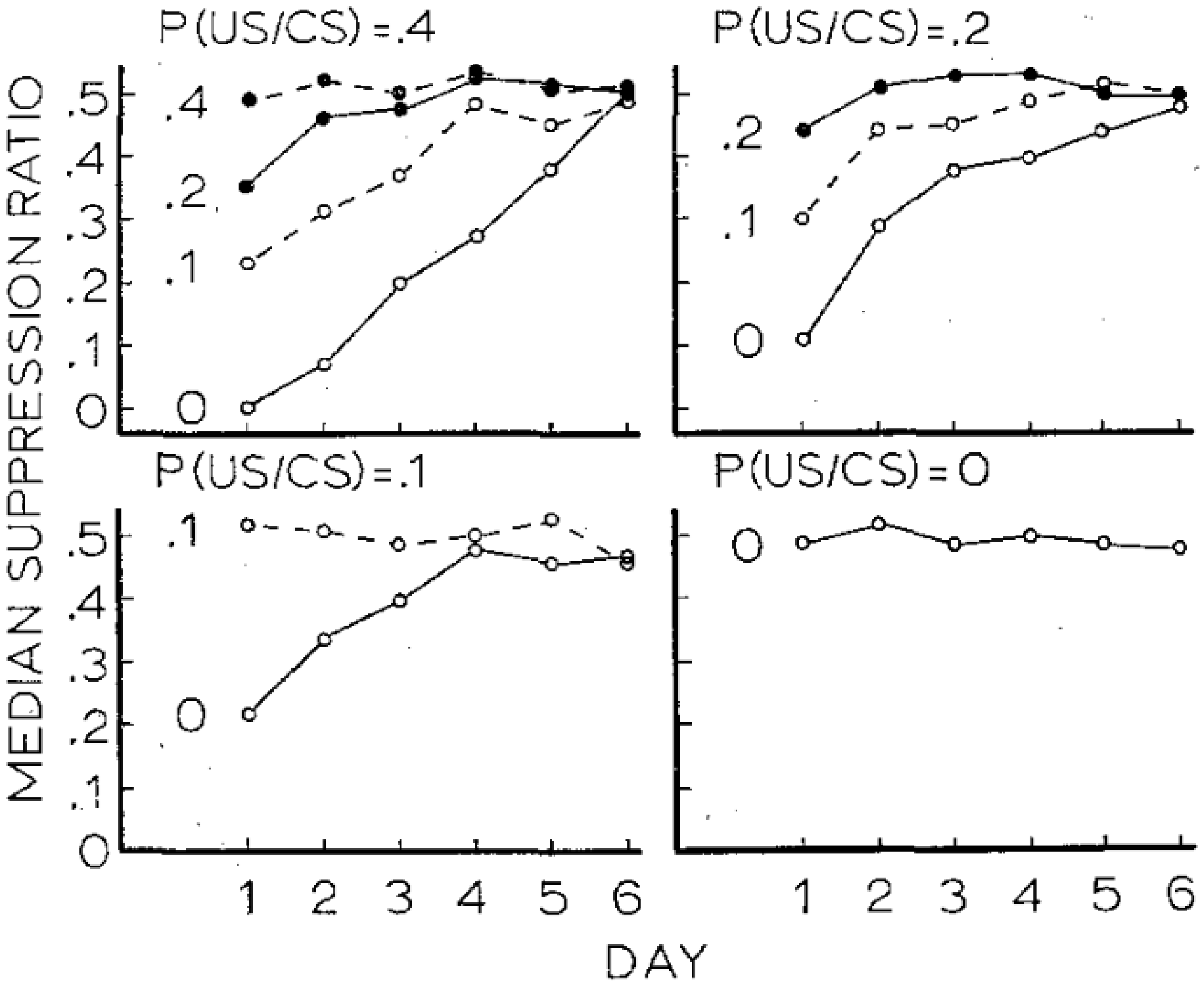


FIG. 3. Median suppression ratio for each group over the six test sessions of Experiment 2. (Within each panel, all groups have the same probability of US during CS; the parameter in each panel is the probability of US in the absence of CS.)

An additional experiment

Within each panel, all groups have the:

- **same p(US/CS)**

BUT

- **different p(US/no-CS)**

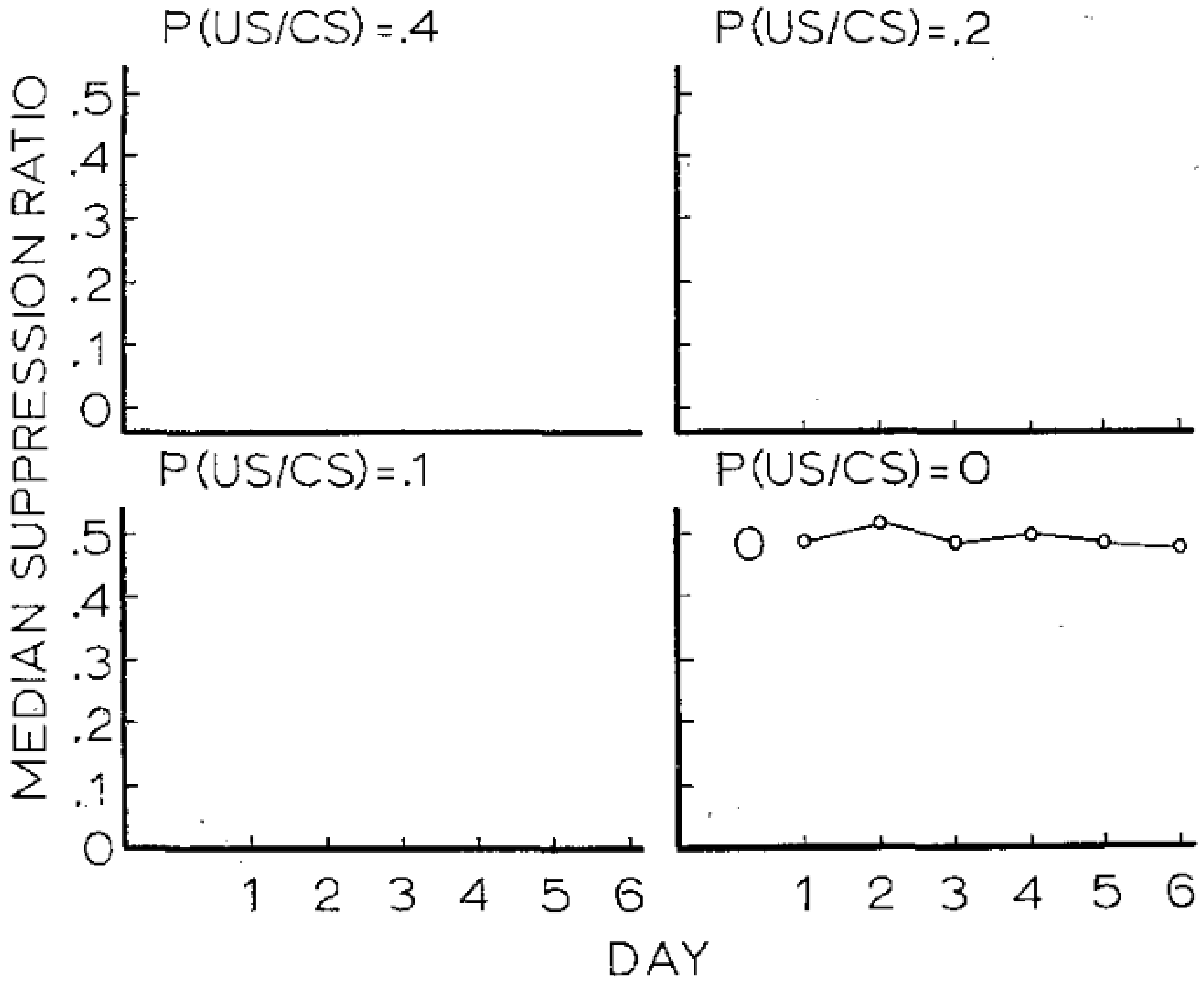


FIG. 3. Median suppression ratio for each group over the six test sessions of Experiment 2. (Within each panel, all groups have the same probability of US during CS; the parameter in each panel is the probability of US in the absence of CS.)

An additional experiment

Within each panel, all groups have the:

- **same p(US/CS)**

BUT

- **different p(US/no-CS)**

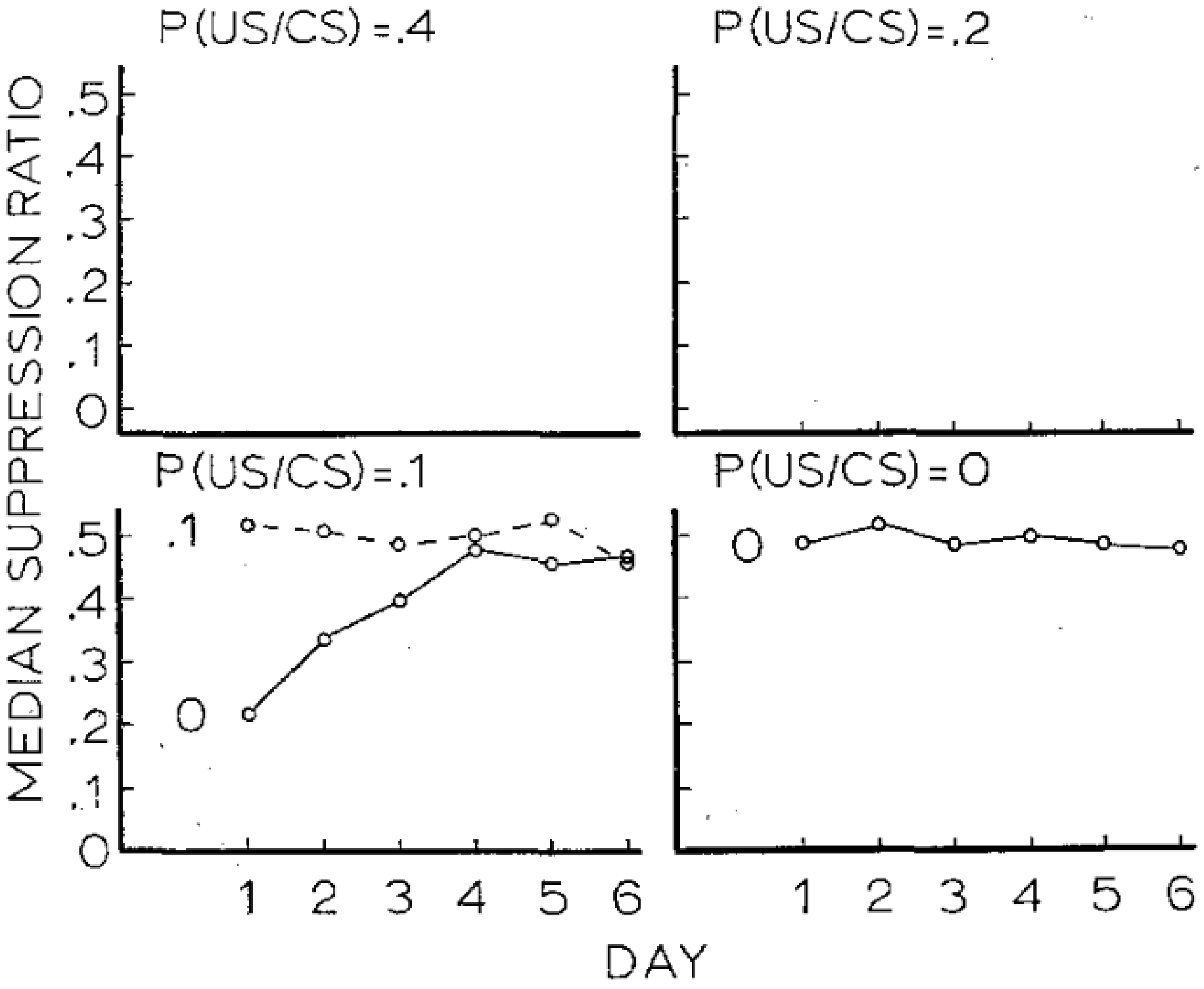


FIG. 3. Median suppression ratio for each group over the six test sessions of Experiment 2. (Within each panel, all groups have the same probability of US during CS; the parameter in each panel is the probability of US in the absence of CS.)

An additional experiment

Within each panel, all groups have the:

- **same p(US/CS)**

BUT

- **different p(US/no-CS)**

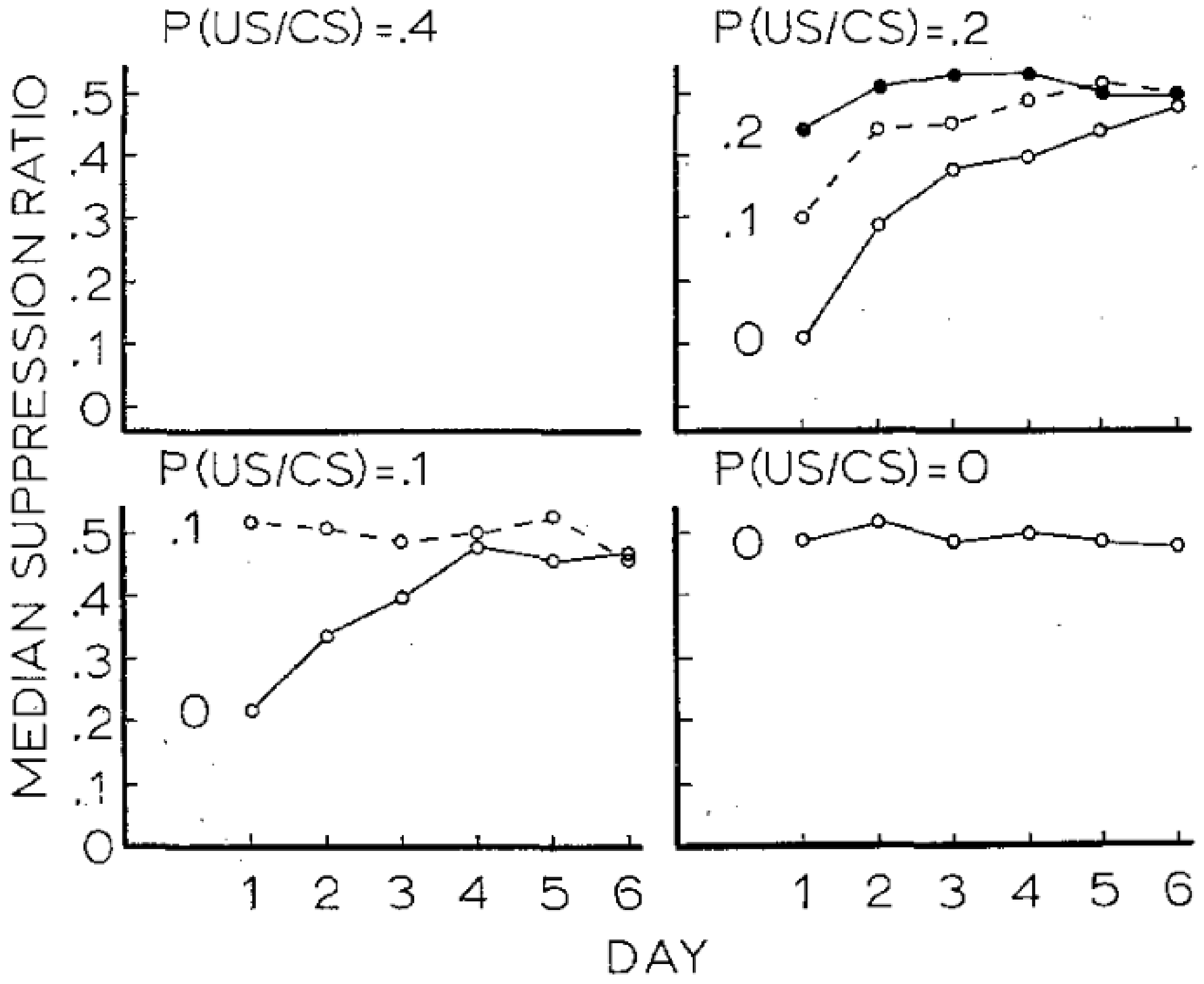


FIG. 3. Median suppression ratio for each group over the six test sessions of Experiment 2. (Within each panel, all groups have the same probability of US during CS; the parameter in each panel is the probability of US in the absence of CS.)

An additional experiment

Within each panel, all groups have the:

- same p(US/CS)

BUT

- different p(US/no-CS)

Conditioned suppression at test occurs only if

during Pavlovian conditioning
p(US/CS)>p(US/no-CS)

Also, the greater the difference
p(US/CS) - p(US/no-CS)
the stronger the conditioned suppression

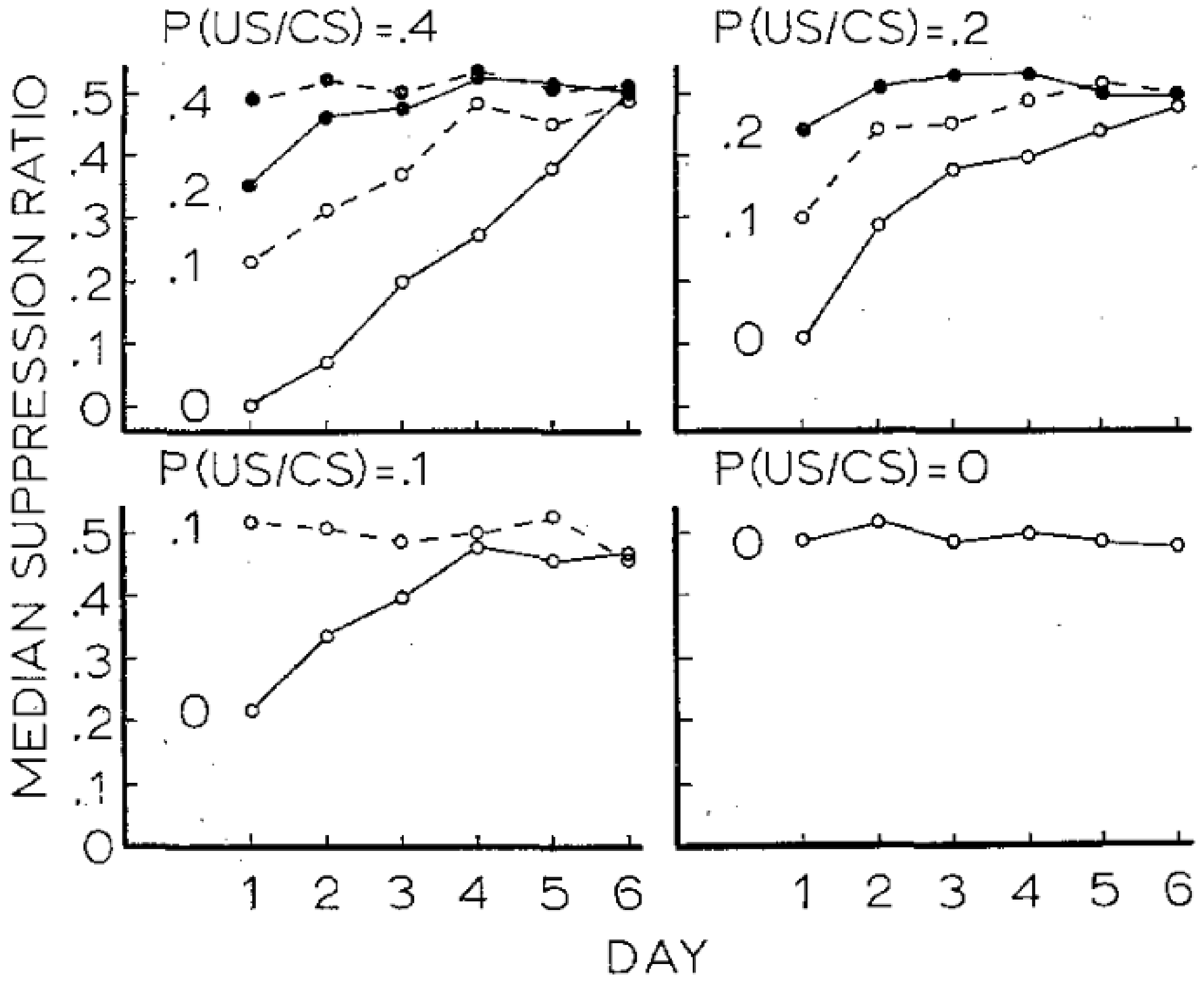


FIG. 3. Median suppression ratio for each group over the six test sessions of Experiment 2. (Within each panel, all groups have the same probability of US during CS; the parameter in each panel is the probability of US in the absence of CS.)

Reinforcement learning is more than learning about contiguity. It's also learning about contingency

Contingency:

- causal relationship between stimulus/behavior and outcome.
- $p(US/CS) > p(US/no\text{-}CS)$

Reinforcement learning is

- more than just learning the contiguous pairing between the CS and US.
- learning about the **causal relationship** between the stimuli/behavior and outcome (Dickinson 1980; Rescorla 1988)
- learning about **predictive relations**

Whether the CS is **informative** (or not) to predict the occurrence of the US is what matters for learning to happen.

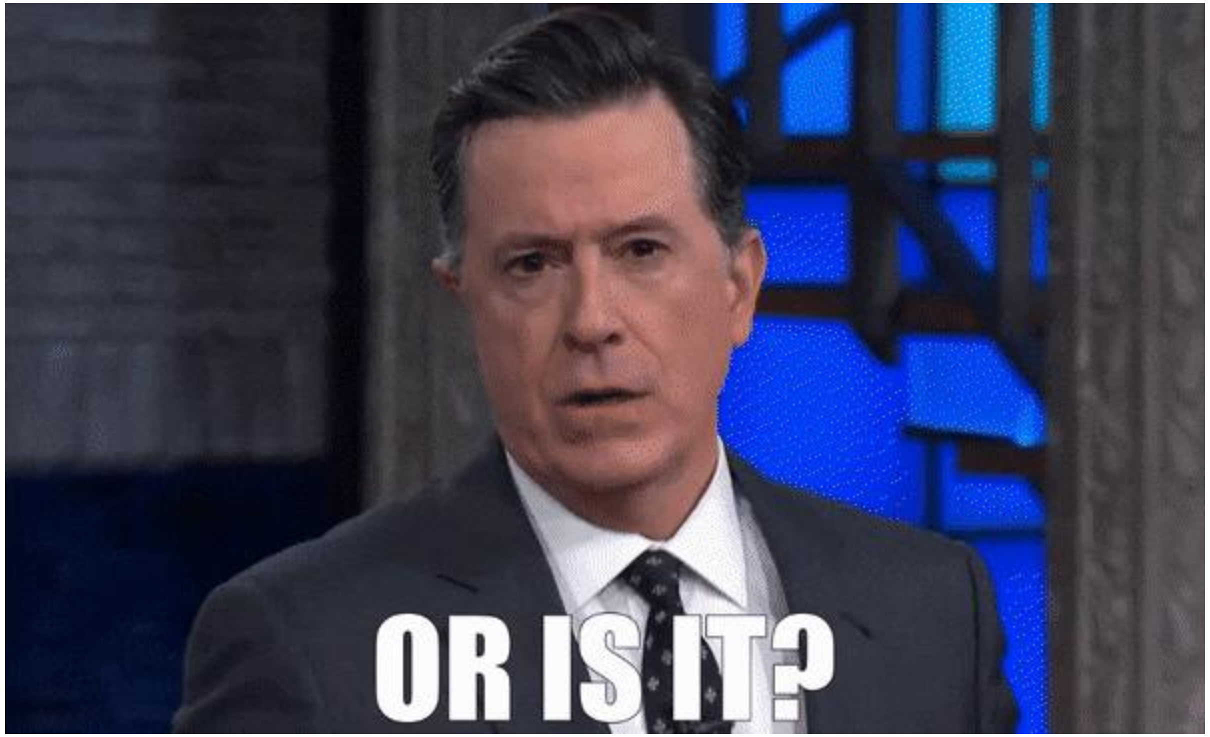
Reinforcement learning is learning about contiguity and contingency...

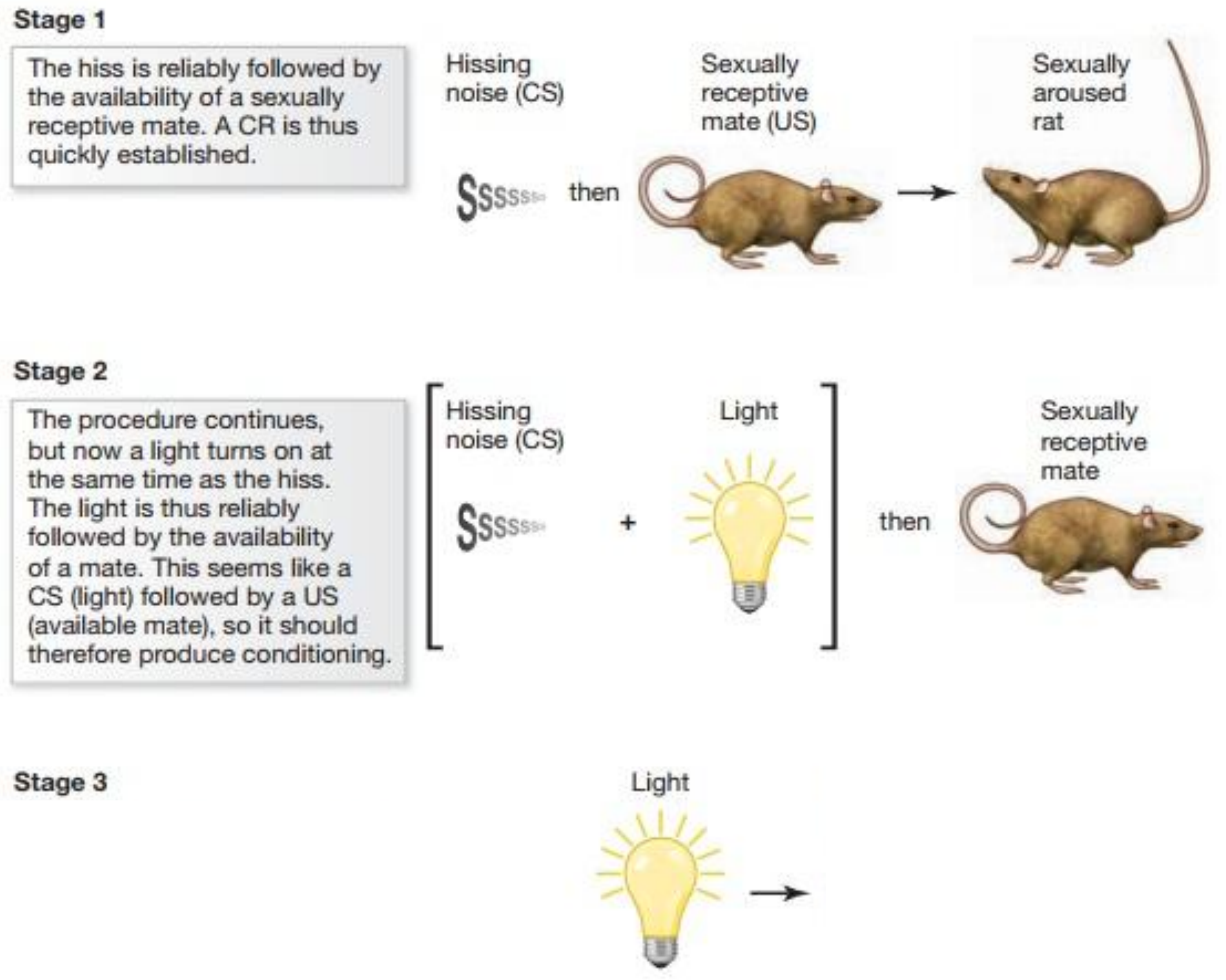
Contiguity: closeness in time between stimulus/behavior and outcome. Stimuli that are close to one another in time become associated. Not sufficient for learning.

Contingency: causal relationship between stimulus/behavior and outcome. When one stimulus depends on the other, they will become associated --> Predictive value critical

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA
CAMPUS DI CESENA


That's all Folks!


OR IS IT?



7.14 The blocking effect procedure

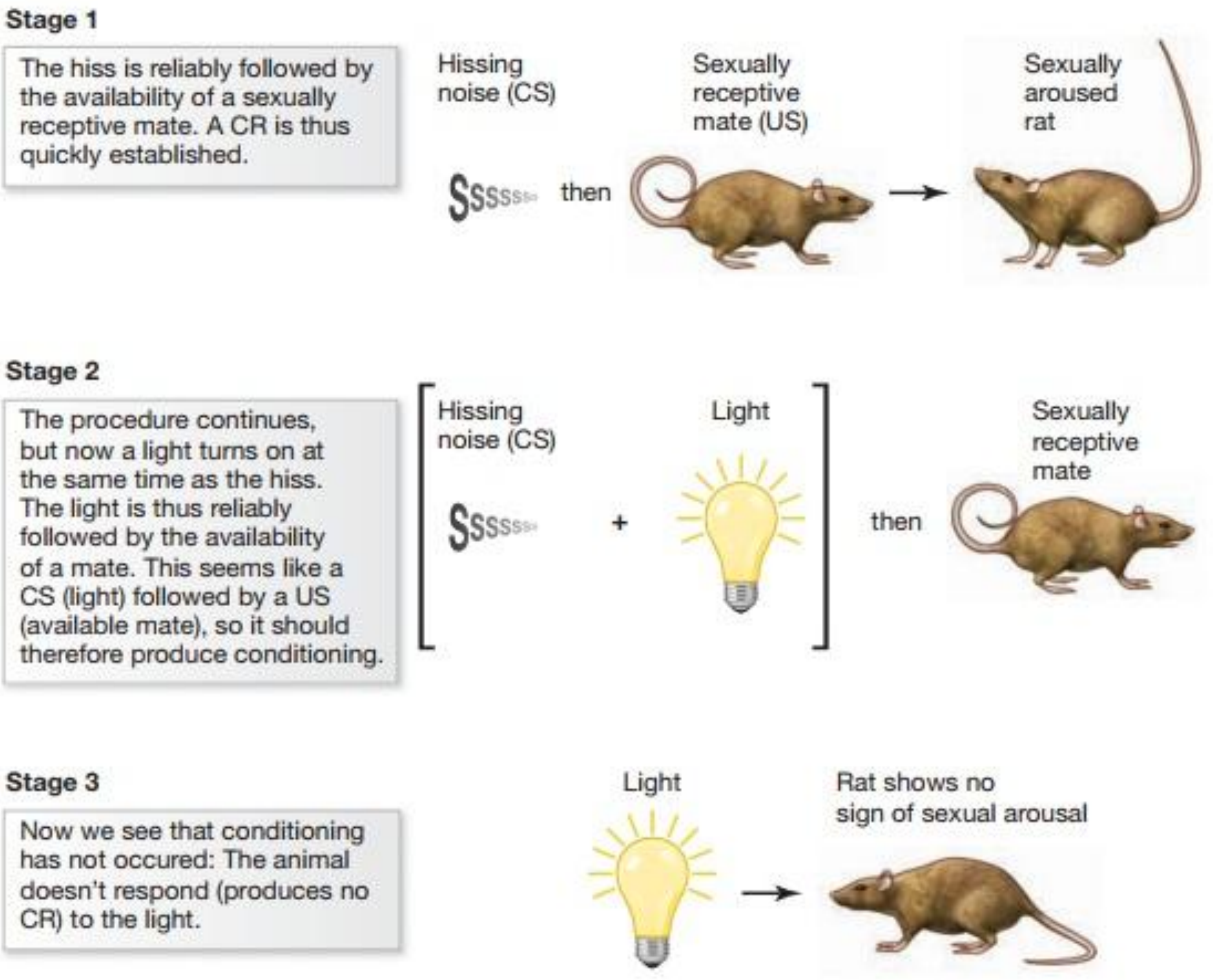
 https://www.brainkart.com/article/Classical-Conditioning--Role-of-Surprise_29314/

We learn only from surprising/unpredicted events: the blocking effect

No learning occurs to a stimulus if it is combined with a previously conditioned stimulus during conditioning trials.

The light provided only redundant information. The rat already knew from the hissing noise that the US was about to be presented, and so it wasn't at all surprised when the US did arrive.

P(US | [compound CS])≠P(US | original CS)



7.14 The blocking effect procedure What prevented the conditioning in this sequence? In Stage 2, the light provided no new information, because the hiss told the animal that the US would soon arrive. Conditioning does not occur with uninformative stimuli!

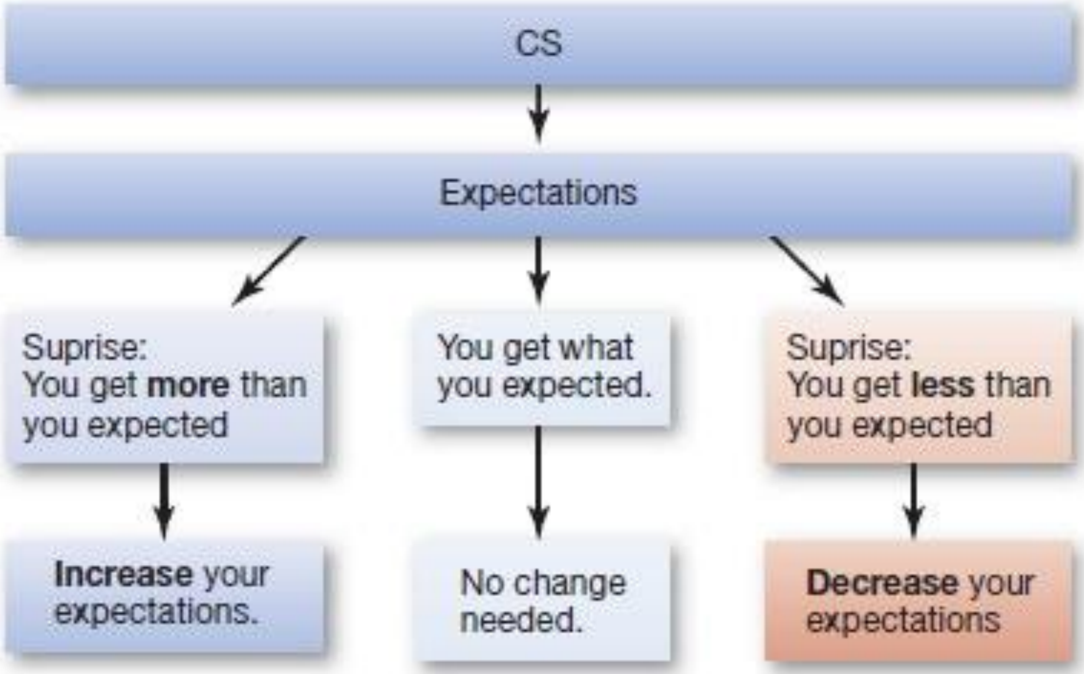
 https://www.brainkart.com/article/Classical-Conditioning--Role-of-Surprise_29314/

Both contiguity and contingency are not sufficient to learn, you also need surprise

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA
CAMPUS DI CESENA

We learn only from surprising/unpredicted events

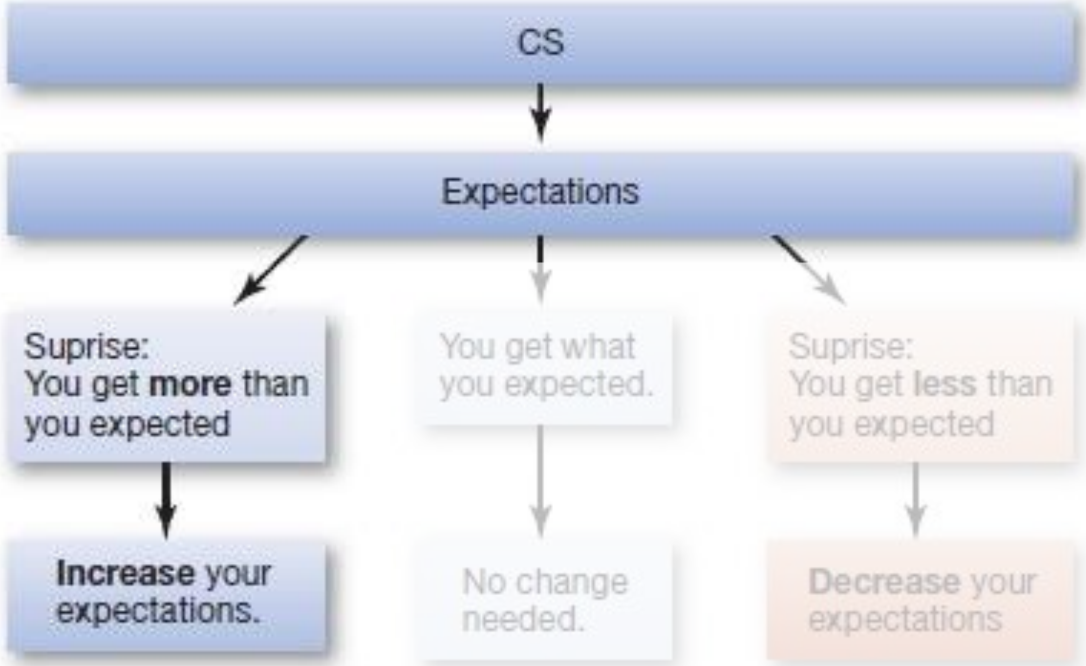
- Predictive learning depends on what is already known about those events.
- If little is known about the relation between the events, so that the US is not predicted by the CS, then learning occurs.
- If much is known about this relation, so that the US is adequately predicted by the CS, then learning fails.



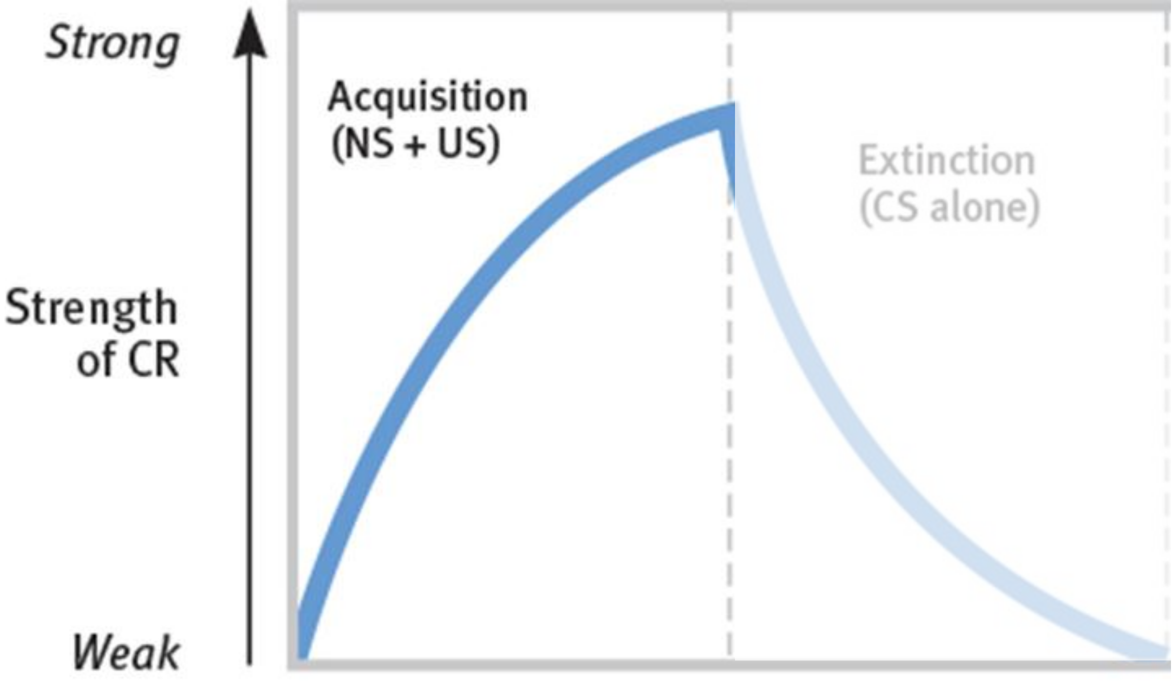
7.13 The role of expectations and surprise This figure shows the (automatic, unconscious) process through which expectations can be adjusted, trial by trial, in a classical conditioning experiment. The one complication not shown here is that bigger surprises (greater departures from expectations) will trigger larger adjustments; smaller surprises will lead to smaller adjustments.

https://www.brainkart.com/article/Classical-Conditioning--Role-of-Surprise_29314/

We learn only from surprising/unpredicted events

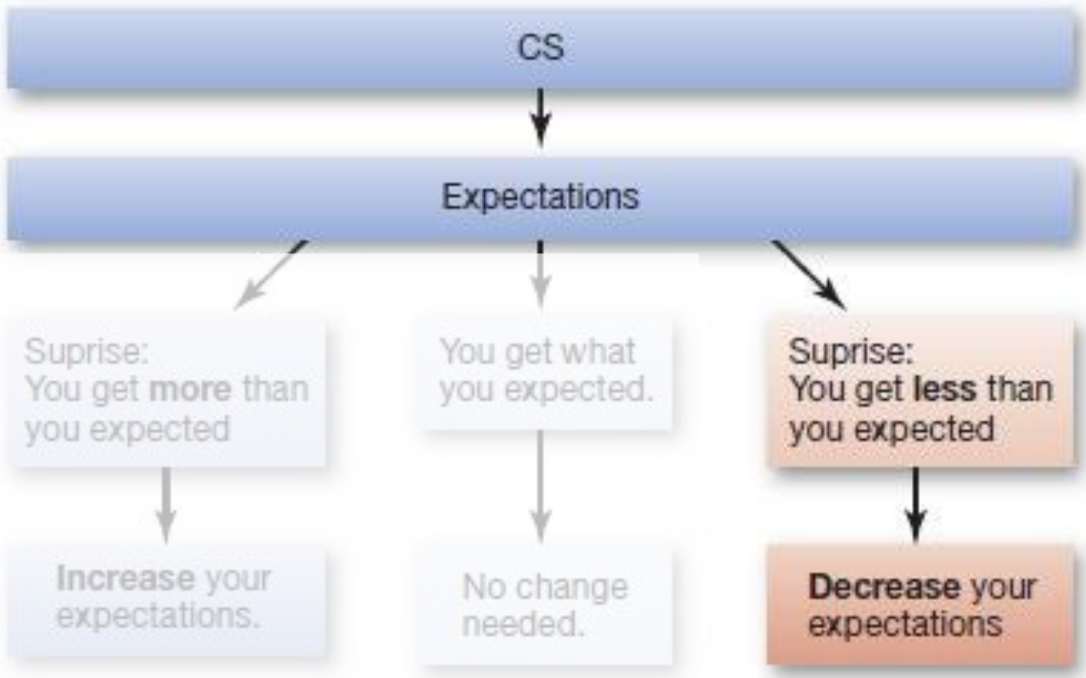


7.13 The role of expectations and surprise This figure shows the (automatic, unconscious) process through which expectations can be adjusted, trial by trial, in a classical conditioning experiment. The one complication not shown here is that bigger surprises (greater departures from expectations) will trigger larger adjustments; smaller surprises will lead to smaller adjustments.

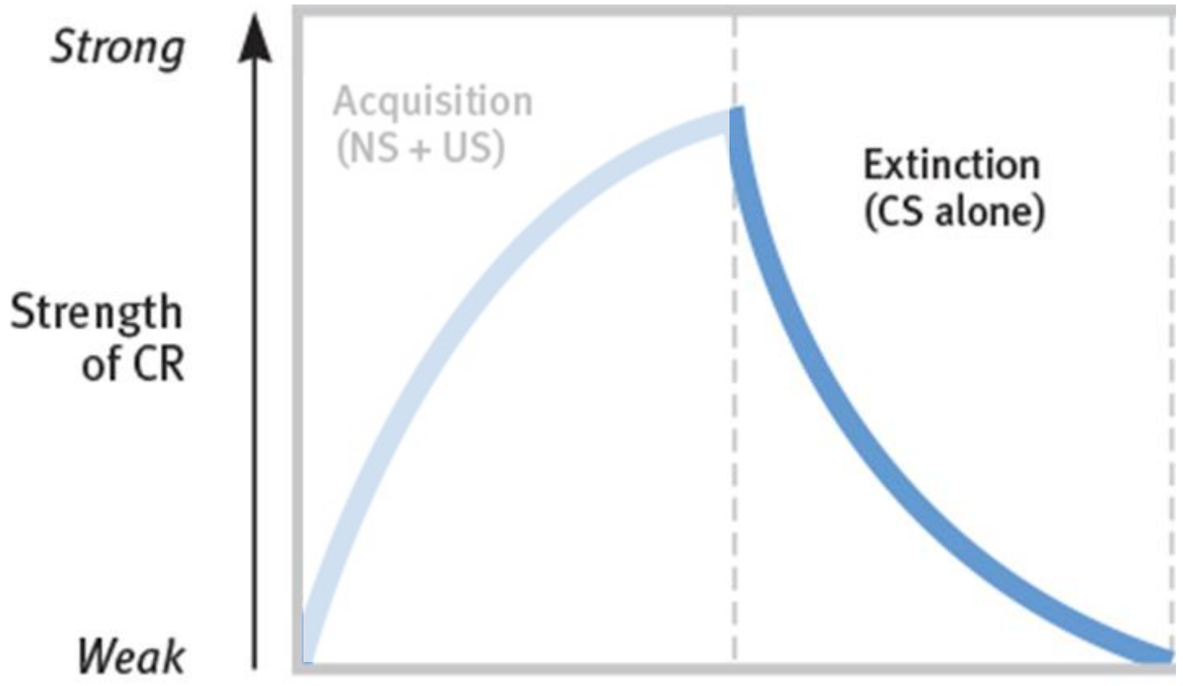


https://www.brainkart.com/article/Classical-Conditioning--Role-of-Surprise_29314/

We learn only from surprising/unpredicted events

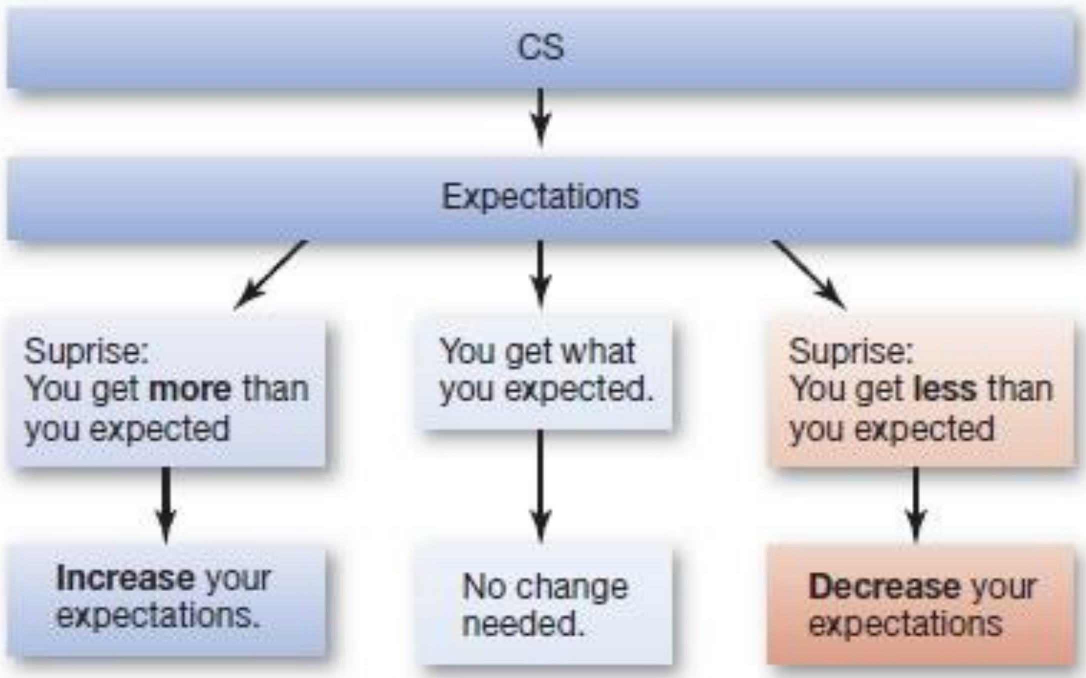


7.13 The role of expectations and surprise This figure shows the (automatic, unconscious) process through which expectations can be adjusted, trial by trial, in a classical conditioning experiment. The one complication not shown here is that bigger surprises (greater departures from expectations) will trigger larger adjustments; smaller surprises will lead to smaller adjustments.



https://www.brainkart.com/article/Classical-Conditioning--Role-of-Surprise_29314/

We learn only from surprising/unpredicted events



7.13 The role of expectations and surprise This figure shows the (automatic, unconscious) process through which expectations can be adjusted, trial by trial, in a classical conditioning experiment. The one complication not shown here is that bigger surprises (greater departures from expectations) will trigger larger adjustments; smaller surprises will lead to smaller adjustments.

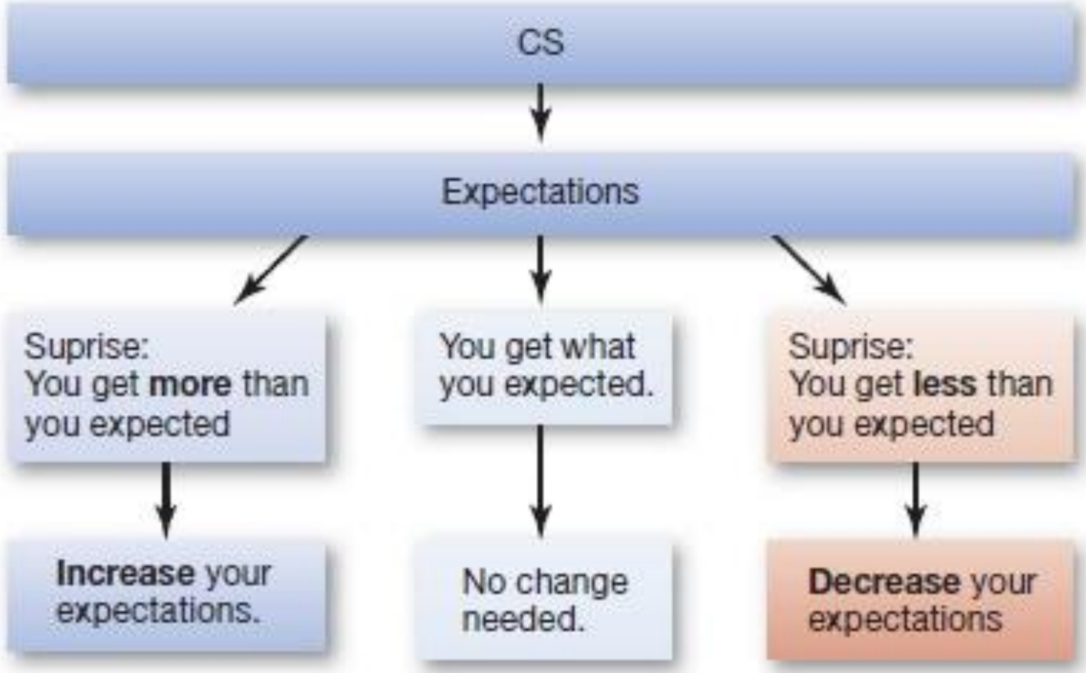
Reinforcement learning is not instructed by a simple sensory representation of the outcome,

but instead by an **error signal measuring the difference between the outcome expected** and that **actually present**, i.e. the **Prediction error**

(McNally, 2011)

https://www.brainkart.com/article/Classical-Conditioning--Role-of-Surprise_29314/

We learn only from surprising/unpredicted events



7.13 The role of expectations and surprise This figure shows the (automatic, unconscious) process through which expectations can be adjusted, trial by trial, in a classical conditioning experiment. The one complication not shown here is that bigger surprises (greater departures from expectations) will trigger larger adjustments; smaller surprises will lead to smaller adjustments.

Reinforcement learning is regulated by prediction errors

Prediction error

- is a quantitative discrepancy between the **outcome expected** when the cue was presented, and the **outcome** that was actually **experienced**

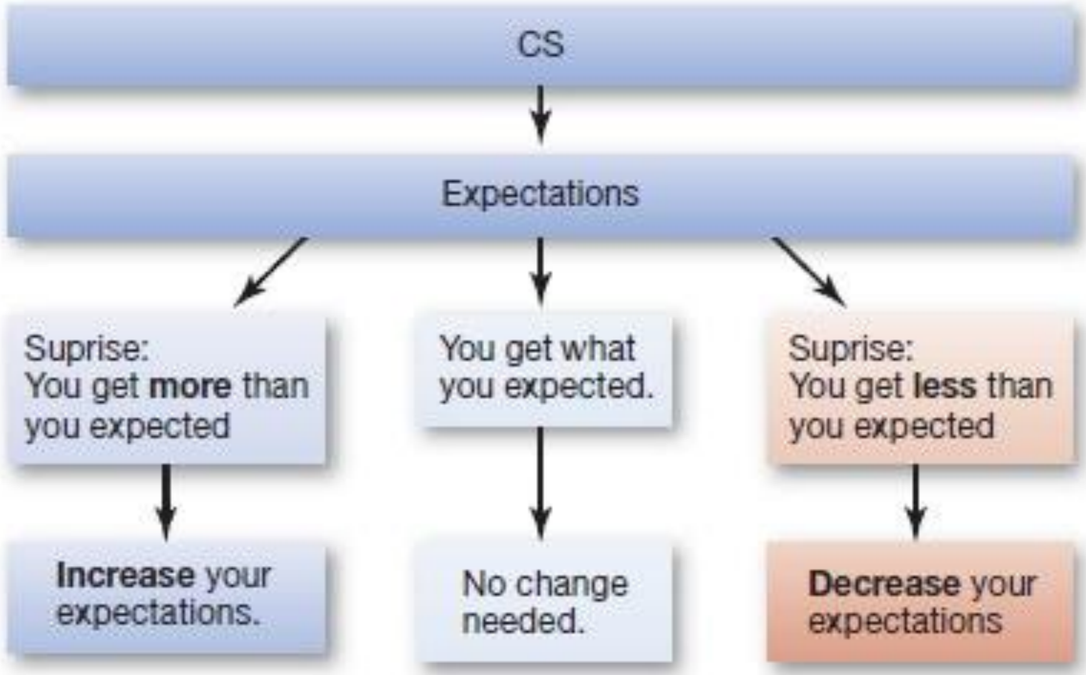
- functions as **teaching signal** to update expectations and reduce following prediction errors

(McNally, 2011)

https://www.brainkart.com/article/Classical-Conditioning--Role-of-Surprise_29314/

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA
CAMPUS DI CESENA

We learn only from surprising/unpredicted events



7.13 The role of expectations and surprise This figure shows the (automatic, unconscious) process through which expectations can be adjusted, trial by trial, in a classical conditioning experiment. The one complication not shown here is that bigger surprises (greater departures from expectations) will trigger larger adjustments; smaller surprises will lead to smaller adjustments.

Learning is regulated by prediction errors

Prediction error

- is a quantitative discrepancy between the **outcome expected** when the cue was presented, and the **outcome** that was actually **experienced**
- Functions as **teaching signal** to update expectations and reduce following prediction errors

https://www.brainkart.com/article/Classical-Conditioning--Role-of-Surprise_29314/

Learning is regulated by prediction errors

- Error signal measuring the difference between the outcome actually present and that expected

- Occurs when the outcome of a conditioning trial is different from that which is predicted by the conditioned stimuli that are present on the trial (i.e., when the US is surprising)

Prediction error is **necessary for associative learning**

- Dictates variations in the effectiveness of the US in supporting learning
- if the difference is large, predictions did not match observations, and there is a need for more learning to update those predictions
- If predictions match observations, there is no prediction error and no learning occurs

Computational models

Mathematical formalization of learning theories

Computational models: the Rescorla-Wagner model

- Change in value is proportional to the difference between predicted and actual outcome --> error driven learning

Prediction error

$$\delta_t = R_t - V_t$$

- R is the value of US
    - R=1 if US is delivered
    - R=0 if US is omitted
- V is the expectancy of US or the expected value of a given CS
- then the error signal is computed as R−V
- the expected value V of a given CS is updated based on the sum of the current expected value V and the prediction error at trial t

$$V_{t+1} = V_t + \alpha\delta_t$$

- Where α is a learning rate parameter, which determines the size of the update step (α assumed between 0 and 1)

Excel file

Computational models: the Rescorla-Wagner model

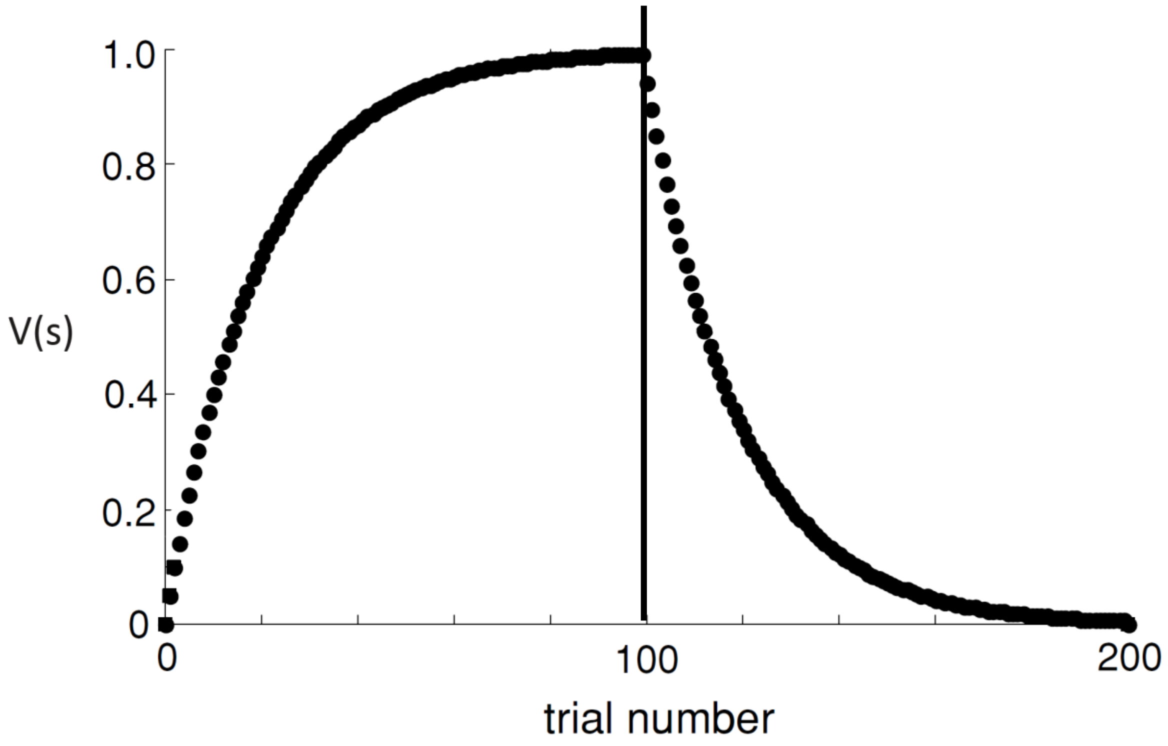
Prototypical acquisition and extinction learning curves for Pavlovian conditioning as predicted by the Rescorla-Wagner model.

The filled circles show the time **evolution of the value V** of a CS over 200 trials.

- In the first 100 trials, a reward R was paired with the CS (acquisition)
- in trials 100-200 no reward was paired (extinction)

Learning is proportional to prediction error δ, which is larger at the start of training when reward is unexpected, and gets smaller, eventually 0, as training progresses and reward is fully predicted.

$$\delta_t = R_t - V_t$$



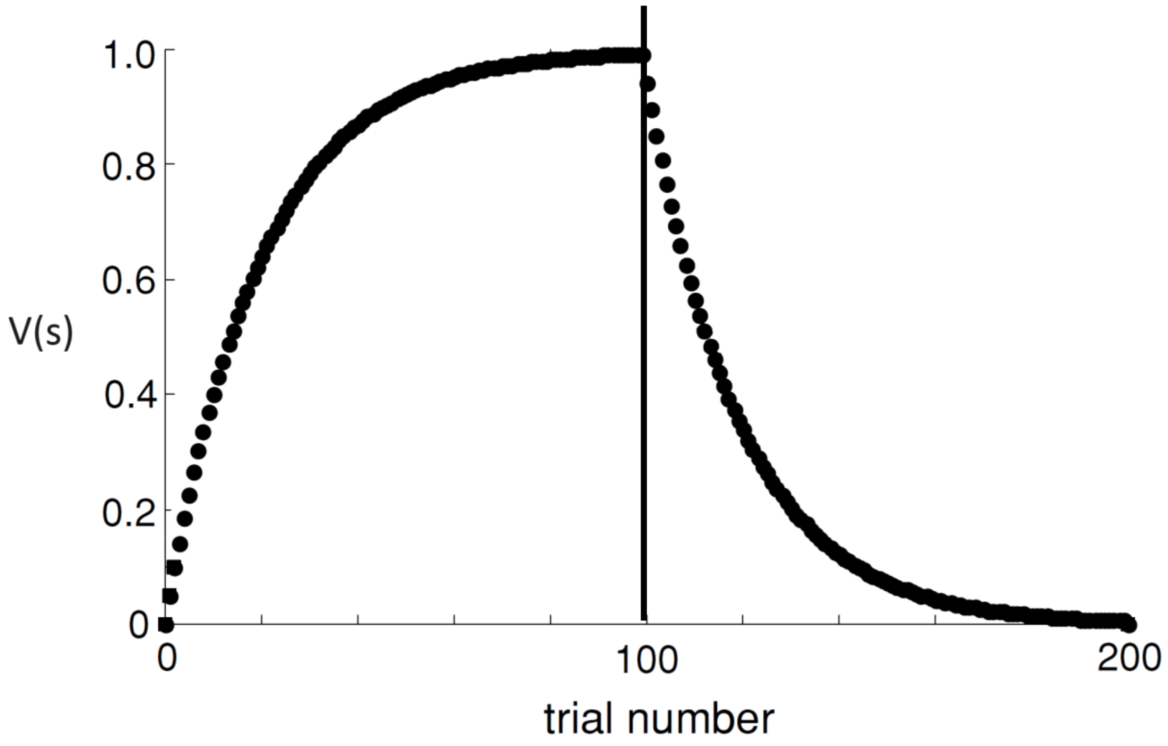
Computational models: the Rescorla-Wagner model

Learning is proportional to prediction error δ

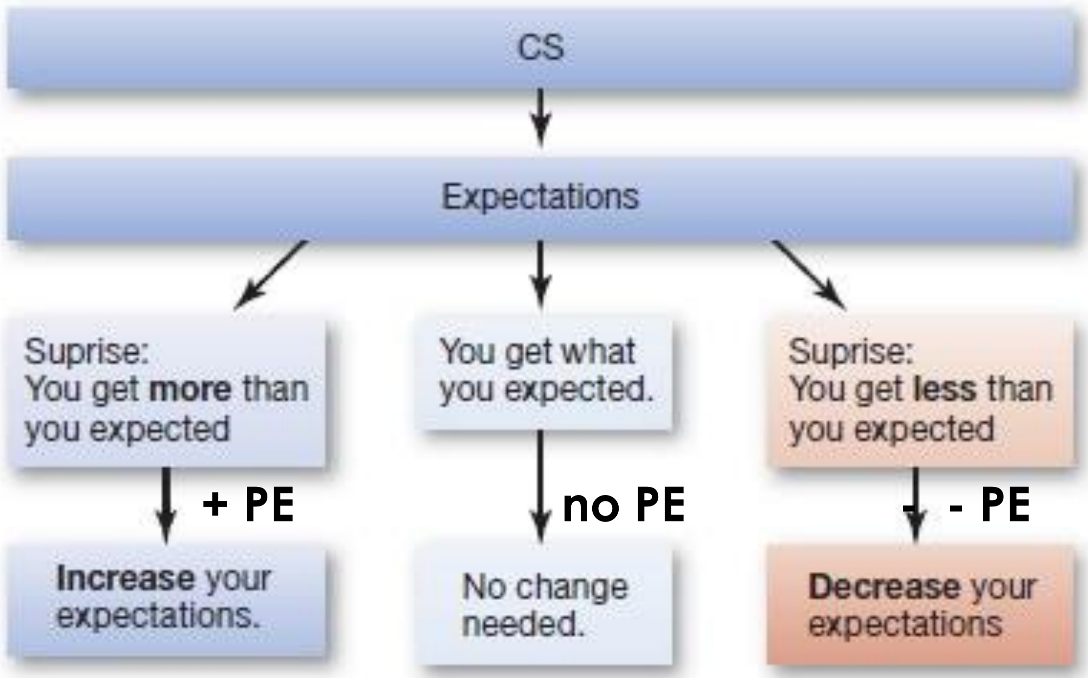
- δ is **positive** (more than expected) and larger at the start of acquisition when reward delivery is unexpected
- δ gets smaller, eventually 0, as acquisition progresses and reward delivery is fully predicted

- δ is **negative** (less than expected) and larger at the start of extinction when reward omission is unexpected
- δ gets smaller, eventually 0, as extinction progresses and reward omission is fully predicted

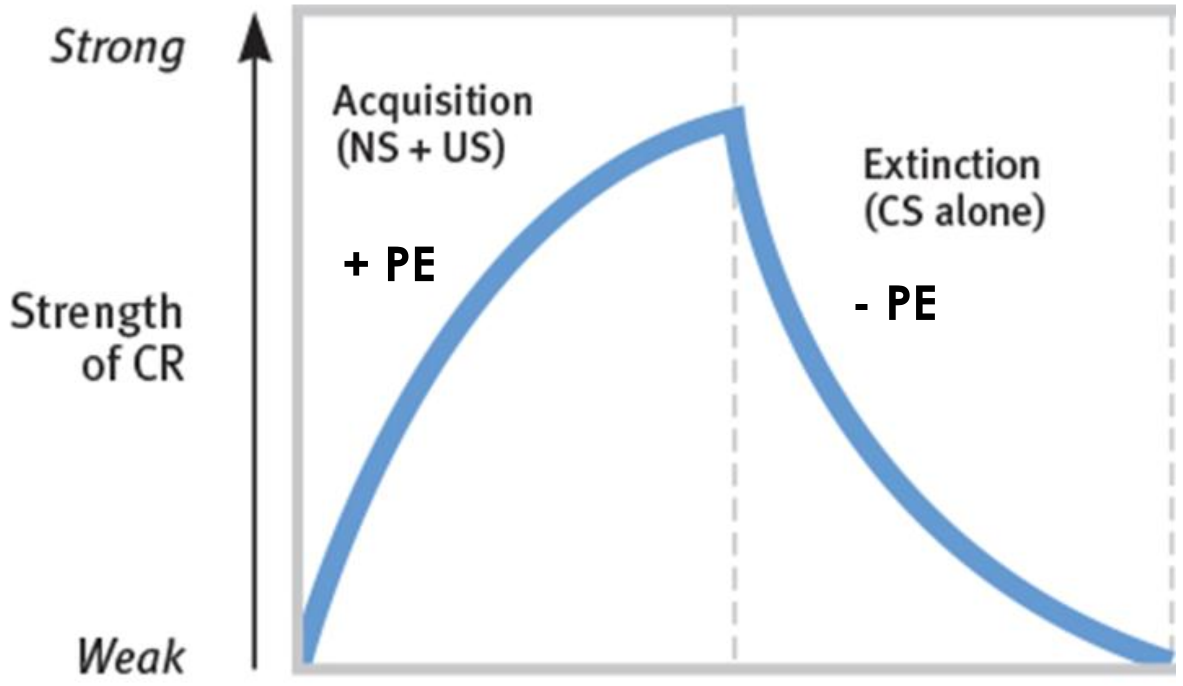
$$\delta_t = R_t - V_t$$



Rescorla-Wagner model (and TD model – see next slides) relies upon signed prediction errors

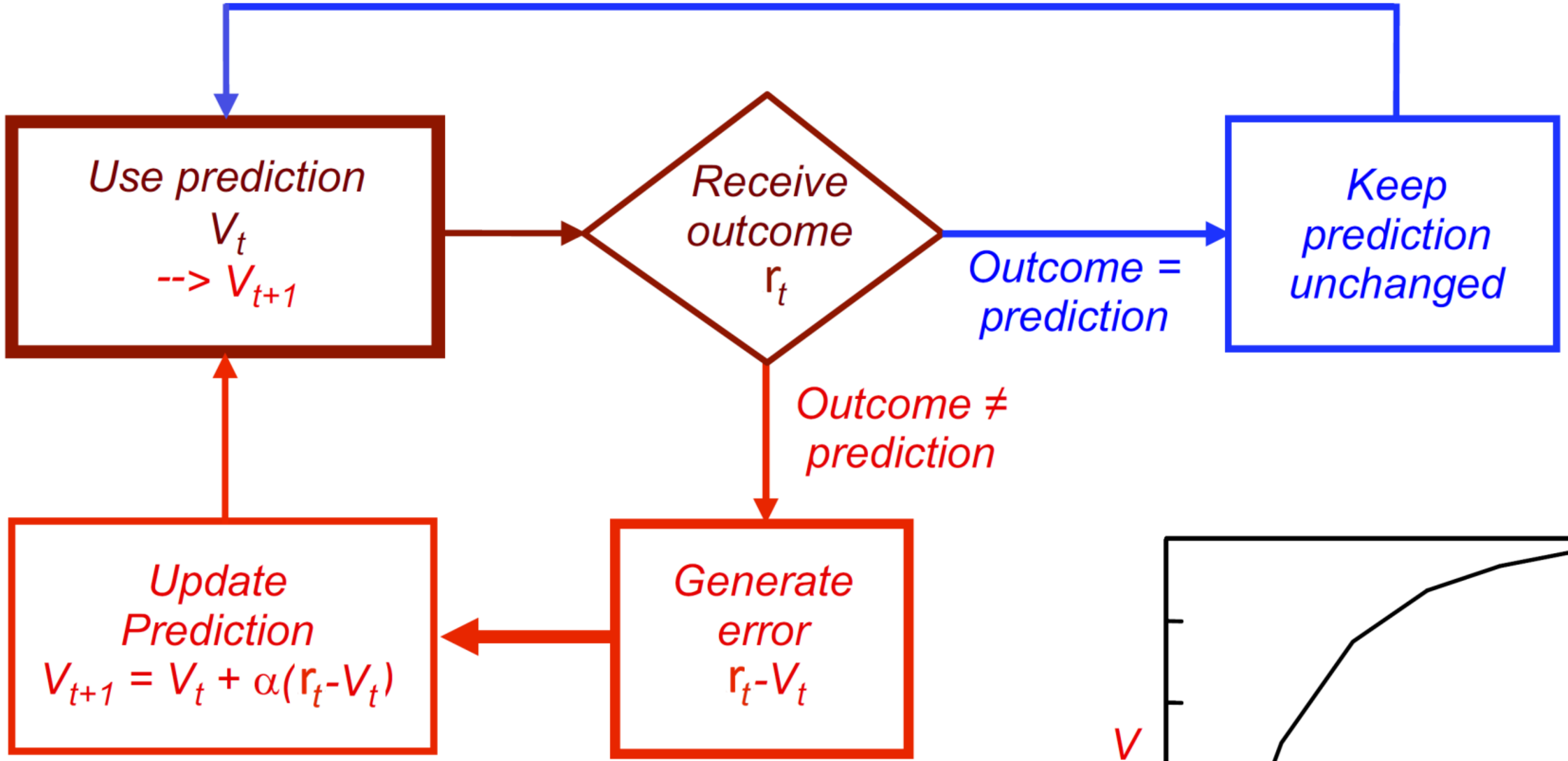


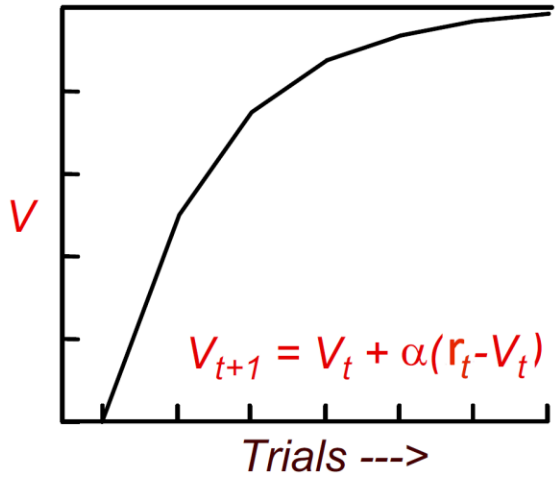
7.13 The role of expectations and surprise This figure shows the (automatic, unconscious) process through which expectations can be adjusted, trial by trial, in a classical conditioning experiment. The one complication not shown here is that bigger surprises (greater departures from expectations) will trigger larger adjustments; smaller surprises will lead to smaller adjustments.



https://www.brainkart.com/article/Classical-Conditioning--Role-of-Surprise_29314/

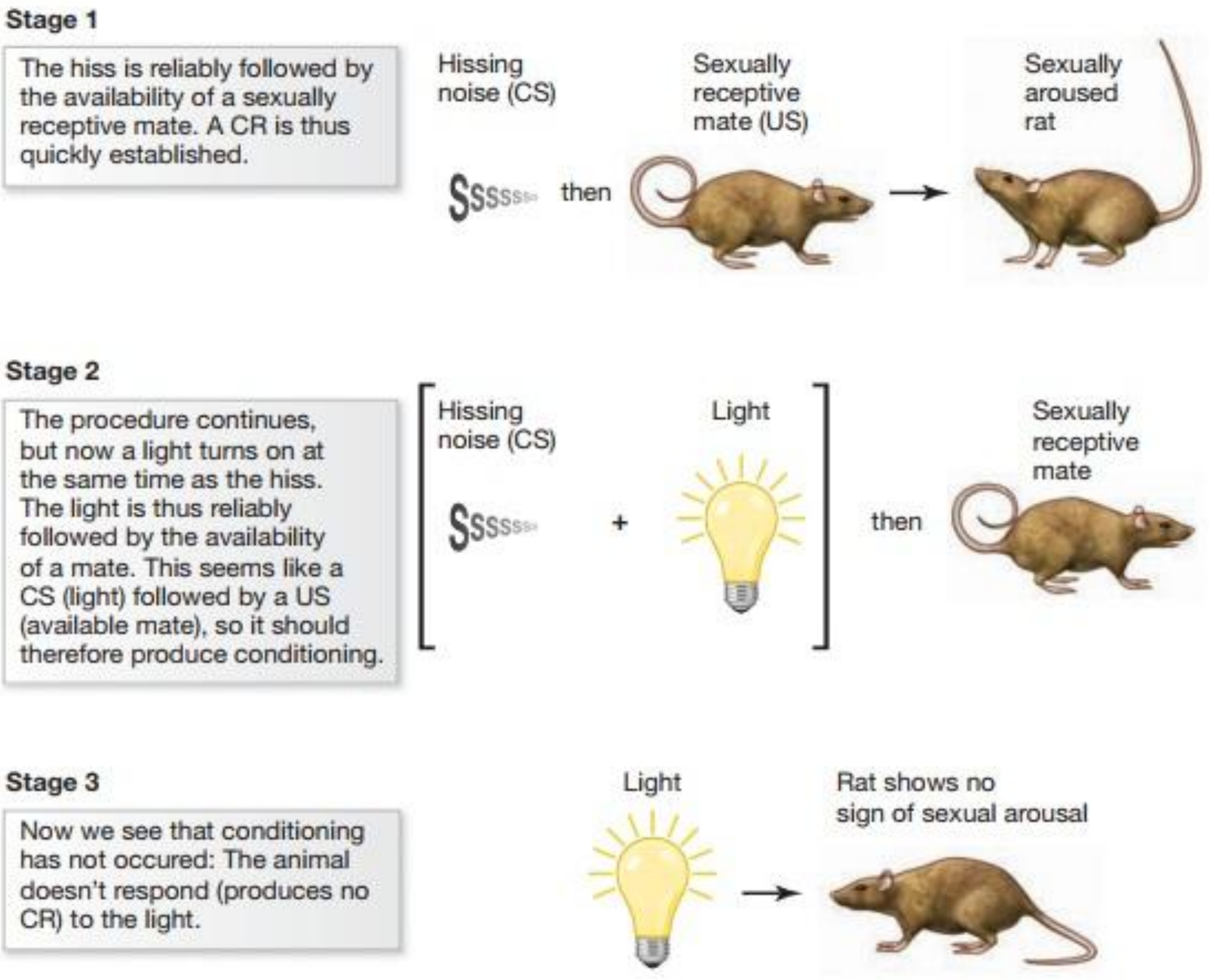
Computational models: the Rescorla-Wagner model





Computational models: the Rescorla-Wagner model & blocking

The Rescorla-Wagner model captures blocking by specifying that when multiple stimuli are observed (e.g. light and tone), the animal computes a single net prediction and prediction error, from the sum of all predictions concerning the stimuli present on a given trial.



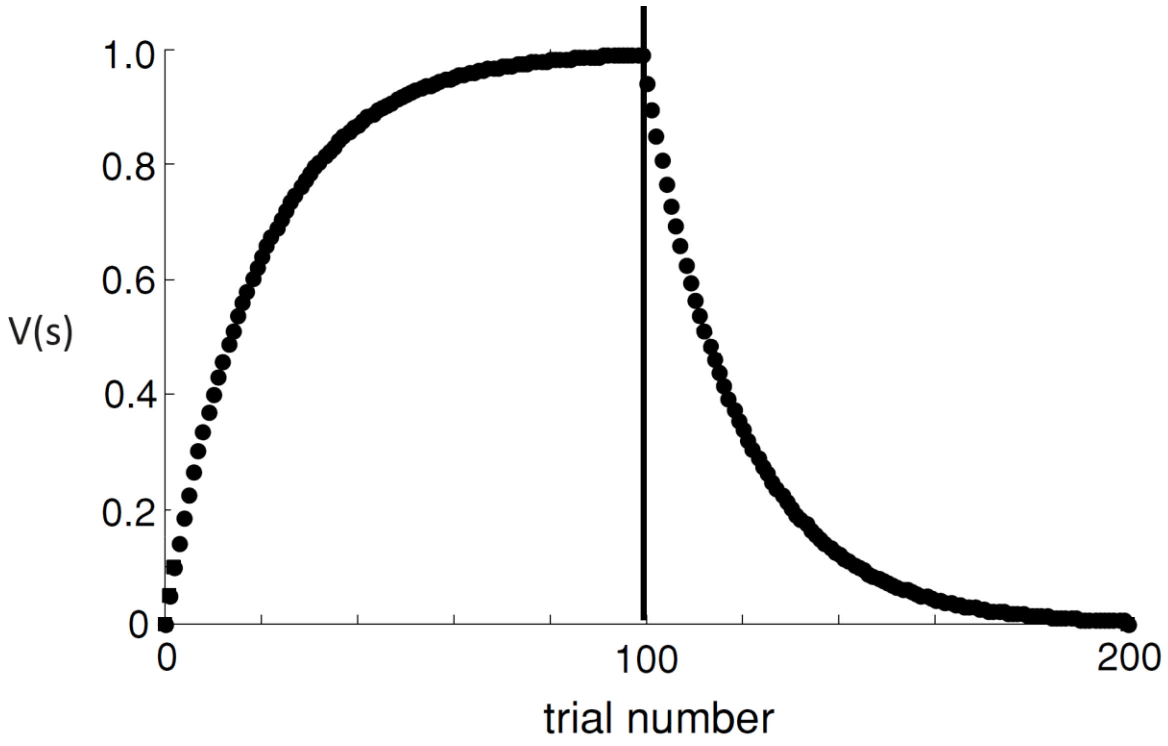
7.14 The blocking effect procedure What prevented the conditioning in this sequence? In Stage 2, the light provided no new information, because the hiss told the animal that the US would soon arrive. Conditioning does not occur with uninformative stimuli!

Computational models: the Rescorla-Wagner model

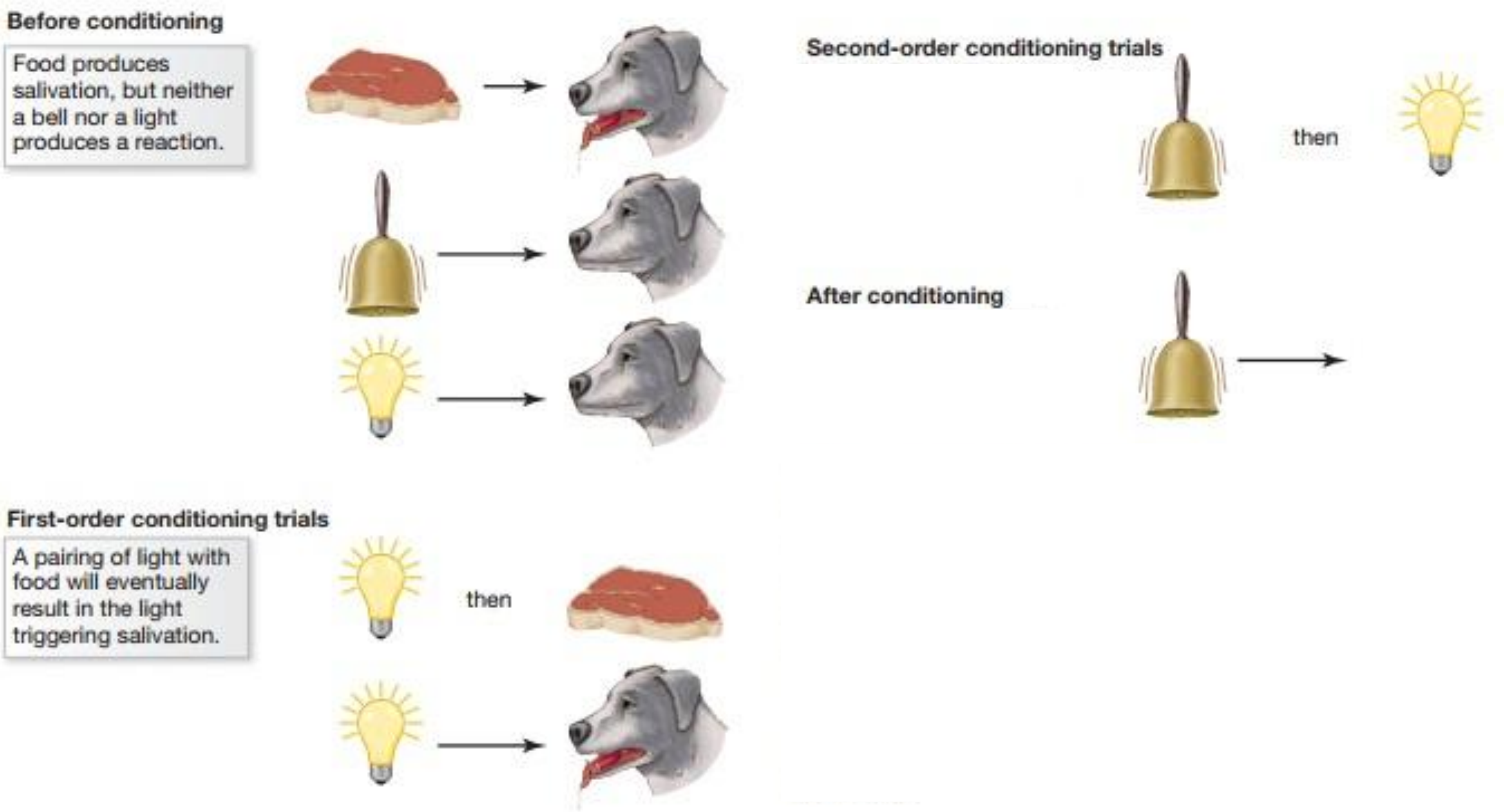
It is a **trial-level** model

- it deals with how associative strengths change from trial to trial without considering any details about what happens within and between trials
- The model does not apply to details about what happens during the time a trial is taking place, or what might happen between trials.
- Within each trial an animal might experience various stimuli whose onsets occur at particular times and that have particular durations. These timing relationships strongly influence learning.

$$\delta_t = R_t - V_t$$



ALMA MATER STUDIORUM UNIVERSITÀ DI BOLOGNA CAMPUS DI CESENA



https://www.brainkart.com/article/Classical-Conditioning--Acquisition-of-Conditioned-Responses_29307/

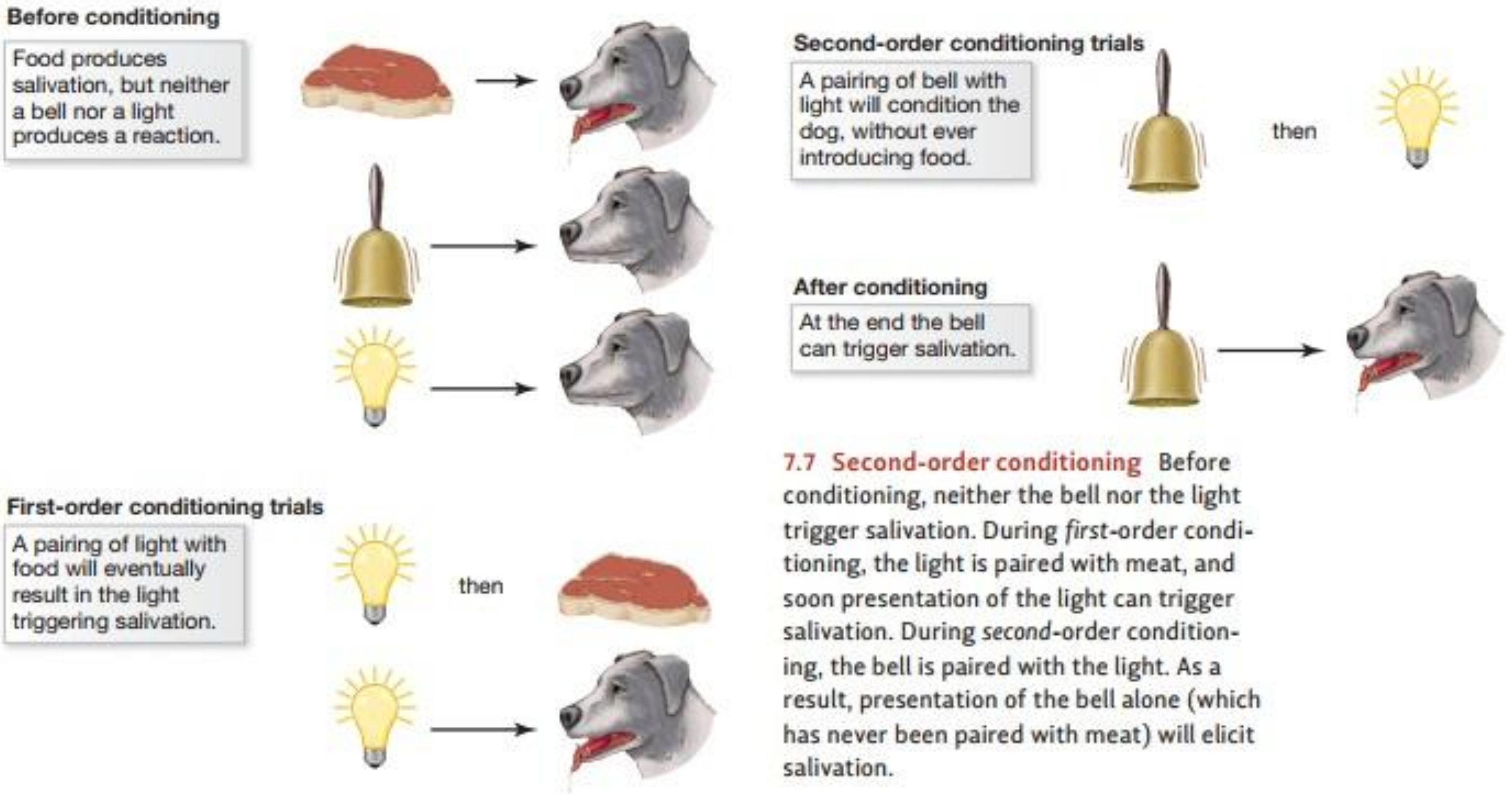
Second-order conditioning

RW model predicts no learning because there is no US in the second-order conditioning trials

BUT

the animal indeed learns

The RW model fails to explain the data



7.7 Second-order conditioning Before conditioning, neither the bell nor the light trigger salivation. During *first*-order conditioning, the light is paired with meat, and soon presentation of the light can trigger salivation. During *second*-order conditioning, the bell is paired with the light. As a result, presentation of the bell alone (which has never been paired with meat) will elicit salivation.

https://www.brainkart.com/article/Classical-Conditioning--Acquisition-of-Conditioned-Responses_29307/

Computational models: the temporal difference model

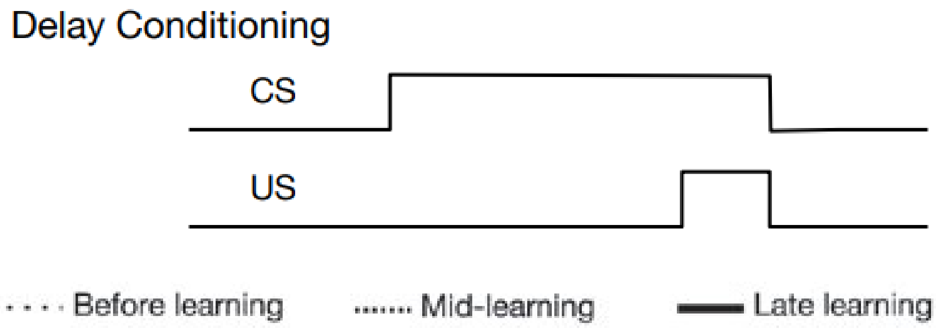
From the point of view of the animal, a trial is just a fragment of its continuing experience interacting with its world

- it is a real-time model
- it now labels time steps within a trial instead of complete trials
- **provide a prediction, for each time *t* in the trial during which a CS is presented, of the total future reward that will be gained in the trial from time *t* to the end of the trial**
- the prediction error here
    - compares the predicted value at time t to the predicted value at time t-1

$$V_{t+1} = V_t + \alpha\delta_t$$

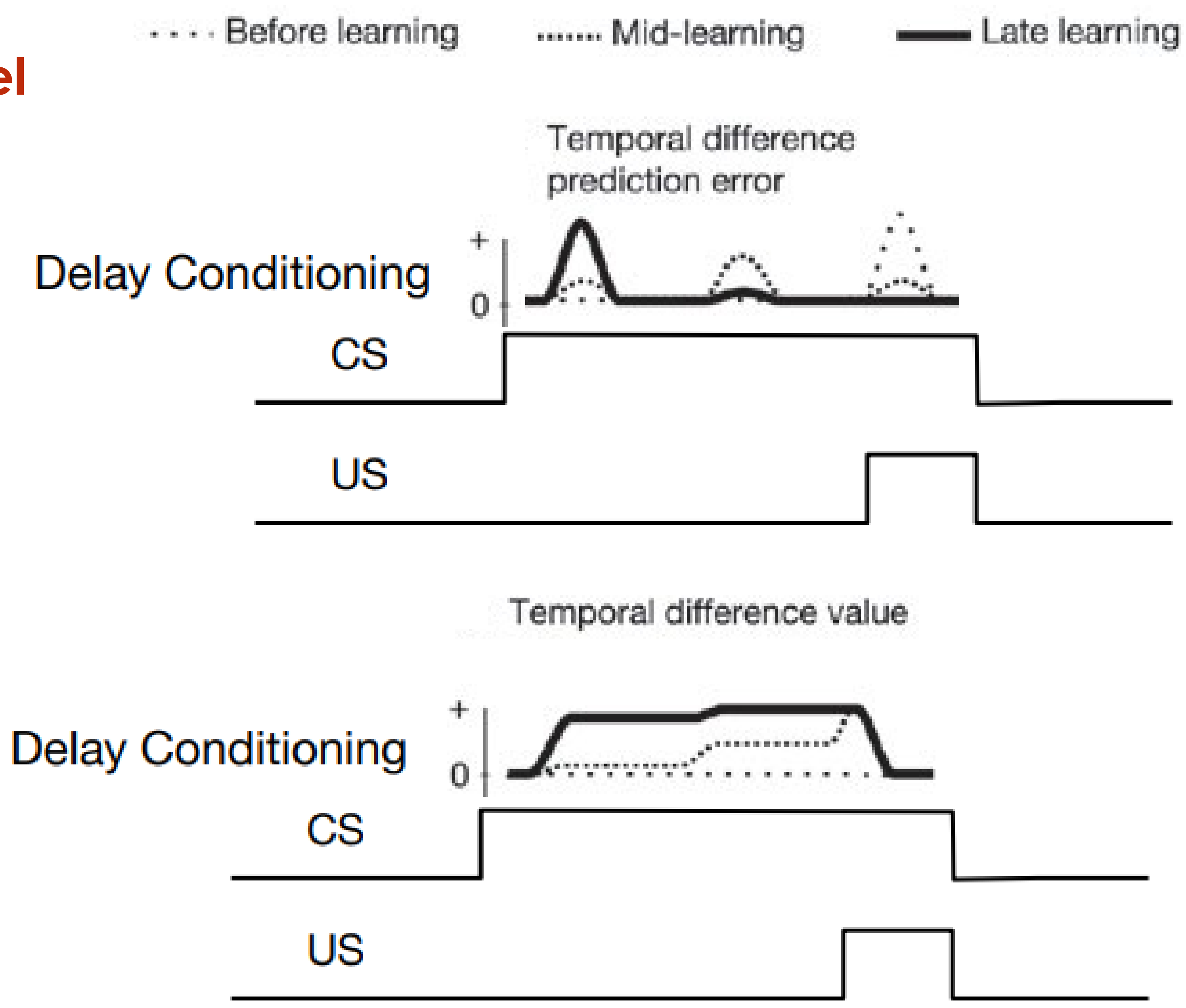
Prediction error

$$\delta_t = R_t + V_t - V_{t-1}$$



Computational models: the temporal difference model

- **provide a prediction, for each time *t* in the trial during which a CS is presented, of the total future reward that will be gained in the trial from time *t* to the end of the trial**
- the prediction error here
    - compares the predicted value at time t to the predicted value at time t-1



Computational models: the temporal difference model

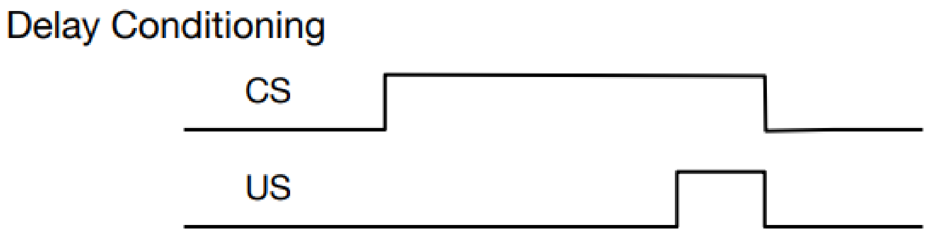
From the point of view of the animal, a trial is just a fragment of its continuing experience interacting with its world

$$V_{t+1} = V_t + \alpha\delta_t$$

- At the **beginning** of learning, the predicted reward ***V*(*t*) is zero for each time *t* until** the time at which the reward or UCS is delivered (t_{UCS})
- On the **next learning trial**, a comparison between $\boldsymbol{V(t_{UCS})}$ **and** $\boldsymbol{V(t_{UCS} - 1)}$ generates a positive prediction error that is used to **increment the V at time** $\boldsymbol{t_{UCS} - 1}$ (in proportion to an arbitrary learning rate).
- On **subsequent learning trials**, ***V*(*t*) is updated for each time *t* ranging from** $\boldsymbol{t_{UCS}}$ **back to** $\boldsymbol{t_{CS}}$ (the earliest time at which the CS is presented).

Prediction error

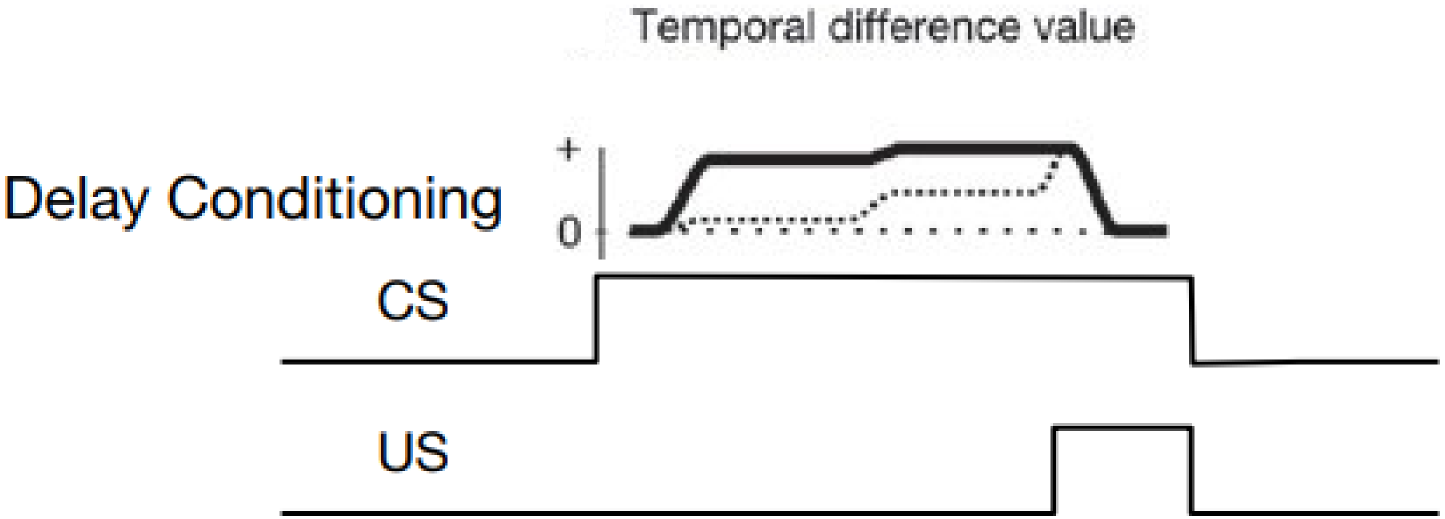
$$\delta_t = R_t + V_t - V_{t-1}$$



Prediction errors in computational models

- **Signed difference between the expected and delivered outcome**
- **Used to update predicted value**
    - The value signal produced by the reward itself transfers back to events that reliably precede reward delivery (e.g. CSs)
    - Thus, **the rewarding value transfers from the reward to the CS that predicts reward**

In this manner, **reinforcement learning algorithms explicitly state that the quantitative value inherent in reward transfers back to the antecedent cue predicting its delivery**. That is, the predictive cue becomes endowed with the scalar value of the reward.



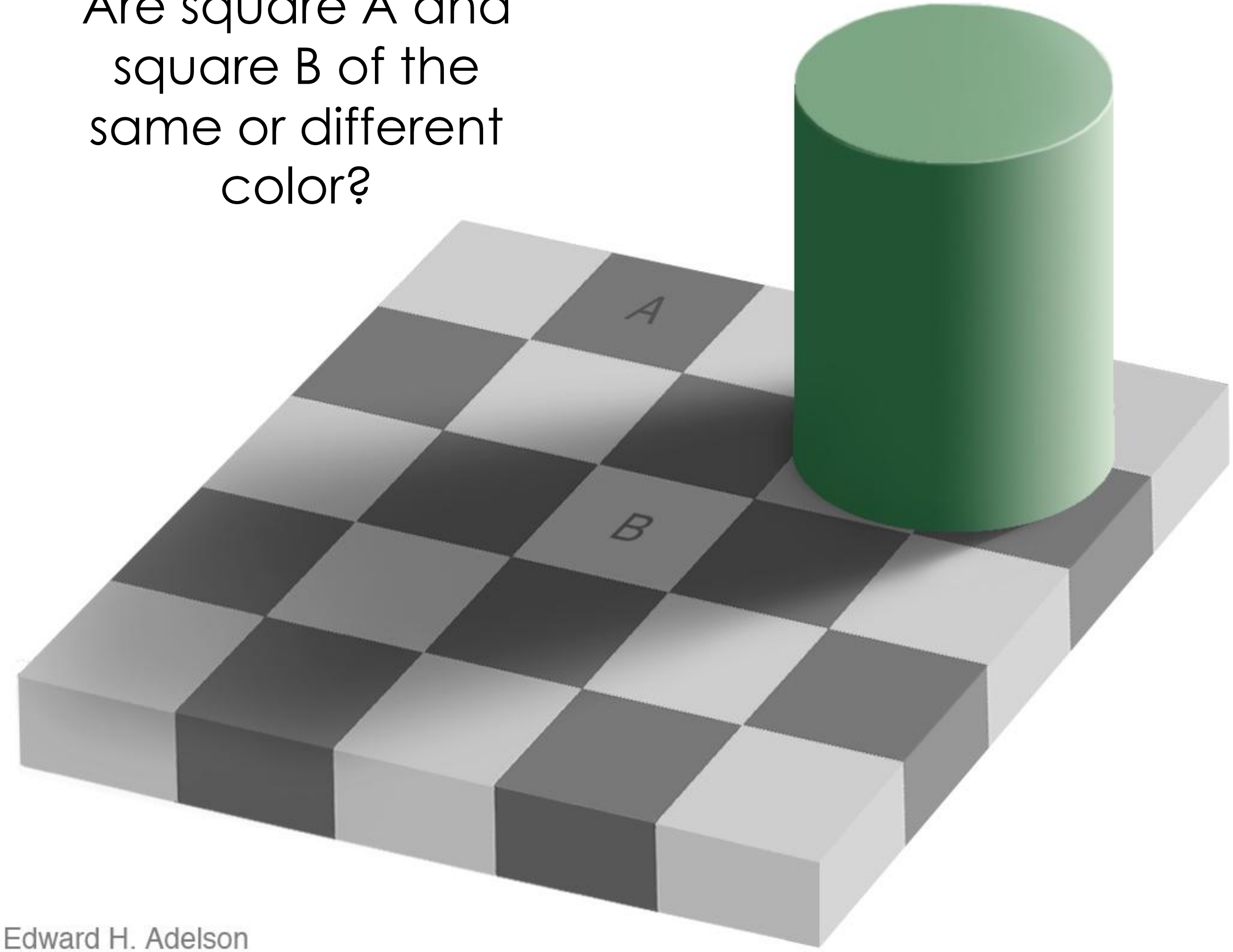
ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA
CAMPUS DI CESENA

wooclap

We make predictions all the time

- Predictions are made also in the realm of sensory perception
- Sensory perception is actively constructed
- Sensory perception reflects the brain's prediction about the causes of sensory signals

Are square A and square B of the same or different color?

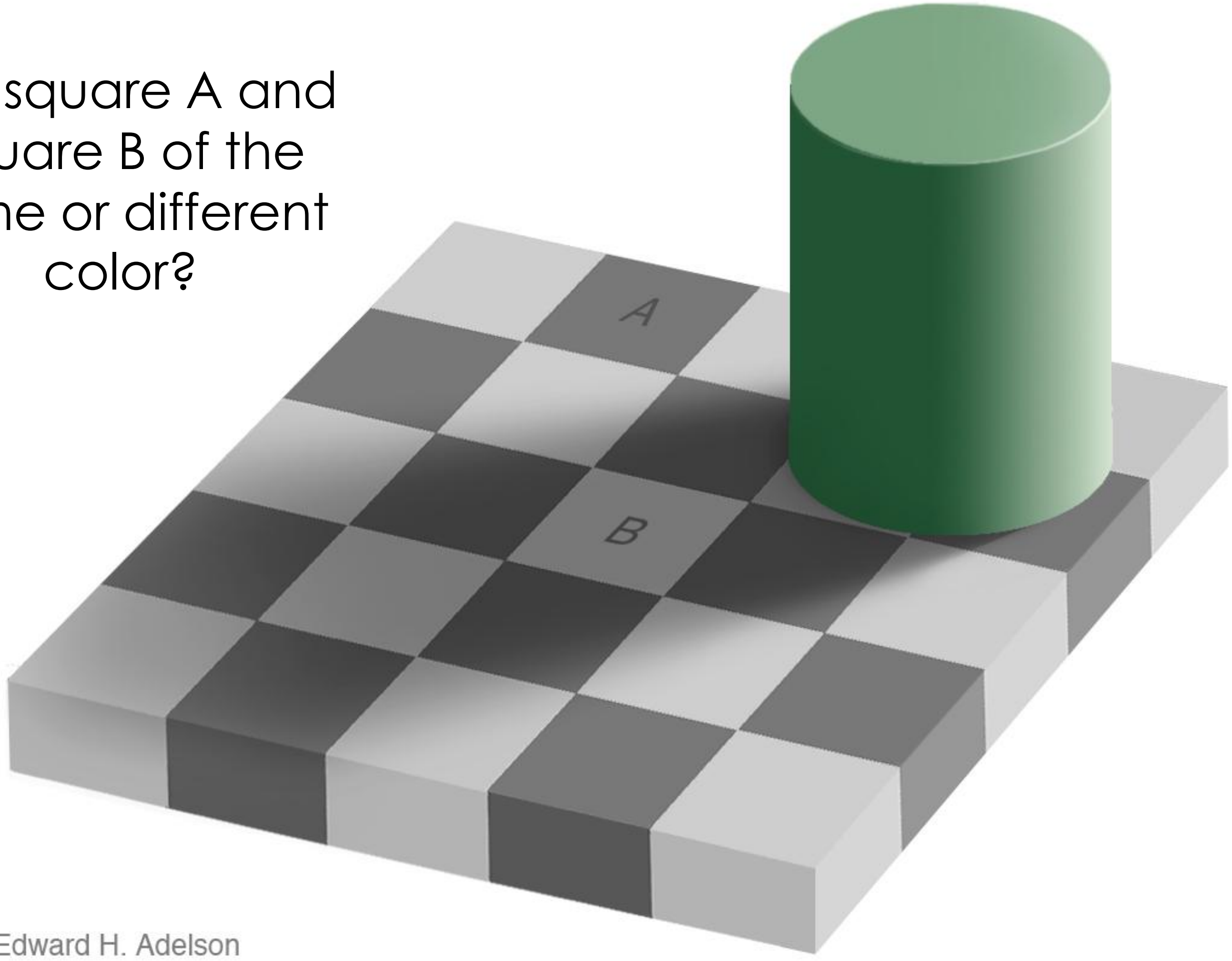


Edward H. Adelson

http://persci.mit.edu/gallery/checkershadow

We make predictions all the time

Are square A and square B of the same or different color?



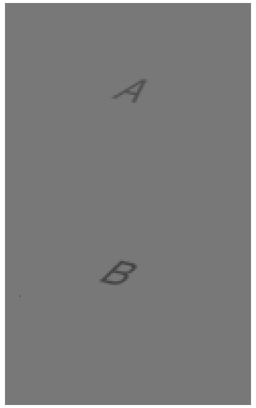
Edward H. Adelson

Let’s test our hypothesis

http://persci.mit.edu/gallery/checkershadow

We make predictions all the time

Sensory perception
is
active
and
predictive

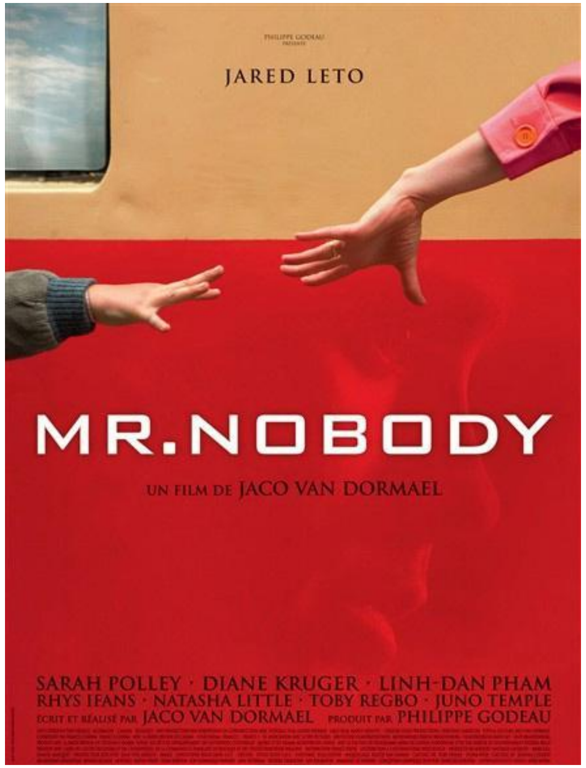


http://persci.mit.edu/gallery/checkershadow

A drawback of the predictive brain...

https://youtu.be/NCtF4aVlxgU

A drawback of the predictive brain...



Skinner, B. F. (1948). 'Superstition' in the pigeon. *Journal of Experimental Psychology, 38*(2), 168–172. https://doi.org/10.1037/h0055873

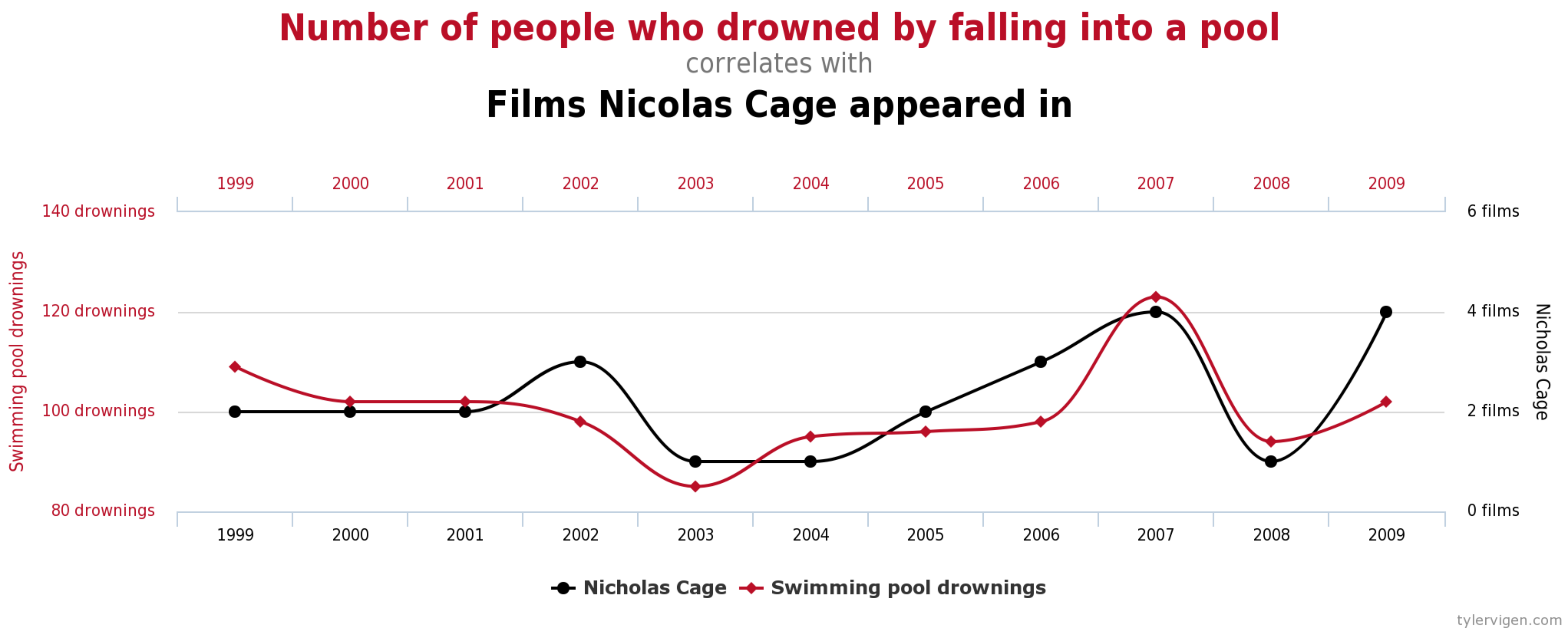
Abstract

"A pigeon is brought to a stable state of hunger by reducing it to 75 percent of its weight when well fed. It is put into an experimental cage for a few minutes each day. A food hopper attached to the cage may be swung into place so that the pigeon can eat from it. A solenoid and a timing relay hold the hopper in place for five sec. at each reinforcement. If a clock is now arranged to present the food hopper at regular intervals *with no reference whatsoever to the bird's behavior*, operant conditioning usually takes place." The bird tends to learn whatever response it is making when the hopper appears. The response may be extinguished and reconditioned. "The experiment might be said to demonstrate a sort of superstition. The bird behaves as if there were a causal relation between its behavior and the presentation of food, although such a relation is lacking." (PsycINFO Database Record (c) 2016 APA, all rights reserved)

https://youtu.be/NCtF4aVlxgU

A drawback of the predictive brain...

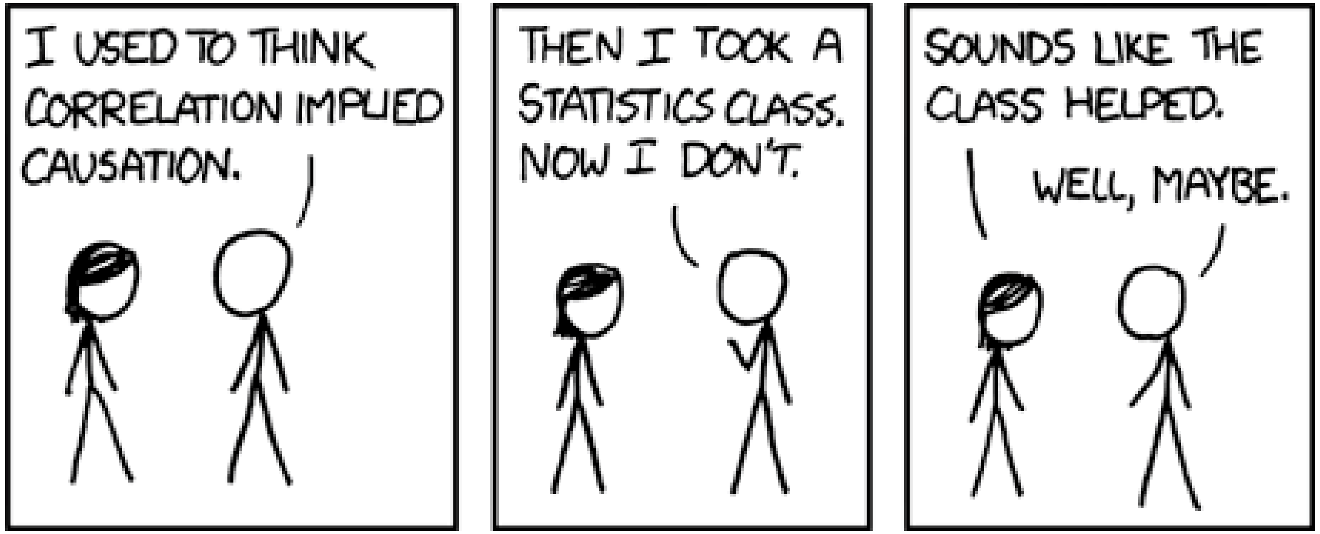
A drawback of the predictive brain...



https://www.tylervigen.com/spurious-correlations

A drawback of the predictive brain...

Correlation is not causation!



Recommended readings

- Daw, N. D., & Tobler, P.N. (2014). Value Learning through Reinforcement: The Basics of Dopamine and Reinforcement Learning. In *Neuroeconomics* (Chapter 15, pp. 283-298). Academic Press.
- Sutton, R. S., & Barto, A. G. (2018). Reinforcement learning: An introduction. MIT press.
    - Chapter 14
        - Sections 14.1, 14.2

Revision questions

Answer the following questions developing a thesis statement and supporting arguments using evidence from the provided source material.

- Critically evaluate the historical shift in understanding the conditions necessary for learning, from an emphasis on contiguity to the recognition of the importance of contingency and surprise. Use specific experimental examples to support your arguments.
- Discuss the role of prediction error in reinforcement learning. How is this concept formalized in computational models like the Rescorla-Wagner model, and what are the implications for understanding how organisms learn from their experiences?
- Compare and contrast the Rescorla-Wagner and Temporal Difference models of learning. In what ways does the TD model offer a more nuanced account of learning, particularly with respect to the temporal aspects of conditioning?
- Explain the blocking effect and its significance for our understanding of reinforcement learning. How does this phenomenon demonstrate that learning is not simply about forming associations between contiguous stimuli? What does it tell us about the role of prediction and surprise?
- [Extra question] Discuss the broader implications of predictive learning beyond classical conditioning, particularly in the domain of sensory perception. Provide examples of how the brain actively constructs our experience by generating and updating predictions about the sensory world.

Glossary of Key Terms

- Conditioned Response (CR): A learned response to a previously neutral stimulus (the conditioned stimulus) that occurs after repeated pairings with an unconditioned stimulus.
- Conditioned Stimulus (CS): A previously neutral stimulus that, after repeated association with an unconditioned stimulus, elicits a conditioned response.
- Contiguity: The temporal closeness between two events, such as a conditioned stimulus and an unconditioned stimulus.
- Contingency: The predictive relationship between two events; the probability of one event occurring given the occurrence (or non-occurrence) of another.
- Delay Conditioning: A classical conditioning procedure in which the onset of the conditioned stimulus precedes that of the unconditioned stimulus and the offset of the conditioned stimulus occurs after unconditioned stimulus onset (i.e. CS and US overlap).
- Extinction: The gradual weakening and eventual disappearance of a conditioned response as a result of the conditioned stimulus being repeatedly presented without the unconditioned stimulus.
- Hebbian Plasticity: The principle that synaptic connections between neurons become stronger when the neurons fire simultaneously ("neurons that fire together, wire together").

Glossary of Key Terms

- Prediction Error: The difference between the expected outcome and the actual outcome. It plays a critical role in driving learning.
- Predictive Learning: The view that learning is fundamentally about forming and updating predictions about future events based on past experiences.
- Rescorla-Wagner Model: A computational model of classical conditioning that explains learning in terms of changes in the associative strength between a CS and a US based on prediction error.
- Second-Order Conditioning: A form of learning in which a neutral stimulus is paired with a previously conditioned stimulus (rather than an unconditioned stimulus) and comes to elicit a conditioned response.
- Surprise: The degree to which an outcome is unexpected or violates existing predictions. Surprise often enhances learning.

Glossary of Key Terms

- Temporal Difference (TD) Model: A real-time computational model of learning that extends the concept of prediction error to individual time steps within a trial, allowing for learning about the timing of events and rewards.
- Trace Conditioning: A classical conditioning procedure in which the conditioned stimulus is presented and terminated before the onset of the unconditioned stimulus, leaving a "trace" interval.
- Truly Random Control Procedure: An experimental design used to demonstrate the importance of contingency in classical conditioning, where a CS and US are presented randomly with respect to each other.
- Unconditioned Response (UR): An innate, automatic response to an unconditioned stimulus.
- Unconditioned Stimulus (US): A stimulus that naturally and automatically elicits an unconditioned response without prior learning.