An Introduction to Learning

Lecture 4/15

Todd M. Gureckis, Lila Davachi
Department of Psychology
New York University
{todd.gureckis,lila.davachi}@nyu.edu
Agenda for Today

- Foundations of Classical Conditioning
- Empirical Phenomena
- Theoretical Accounts
Pavlov and the Dogs
1. Before conditioning
   - Food
     - Response
     - Unconditioned stimulus (US)
     - Unconditioned response (UR)
   - Bell
     - Response
     - Neutral stimulus
     - No salivation
     - No conditioned response

2. During conditioning
   - Bell + Food
     - Response
     - Unconditioned response (UR)

3. After conditioning
   - Bell
     - Salivation
     - Conditioned response (CR)
1 Terminology.

“the beginning of wisdom is to call things by their right name.” - chinese proverb
1 Conditioned Stimulus (CS).

2 Unconditioned Stimulus (US)

3 Unconditioned Response (UR)

4 Conditioned Response (CR)
1 Conditioned Stimulus (CS).

An initially neutral stimulus that later comes to elicit a condition response (e.g., bell, tone)

2 Unconditioned Stimulus (US)

3 Unconditioned Response (UR)

4 Conditioned Response (CR)
1 Conditioned Stimulus (CS).

An initially neutral stimulus that later comes to elicit a condition response (e.g., bell, tone)

2 Unconditioned Stimulus (US)

A stimulus that triggers some innate reflex (e.g., food, shock), also know as the reinforcer (what does that mean?)

3 Unconditioned Response (UR)

4 Conditioned Response (CR)
1 Conditioned Stimulus (CS).

An initially neutral stimulus that later comes to elicit a condition response (e.g., bell, tone)

2 Unconditioned Stimulus (US)

A stimulus that triggers some innate reflex (e.g., food, shock), also know as the *reinforcer* *(what does that mean?)*

3 Unconditioned Response (UR)

The automatic reaction triggered by the unconditioned stimulus (salivation, digestion, freezing/fear, etc...)

4 Conditioned Response (CR)
1 **Conditioned Stimulus (CS).**

An initially neutral stimulus that later comes to elicit a condition response (e.g., bell, tone)

2 **Unconditioned Stimulus (US)**

A stimulus that triggers some innate reflex (e.g., food, shock), also know as the *reinforcer* *(what does that mean?)*

3 **Unconditioned Response (UR)**

The automatic reaction triggered by the unconditioned stimulus (salivation, digestion, freezing/fear, etc...)

4 **Conditioned Response (CR)**

Usually similar to the UR, but is the response made to the CS once it has been effectively conditioned (e.g., salivation, digestion, freezing/fear, etc...
5 Appetitive conditioning

1. Before conditioning
   Food
   \(\rightarrow\)
   Salivation
   Unconditioned stimulus (US)
   Unconditioned response (UR)
   Response

2. During conditioning
   Bell
   \(\rightarrow\)
   Salivation
   Neutral stimulus
   No salivation
   No conditioned response

3. After conditioning
   Salivation
   Unconditioned response (UR)
   Conditioned response (CR)

6 Aversive conditioning
4 Conditioned Response (CR)

**Appetitive** - usually obvious (salivation), orienting response

**Aversive** - musculo-skeletal response (freezing)

*condition emotional response (CER)* - Galvanic skin response, heart rate

Also, disruption of a otherwise normal or ongoing behavior
4 Conditioned Response (CR)

- Estes and Skinner (1941)

First trained rats to press a lever at a particular rate (*fixed interval schedule*) to get a food pellet

- Then for six sessions heard a 3 minute tone (CS) then shock (US)

- Fear response measured indirectly by drop in pressing lever while the CS was on

aka *conditioned suppression*

---

*Figure 3-1.* Idealized diagram showing suppression of a CR (lever pressing) during presentation of a fear inducing CS.

---

Eyeblink Conditioning
A simple, popular paradigm for studying classical conditioning in humans and other animals

B. Conditioned Responses

- **Day 1:** Eyeblink UR
- **Day 3:** Weak Eyeblink CR
- **Day 5:** Strong Eyeblink CR

---

**Original Reflex:**
- CS (Tone)
- US (Airpuff to eye)
- UR (Eyeblink to protect eye)

**Conditioning:**
- CR (Anticipatory blink, at time US is expected to arrive)

**Anticipatory Response:**
- US (Airpuff to eye)
- UR (Eyeblink to protect eye)

---

*Fig. 9-3 A subject wearing the headgear for recording eyelid movement*
Classical Conditioning

1. Before conditioning
   - Food
   - Response
   - Salivation
   - Unconditioned stimulus (US)
   - Unconditioned response (UK)

2. During conditioning
   - Bell
   - Response
   - Salivation
   - Neutral stimulus
   - No conditioned response

3. After conditioning
   - Salivation
   - Conditioned response (CR)

Classical conditioning does not depend on the animals' response. The CR is unavoidable, like a reflex, not utilitarian or flexible; direct result of prediction.

Habituation/Sensitization

(A) Mantle shelf (retracted)
   - Siphon
   - Gill

(B) Sensory neuron
   - Interneurons
   - Excitatory
   - Inhibitory
   - Motor neuron
   - Stimulus
   - Siphon skin
   - Gill

Instrumental conditioning
Why classical condition? What is the function?

Prediction. Classical conditioning allows organism to anticipate future events and make responses appropriate in anticipation of those events.
2 Basic Procedures and Results.
Delay Conditioning

1. Delay conditioning

- CS on for short period
- US on at the end of that period
- Delay from onset of CS to onset of US is the interstimulus interval (ISI)
- Time before next CS onset is intertrial interval (ITI)
Trace Conditioning

2. Trace conditioning

CS and US separated in time leaving a trace interval
Simultaneous Conditioning

3. Simultaneous conditioning

CS and US on together.

Might think this would be most effective since correlation high. Not true.
Backwards Conditioning

4. Backward conditioning

Reverse order of trace conditioning. US on first then CS. A test to see if the order of prediction matters for the acquisition of the conditioned response.
Importance of Control Trials

- Shock or air-puffs might sensitize the organism making them more responsive to many stimuli (including the CS)

- Control trials: **CS alone, US alone, explicitly unpaired CS/US***, or random control
A comment on learning curves

(in classical conditioning this is usually like probability of CR)
A comment on learning curves

(Notice the classical conditioning, this is usually like probability of CR)
A comment on learning curves

(in classical conditioning this is usually like probability of CR)

Steep learning curve??

Number of trials or attempts at learning
4 Some basic phenomena and variables that affect conditioning
The Stimulus and Response

- **Properties of US:** Some USs (e.g., stronger, more intense) lead to faster learning
  - Might reach asymptote in 10 trials in a shock experiment, 100 trials in eyeblink conditioning experiment
  - Also asymptotic CR response rates can differ as a function of stimulus

- **Properties of CS:** more intense, more attention-grabbing, salient are learned better.
  - Species specific differences (people response to visual stimuli more than auditory ones, same for pigeons, less so for rats)
  - Might imply an organization to CS properties that predict learning alone... not the case...
The Stimulus and Response

- **CS-US-CR compatibility:** Different USs are more easily conditioned to certain CSs

- Garcia & Koelling (1966) showed that rats learned to avoid drinking sugar water when injected with drug to make them sick, but not when shocked afterwards (easier to learn **food->sick** than **food->shock**).

  - Also depends on the animal species: pigeons more easily associate color with illness, rats associate outcomes with flavors

  - Evolutionary Adaptive constraints for learning particular forms of association
The Stimulus and Response

FIGURE 4-3. Photograph showing pigeons pecking keys when receiving water reward (left) or food (right). Pecks to the water-related key included drinking-like movements such as licking. Pecks to the food-related key were made with the beak slightly open, as if to seize a piece of grain.

Photo Courtesy of Bruce Moore
Timing

(A)

Conditioned responding

CS-US interval (msec)

(B)

1,000 msec
Tone CS
Air puff US
Blink UR
CR

750 msec
Tone CS
Air puff US
CR

Shorten the ISI...
Presenting the CS alone repeatedly leads to a decrease in the CR (a process known as extinction).

Like habituation in that loss of sensitivity to a stimulus on repeated presentation. Key difference is if the habituated response is a kind of innate, reflex type behavior or a previously learned association.
Is extinction just forgetting?

No, much faster than the forgetting that happened in classical conditioning (only a few trials of CS alone needed to remove response)
Is extinction “unlearning”? 

- Some evidence suggests no.
- Even if response completely extinguished in one session, can return the next day. (Spontaneous Recovery)
- Also appears context dependent (e.g., Bouton & Bolles, 1979) ABA renewal (training context A, extinguish in context B, return to A and immediately recover CR)
Pavlov’s hypothesis was that there were both **excitatory** and **inhibitory** processes going on. Extinction involved the creation of inhibitory links, which (critically) are more fragile/easy to forget which explains recovery in each session.

Another view is that attention to the CS drops on non-paired trials. Eventually the animal just ignores the CS. In new session/next day the CS is interesting again and grabs attention.
Is extinction “unlearning”?

- **Disinhibition**: After extinction some other arousing stimulus can reactive the effectiveness of the CS in eliciting the CR. Learn tone->eyeblink, extinguish, then ring loud bell. The tone becomes a effective cue for the CR again. (could be attentional story again)

- **Rapid reacquisition**: relearning is fast for previously learned then extinguished pairings. This is true even if you extinguish out so that spontaneous recovery doesn’t happen any more.
Intermittent Pairing

Intermittent pairing effects primarily the asymptotic level of reinforcement.

Effect seems to be primarily on motivation rather than on the strength of the association.

Figure 3-3. Mean percentage of conditioned eyeblinks in successive blocks of ten trials for each of the experimental groups in Ross's study. Group C had continuous pairing of the CS and US throughout training. Group I had 50 percent intermittent pairing throughout. Group C 100 I switched from continuous to 50 percent intermittent pairing after 100 trials. Group C 20 I switched from continuous to 50 percent intermittent pairing after 20 trials. Group I 40 C switched from 50 percent intermittent to continuous pairing after 40 trials. (Copyright 1959 by the American Psychological Association. Reprinted by Permission.)
Conditioned Inhibition

Things can be conditioned to STOP the CR too.

First...

Tone (CS+) → Food (US)

Then...

Tone (CS+) & Light (CS-) → no food

Tone along still controls salivary response, but Tone+Light compound doesn’t (summation test). Light is a “conditioned inhibitor”
Generalization & Discrimination

- Conditioning with one stimulus tends to spread to similar ones (like tone at particular frequency)

- Shape of generalization can be warped by the specifics of training. If certain frequencies predict food, and others are presented in training that don’t, animal will discriminate by only showing CR to particular CS.
4 What IS learned in classical conditioning?
**Artistotelian Concept of “Associations”**

A linkage between pairs of events, sensations, or ideas such that the activation or elicitation of one (EVENT A) results in the retrieval or anticipation of the other (EVENT B).

Three key principals

- **Contiguity**: Events experience at similar point in time or space are more easily associated

- **Frequency**: the more often we experience something, the more strongly we associate them

- **Similarity**: when two things that are similar, thinking of one can trigger the other (e.g., chair/table, dog/cat, iPad/awesomeness)
What IS learned in classical conditioning?

Pavlov’s *Stimulus Substitution* idea - The CS becomes a substitute for the US in invoking the UR (stimulus-response learning)
An alternative view

**acquired**

* S-S association

Stimulus-Stimulus (SS) associations - The CS becomes associated with the US. Presentation of the CS makes the subject “think of” the US which in turns activates the CR.
Ways to tell them apart

- Use mechanical means or drugs to induce paralysis (can’t execute response). In these situations, learning still happens by evidence of the CR presence after drugs/mechanical device removed.

- **Sensory preconditioning** - Take two stimuli (S1 and S2, say tone and light) and repeatedly pair them without US in a pre-training phase. Later associate one of the paid (S1) to a US (shock). Later test S2, find evidence of transfer of CR to the other stimulus. This is further evidence of some kind of S-S relationships.

- **Devaluation studies** - Rescorla (1973) first trained rats to press a lever for rewards (instrumental conditioning). Then, training a light to predict a loud sound causing them to stop pressing the reward button for a few seconds. Then devalued the US by repeatedly presenting the noise. Then measured pressing response to light. Rats who were exposed to noise continued pressing suggesting mediation by the US in the CR response (stimulus-stimulus).
A **linkage between pairs of events, sensations, or ideas** such that the activation or elicitation of one (EVENT A) results in the retrieval or anticipation of the other (EVENT B).

**Three key principals**

- **Contiguity**: Events experienced at similar point in time or space are more easily associated

- **Frequency**: the more often we experience something, the more strongly we associate them

- **Similarity**: when two things that are similar, thinking of one can trigger the other (e.g., chair/table, dog/cat, iPad/awesomeness)
Is contiguity sufficient?  
from Rescorla (1988)

We start out by taking an unconditioned stimulus (UCS) that produces the desired response without training. . . . We pair the UCS with a conditioned stimulus (CS). . . . This procedure, when repeated several times . . . will ultimately result in the occurrence of the response following the CS alone. (Klatsky, 1980, p. 281)

The essential operation in conditioning is a pairing of two stimuli. One, initially neutral in that it elicits no response, is called the conditioned stimulus (CS); the other, which is one that consistently elicits a response, is called the unconditioned stimulus (US). The response elicited by the unconditioned stimulus is the unconditioned response (UR). As a result of the pairing of the conditioned stimulus (CS) and the unconditioned stimulus (US), the previously neutral conditioned stimulus comes to elicit the response. Then it is called the conditioned response (CR). (Morgan & King, 1966, pp. 79–80)

A widely used developmental text agrees, calling conditioning a “form of learning in which a neutral stimulus, when paired repeatedly with an unconditioned stimulus, eventually comes to evoke the original response” (Gardner, 1982, p. 594). Similarly, a best-selling textbook of abnormal psychology describes a conditioned stimulus as “a stimulus that, because of its having been paired with another stimulus (unconditioned stimulus) that naturally provokes an unconditioned response, is eventually able to evoke that response” (Rosenhan & Seligman, 1984, p. 669).
Is contiguity sufficient?

from Rescorla (1988)
Is contiguity sufficient?
from Rescorla (1988)

Figure 1
Schematic of Two Conditioned Stimulus/Unconditioned Stimulus (CS/US) Relations That Share the Same Contiguity but Differ in the Information the CS Gives About the US

P(US|CS) < P(US | no CS)
Is contingency sufficient?
from Rescorla (1988)

Table 7.2

<table>
<thead>
<tr>
<th>The Blocking Paradigm</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Group</strong></td>
</tr>
<tr>
<td>Control group</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Experimental “pre-trained” group</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>
Is contingency sufficient?

from Rescorla (1988)

**Table 7.2**

<table>
<thead>
<tr>
<th>Group</th>
<th>Phase 1</th>
<th>Phase 2</th>
<th>Phase 3 (test)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control group</td>
<td>Rat sits in chamber; no training</td>
<td>Tone CS combined with light CS — shock US</td>
<td>Tone CS → medium CR</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Light CS — shock US</td>
<td>Light CS → medium CR</td>
</tr>
<tr>
<td>Experimental “pre-trained” group</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Light CS — shock US</td>
<td>Tone CS combined with light CS — shock US</td>
<td>Tone CS → little or no CR</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Light CS →</td>
<td>Need surprise!</td>
</tr>
</tbody>
</table>

Kamin’s (1969) Blocking Effect!

\[ P(US|CS_1 + CS_2) \neq P(US | CS_1) \]
Are contingency and surprise sufficient?

Reynold’s overshadowing effect

thanks to yael niv who thanks randy gallistel
Are contingency and surprise sufficient?

Stimuli compete

Reynold’s overshadowing effect

Not contingency!
Need surprise!
Even then may not learn!

thanks to yael niv who thanks randy gallistel
A Learning Theory

The secret to understanding learning is...
SURPRISE!

IT'S CATURDAY!
No, seriously, *surprise*. That’s the secret.
A Learning Theory
Rescorla & Wagner (1972)

- One of the most influential theories of learning. Accounts for gobs of data on how people and animals learn.

- Independently discovered twice (Rescorla-Wagner - working on animal learning/conditioning; Widrow-Hoff working on adaptive filtering).

- As noted in the Rescorla (1988) reading, the basic principal (also instantiated as the “delta-rule”) strongly influenced contemporary “connectionist” modeling approaches.

- At the same time, now viewed (by some) as the kind of null model. If you finding some learning phenomena not easily accounted for by RW theory, you have a paper since basically everyone things RW is right!

- So what is it? A mathematical model of the trial-by-trial changes in expectation for future events that result in a conditioning experiment.
3 A Theory of Pavlovian Conditioning: Variations in the Effectiveness of Reinforcement and Nonreinforcement

THE BASIC THEORY

The generalization which applies to all of the results in the previous section is that the effect of a reinforcement or nonreinforcement in changing the associative strength of a stimulus depends upon the existing associative strength, not only of that stimulus, but also of other stimuli concurrently present. It appears that the changes in associative strength of a stimulus as a result of a trial can be well-predicted from the composite strength resulting from all stimuli present on that trial. If this composite strength is low, the ability of a reinforcement to produce increments in the strength of component stimuli will be high; if the composite strength is high, reinforcement will be relatively less effective. Similar generalizations appear to govern the effectiveness of a nonreinforced stimulus presentation. If the composite associative strength of a stimulus compound is high, then the degree to which a nonreinforced presentation will produce decrements in the associative strength of the components will be large; if the composite strength is low, the effect of a nonreinforcement will be reduced.
What does it mean?
Key assumptions:

Assumption 1: Each CS has an association weight between it and the US which we can represent as:

\[ V_{s_1} \quad V_{s_2} \quad V_{\text{light}} \quad V_{\text{tone}} \]

Assumption 2: Prediction of the US is the sum of all the weights presented during a trial:

\[ V_{s_1s_2} = V_{s_1} + V_{s_2} \quad V_{\text{light,tone}} = V_{\text{light}} + V_{\text{tone}} \]

Assumption 3: Learning on each trial is proportional to the difference between the outcome the animal expects and what actually occurs. This is the prediction error that measures how surprising the outcome on any trial actually was. ERROR DRIVES LEARNING!

\[ \delta = R_{US} - V_{s_1s_2} \]
Rescorla & Wagner (1972)

- Prediction error all the way!

\[ \delta = R_{US} - V_{s_1,s_2} \]

Prediction error = actual US - expected US
Learning for each component is given by

\[ \Delta V_{s_1} = \alpha_{s_1} \beta_1 (R_{US} - V_{s_1,s_2}) \]

\[ \Delta V_{s_2} = \alpha_{s_2} \beta_1 (R_{US} - V_{s_1,s_2}) \]
Learning for each component is given by

\[ \Delta V_{s_1} = \alpha_{s_1} \beta_1 \delta \]

\[ \Delta V_{s_2} = \alpha_{s_2} \beta_1 \delta \]
Learning for each component is given by

\[ \Delta V_{s_1} = \alpha_{s_1} \beta_1 \delta \]

\[ \Delta V_{s_2} = \alpha_{s_2} \beta_1 \delta \]

\( \alpha \) = learning rate for s_x  \quad \beta = learning rate for US
Rescorla & Wagner (1972)

Learning for each component is given by

\[ \Delta V_{s_1} = \alpha_{s_1} \delta \]
\[ \Delta V_{s_2} = \alpha_{s_2} \delta \]

\( \alpha \) = learning rate for \( s_x \)
A few examples

Rescorla & Wagner (1972)

(play with RW in excel)
The value of the CS at any point in time (T) is an weighted average of the rewards on previous trials ($R_i$).

**RECENCY WEIGHTED** - more recent rewards are given more weight/emphasis in current predictions.
<table>
<thead>
<tr>
<th>phenomena explained</th>
<th>RW</th>
<th>a better theory?</th>
</tr>
</thead>
<tbody>
<tr>
<td>acquisition curves</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>extinction curves</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>blocking</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>overshadowing</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>temporal relationships</td>
<td>✗</td>
<td></td>
</tr>
<tr>
<td>overexpectation</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>second-order conditioning</td>
<td>✗</td>
<td></td>
</tr>
<tr>
<td>latent inhibition</td>
<td>✗</td>
<td></td>
</tr>
</tbody>
</table>
Problems? NOOOO!
Rescorla-wagner all the way!
Latent Inhibition

Phase 1

Phase 2

Result: Unpaired presentations of light lead to shock

No learning/prediction error on unpaired trials? RW can’t learn anything then.
Modulations in CS Effectiveness: The Pearce-Hall Model

One possible explanation of latent inhibition was actually implied in our discussion of the RW model (the possibility that the “effectiveness” of the CS could be modulated by the setting of alpha.

Example (Wagner, 1978):

associability. Just as a fully expected US loses its ability to reinforce conditioning, so a fully expected CS loses its ability to become associated with a US. More formally, we may say that the parameter $\alpha$ may change as a result of training, and if $l$ is the asymptotic strength of the context-CS association and $v$ is the current strength of this association, then $\alpha$ is given by Equation 3 as

$$\alpha = \delta(l - v) \quad (3)$$

where $\delta$ is a learning-rate parameter determined by the properties of the CS, which takes a value between 0 and 1. By substituting for $\alpha$ in Equation 2, we derive

$$\Delta V_\alpha = \delta(l - v)(\lambda - V_\Sigma), \quad (4)$$

which describes how changes in $\alpha$ can modify the course of conditioning.$^2$
Key idea: ongoing changes in the “Associability of stimuli”

The associability of a new stimulus is

\[ \alpha_{A}^{n} = \lambda^{n-1} - V_{A}^{n-1} \].

Learning is a function of the associability of the CS and it’s salience, and the salience of the US

\[ \Delta V_{A} = S_{A} \alpha_{A} \lambda, \quad \Delta V_{A}^{n} = S_{A} \left| \lambda^{n-1} - V_{A}^{n-1} \right| \lambda^{n} \].

Very different account... not about changes in US effectiveness (by virtue of being predictable) but in changes in the effectiveness of the CS itself
As Va approaches lambda, the associability of the CS drops to zero, making smaller and smaller changes (just like in RW)
A Model for Pavlovian Learning: Variations in the Effectiveness of Conditioned But Not of Unconditioned Stimuli

John M. Pearce
University of Cambridge, Cambridge, England

Geoffrey Hall
University of York, York, England

Explains latent Inhibition

\[ \Delta V_A^n = S_A \left| \lambda^{n-1} - V_{A}^{n-1} \right| \lambda^n, \]

Since \( \lambda^n = 0 \) during first phase, \( V_a \) will drop to zero and thus will differ from a novel stimulus (assuming alpha is nonzero for the novel stimulus of course)
In phase 1, stimulus A will approach value of lambda. On presentation of the compound stimulus AB, the associability will be low for all but the first trial, meaning that blocking is less successful on the first trial than on subsequent trials. (Basic story is similar to RW in this way)
Second-order conditioning

Why does the RW/PH theory have trouble with this?

**Special status of the US stimulus... it must be the thing linked to the UR**
Timing

RW/PH are silent on issues of timing.
Next time

A more modern approach grounded in computational theory, and some neuroscience
The take-home

• **Learning is about predictions** (as we will see classical conditioning is simply one place to measure this)

• **Predictive learning is an error-correcting process.** We build up representations of what we EXPECT to happen in the world and deviations from those expectations drive learning

• **Rescorla-Wagner is a powerful learning model which explains a lot and is infinitely useful**

• **While highly successful, the RW models misses some key phenomena (latent inhibition, second-order conditioning, effects of timing within a trial) that are topics of current and more modern research**

• **An alternative account (Pearce-Hall model) assumes that learning is the result of changes in the associability of the CS (or the effectiveness of the CS). This is highly relevant to contemporary work on attentional factor in learning.**
Readings


References for Slides


Lecture notes from Yael Niv (http://www.princeton.edu/~yael/PSY338/index.html)

The interweb.