Learning II

Factors that influence CC
- Eye-blink conditioning
- CS Intensity
  - Loud vs. Soft Tone
- US Intensity
  - Hard vs. Soft Puff
- Anxiety
  - High vs. Low

Stimulus Generalization
- Conditioned responses (CRs) occurring to stimuli other than the CS used for training
- Similarity
  - The more similar the second stimulus is to the CS the more generalization will occur
- Critical feature of learning
  - we rarely encounter the exact same stimulus twice

Discrimination
- CS+ ---> UCS
- CS- ---> ____
- The stimuli will come to control responding:
  - The CS+ will elicit a CR
  - The CS- will not elicit a CR

Second-Order Conditioning
- Phase 1:
  - Pair CS₁ --> UCS until learning occurs
- Phase 2:
  - Pair a new stimulus (CS₂) as the CS with the first one (CS₁) as the UCS
  - CS₂ --> CS₁
- Because CS₁ reliably elicits a CR, the new stimulus, CS₂ that is paired with it, will begin to elicit the CR as well

Discrimination
- Hypothetical example
- Initially the organism responds to both stimuli
  - shows generalization

Stimulus Generalization
- The more similar the second stimulus is to the CS the more generalization will occur
- we rarely encounter the exact same stimulus twice
Second-Order Conditioning

- From Rizley & Rescorla, 1972
  - CS1: flashing light
  - CS2: 1800 Hz tone
  - UCS: (shock)
- Phase 1:
  - CS1 --> UCS (8 Trials)
- Phase 2:
  - CS2 --> CS1

Stewart & colleagues (1987)

- Slides
  - Neutral scenes
  - Pleasant scenes
  - Various products
- Experimental group
  - Brand L toothpaste always followed by pleasant scenes
- Control group
  - Brand L toothpaste always followed by neutral scenes
- Experimental students rated Brand L significantly more positively than the Control group did

Taste Aversion

- Chemotherapy
- Give children distinctive-flavored Lifesaver candy (CS) between their evening meal and the chemo session (UCS)
  - 12/15 children ate the food at the meal again later
- Control: no lifesaver
  - 6/15 children would eat that meal again

Treating Phobias

- Peter Jones (1924) brought a rabbit into the same room but far away from Peter while he was eating his cookies and milk snack
  - Rabbit: CS that elicits anxiety
  - Snack: CS that elicits good feelings
- Brought the rabbit closer and closer until there was no fear to the rabbit
  - Eventually the rabbit was put into his lap

Counter-conditioning

- CS presented at the same time as another event that elicits an incompatible response
- Systematic Desensitization (Joseph Wolpe, 1958)
  - Train person in deep relaxation
    - separately
  - Create hierarchy of fear eliciting stimuli
    - from least to most strong example of stimulus
    - Imaginal or in vivo desensitization
  - Pair each item of hierarchy with relaxation
    - without producing fear
    - combines counter-conditioning, generalization, and extinction

Advertising

- Pair products with stimuli that elicit positive emotions
- Second-order conditioning

Counter-conditioning
What leads to conditioning?

- **Contiguity**
  - Stimuli that are close to one another in time and in space become associated
- **Co-occurrence**
  - Proximity critical
- **Contingency**
  - When one stimulus depends on the other, they will become associated
- **Information**
  - Predictive value critical

Overshadowing

- When conditioning involves a *compound stimulus*, one stimulus may acquire more stimulus control than the other
- More salient stimulus interferes with conditioning to less salient one

Overshadowing

- Grice & Hunter, 1964
- Human eyeblink conditioning
- 3 Groups:
  - 100 trials w/ CS (loud tone)
  - 100 trials w/ CS (soft tone)
  - 50 trials w/ CS (loud tone) & 50 trials w/ CS (soft tone)

Which one?

- **Contiguity**
- **Contingency**

Blocking

- Phase 1: Pair CS1 → UCS
- Phase 2: Pair compound stimulus with UCS: CS1CS2 → UCS
- Phase 3: Test element stimuli alone to determine amount of conditioning
  - Conditioning to CS1 will be strong, but conditioning to CS2 will be weak: Blocking

Kamin (1968)

<table>
<thead>
<tr>
<th>Group</th>
<th>Phase 1</th>
<th>Phase 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>16 Trials</td>
<td>8 Trials</td>
</tr>
<tr>
<td>Nonreinforced</td>
<td>Results</td>
<td>Supp. Ratio</td>
</tr>
<tr>
<td>NL</td>
<td>L</td>
<td>.05</td>
</tr>
<tr>
<td>N-NL</td>
<td>L</td>
<td>.45</td>
</tr>
<tr>
<td>N only</td>
<td>L</td>
<td>.44</td>
</tr>
</tbody>
</table>

Suppression Ratio of 0 indicates complete fear conditioning, 0.5 indicates chance behavior (no conditioning).
Contiguity or Contingency?

Contiguity
- Both CSs were paired with a UCS in the Blocking procedure
  - BUT one of the CSs was not learned

Contingency
- CS that was most reliably associated with UCS was learned

Predictive Value: Alarms

Contingency
- Kamin’s study:
  - Group N-NL received 24 shocks during acquisition
    - $p(\text{shock} / \text{Noise}) = 24 / 24 = 1.0$ and
    - $p(\text{shock} / \text{No Noise}) = 0 / 24 = 0$
  - Group NL received 8 shocks during acquisition
    - $p(\text{shock} / \text{Noise}) = 8 / 8 = 1.0$ and
    - $p(\text{shock} / \text{No Noise}) = 0 / 8 = 0$
- Predictive value of noise?
  - Noise is a great predictor for both groups

Contingency
- In other words, a CS is only good as a predictor if the UCS occurs fairly often in the presence of the CS but not very often in its absence

Contingency
- Kamin’s study:
  - Group N-NL received 24 shocks during acquisition;
    - $p(\text{shock} / \text{Light}) = 8 / 24 = .33$ and
    - $p(\text{shock} / \text{No Light}) = 16 / 24 = .67$
  - Group NL received 8 shocks during acquisition;
    - $p(\text{shock} / \text{Light}) = 8 / 8 = 1.0$ and
    - $p(\text{shock} / \text{No Light}) = 0 / 8 = 0$
- Predictive value of Light?
  - Light is a great predictor for Group NL
  - Light is a poor predictor for Group N-NL
  - Consequently, little learning for Light in N-NL group!

Kamin’s (1968) Blocking Study
Latent Inhibition

- **Phase 1:**
  - Present CS alone for several trials

- **Phase 2:**
  - CS \(\rightarrow\) UCS for a limited # of trials

- **Test Phase:**
  - CS \(\rightarrow\) ____ to see if conditioning occurred to the CS

---

Latent Inhibition

- **CER Procedure:**
  - Phase 1: Train thirsty rats to drink from tube
  - Phase 2: Separately present Tone during 3 Sessions; Controls had no Tone while in box
  - Phase 3: All rats had Tone \(\rightarrow\) Shock pairings
  - Test Phase: Present Tone while rats were drinking from water tube

---

Co-occurrence vs. Contingency

<table>
<thead>
<tr>
<th>Group</th>
<th>Probability that US follows CS</th>
<th>Probability that US occurs by itself</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td>0.8</td>
<td>0.8</td>
</tr>
<tr>
<td>(2)</td>
<td>0.8</td>
<td>0.4</td>
</tr>
<tr>
<td>(3)</td>
<td>0.4</td>
<td>0.4</td>
</tr>
<tr>
<td>(4)</td>
<td>0.4</td>
<td>0.0</td>
</tr>
</tbody>
</table>

- **Rescorla-Wagner Theory (1972):**
  - Organisms only learn when events violate their expectations
  - Expectations built up when ‘significant’ events follow a stimulus complex
  - Expectations modified when consequent events disagree with the composite expectation

- **Rescorla-Wagner Model:**
  - Change in associative strength of a stimulus depends on
    - Existing associative strength of that stimulus
    - Associative strength of all other stimuli present
  - Change depends on level of existing associative strength
    - If low, potential change is high
    - If high, very little change occurs
  - Speed and asymptotic level of learning determined by strength of the CS and UCS
Rescorla-Wagner Model

- Rescorla and Wagner used a mathematical model to make their "cognitive" account more rigorous.
  - \( \Delta V_a = \alpha \beta [\lambda - V_{ax}] \)
  - \( \Delta V_a \): Change in associative strength to CS
  - \( V_{ax} \): Current associative strength to CS (context)
  - \( \alpha \): Salience of CS
  - \( \beta \): Salience of UCS used in the experiment
  - \( \lambda \): Maximum associative strength possible

Before conditioning begins:
- \( \lambda = 100 \) (number is arbitrary & based on the strength of the UCS)
- \( V_{ax} = 0 \) (because no conditioning has occurred)
- \( c = .5 \) (c must be a number between 0 and 1.0 and is a result of multiplying the CS intensity by the UCS intensity)

First Conditioning Trial

<table>
<thead>
<tr>
<th>Trial</th>
<th>( c )</th>
<th>( (\lambda - V_{ax}) )</th>
<th>( \Delta V_a )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.5</td>
<td>(100 - 0)</td>
<td>50</td>
</tr>
</tbody>
</table>

\[ c \times (\lambda - V_{ax}) = \Delta V_a \]

Second Conditioning Trial

<table>
<thead>
<tr>
<th>Trial</th>
<th>( c )</th>
<th>( (\lambda - V_{ax}) )</th>
<th>( \Delta V_a )</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>.5</td>
<td>(100 - 50)</td>
<td>25</td>
</tr>
</tbody>
</table>

\[ c \times (\lambda - V_{ax}) = \Delta V_a \]

Third Conditioning Trial

<table>
<thead>
<tr>
<th>Trial</th>
<th>( c )</th>
<th>( (\lambda - V_{ax}) )</th>
<th>( \Delta V_a )</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>.5</td>
<td>(100 - 75)</td>
<td>12.5</td>
</tr>
</tbody>
</table>

\[ c \times (\lambda - V_{ax}) = \Delta V_a \]

8th Conditioning Trial

<table>
<thead>
<tr>
<th>Trial</th>
<th>( c )</th>
<th>( (\lambda - V_{ax}) )</th>
<th>( \Delta V_{cs} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>.5</td>
<td>(100 - 99.22)</td>
<td>.39</td>
</tr>
</tbody>
</table>

\[ c \times (\lambda - V_{ax}) = \Delta V_{cs} \]
1st Extinction Trial

\[
\text{Trial } 1: \quad c_1 \ (\lambda - V_{ax}) = \Delta V_{cs_1} = -49.8
\]

2nd Extinction Trial

\[
\text{Trial } 2: \quad c_2 \ (\lambda - V_{ax}) = \Delta V_{cs_2} = -24.9
\]

Rescorla-Wagner Model

- Describes acquisition and extinction of a conditioned response
- Many other learning phenomena, too!

Rescorla-Wagner

- Overshadowing
  - When multiple stimuli or compound stimulus:
    - \( V_{ax} = V_{cs_1} + V_{cs_2} \)
    - Trial 1:
      - \( \Delta V_{noise} = .2 \times (100 - 0) = (.2)(100) = 20 \)
      - \( \Delta V_{light} = .3 \times (100 - 0) = (.3)(100) = 30 \)
      - Total \( V_{ax} = \text{Current } V_{ax} + \Delta V_{noise} + \Delta V_{light} = 50 \)
  - Blocking
    - Clearly, the first 16 trials in Phase 1 will result in most of the \( \lambda \) accruing to the first CS, leaving very little \( \lambda \) available to the second CS in Phase 2

Rescorla-Wagner Model

- Theory not perfect:
  - Can’t handle second-order conditioning
  - Can’t handle latent inhibition
- But, it has been called the “best” theory of Classical Conditioning

Rescorla-Wagner & Delta Rule in Neural Network Learning

- Rescorla-Wagner Rule:
  - \( \Delta V = \omega \cdot [D - V_{ax}] \)
- The second is Sutton and Barto’s (1981) reformulation of Widrow-Hoff:
  - \( R(t) = [z(t)] - y(t) \cdot x(t) \)
- It simply says that the amount of reinforcement at time \( t \) for weight \( i \) is a function of the difference between desired and actual output, as well as the signal through the weight at that time