

An Information Processing Perspective on Conditioning

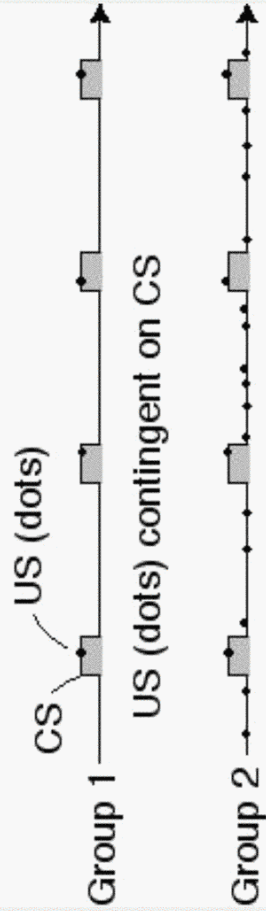
C. R. Gallistel

Rutgers Center for Cognitive Science

Five Classic Experiments

- Rescorla's background conditioning experiment (the truly random control)
- Kamin's blocking experiment
- Reynold's overshadowing experiment
- Wagner's relative cue validity experiment
- Conditioned inhibition experiments

Truly Random Control

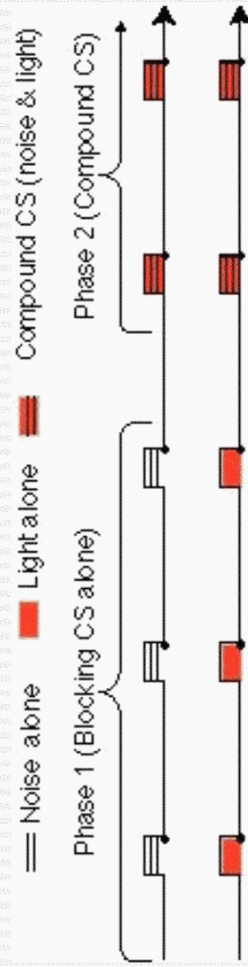


US (dots) not contingent on CS
(but CS-US pairing unchanged)

Result: Only Group 1 conditioned

Implication: Contingency not temporal pairing drives conditioning

Blocking

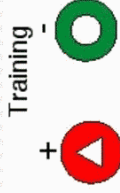


Result: No conditioning to the CS that first appears in compound (the blocked CS)

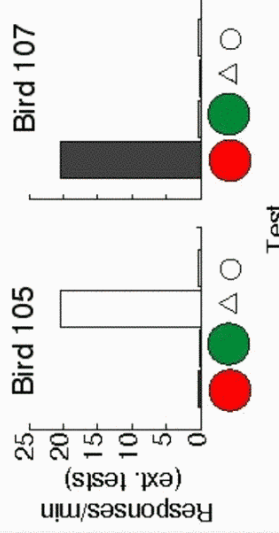
Implications

- Temporal pairing not sufficient
- Conditioning occurs only if US “surprises” subject

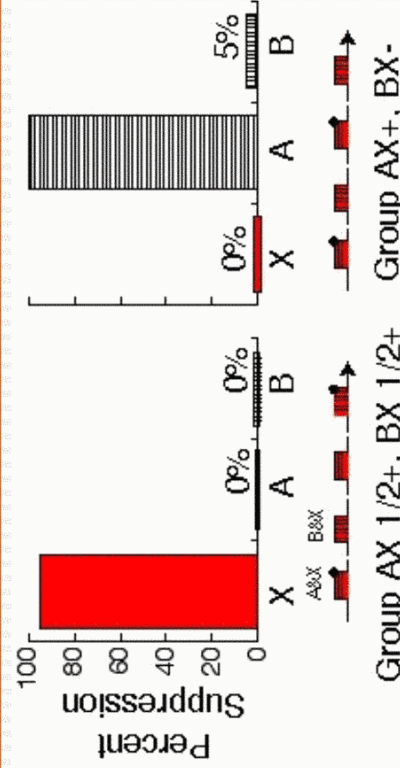
Overshadowing



Implication: When predictors are redundant, one predictor is eliminated



Relative Validity



• Implication: only the cue that carries the most information about the US gets conditioned

Conditioned Inhibition

Explicitly unpaired protocol



Feature negative protocol



Implications:

- Temporal pairing neither necessary nor sufficient for conditioning
- Negative contingency just as effective as positive contingency

So what's the problem?

"We provide the animal with individual events, not correlations or information, and an adequate theory must detail how these events individually affect the animal. That is to say that we need a theory based on individual events."

--Rescorla (1972, p. 10)

But we do present correlations!

- We correlate CS and US occurrences so that the CS provides information about the timing of the US
- Rescorla presumably meant that the conditioning process does not operate at the level of information
- Because information does not increment associations, events do

Shannon meets Pavlov

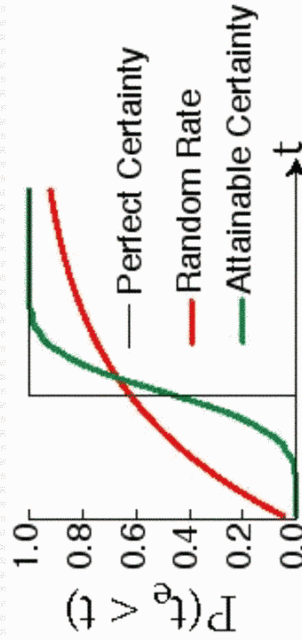
- From an information processing perspective, conditioning is driven by information
- The information an event provides to a subject is measured by the reduction in the subject's uncertainty--Shannon, 1948

Principles

- Subjects respond only to informative CSs
- CSs inform to the extent they change uncertainty about the time to the next US
- Bandwidth maximization by minimizing number of information carrying CSs
- Information in a protocol is carried by the temporal intervals between events and the numbers of events
- Weber's law: uncertainty (noise) in the representation of intervals and numbers is proportional to their magnitude

Tentative Assumption

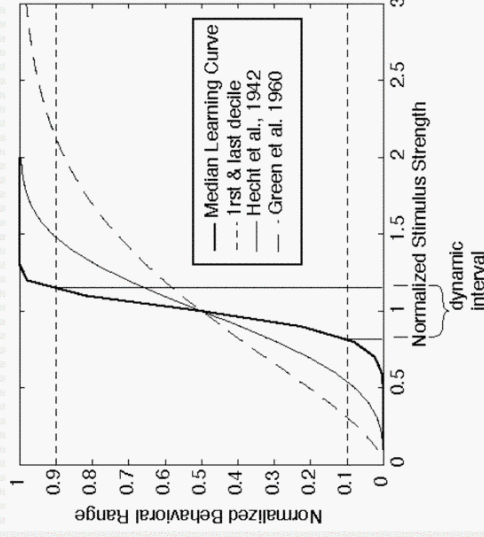
- Uncertainty about the time to the next US is represented by a cumulative probability function



Can Shannon Information Predict Acquisition?

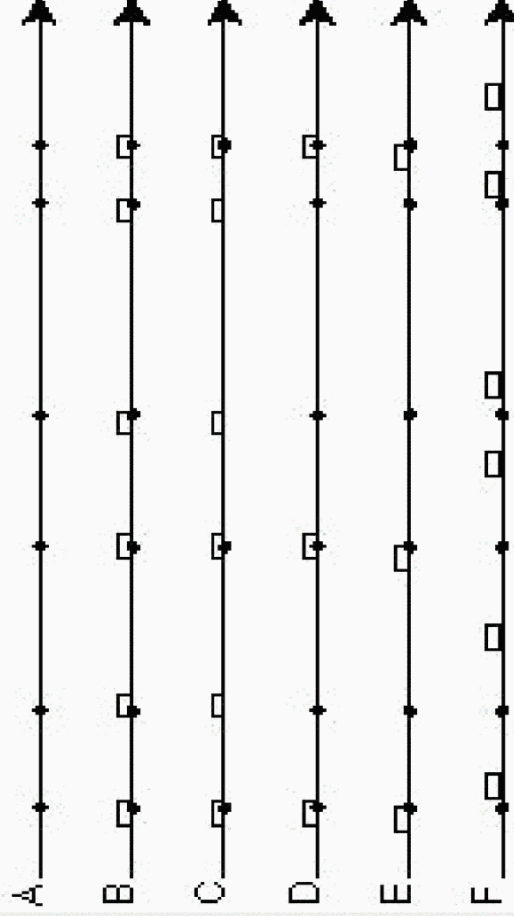
Preliminaries

- The appearance of a conditioned response in the course of conditioning is abrupt
- As if it were the outcome of an evidence-based decision process



Conditioning Protocols

□ CS • US



Information in Random Rate Change Protocols

$$\dot{H} = \lambda \log_2 \left(\frac{e}{\lambda \Delta \tau} \right)$$

$$\bar{H} = \frac{1}{\lambda} \lambda \log_2 \left(\frac{e}{\lambda \Delta \tau} \right) = k - \log_2 \lambda$$

$$\bar{H}_B - \bar{H}_{CS} = (k - \log_2 \lambda_B) - (k - \log_2 \lambda_{CS}) = \log_2 \lambda_{CS} - \log_2 \lambda_B$$

$$\log_2 \left(\frac{\lambda_{CS}}{\lambda_B} \right)$$

Bits Communicated per Reinforced CS

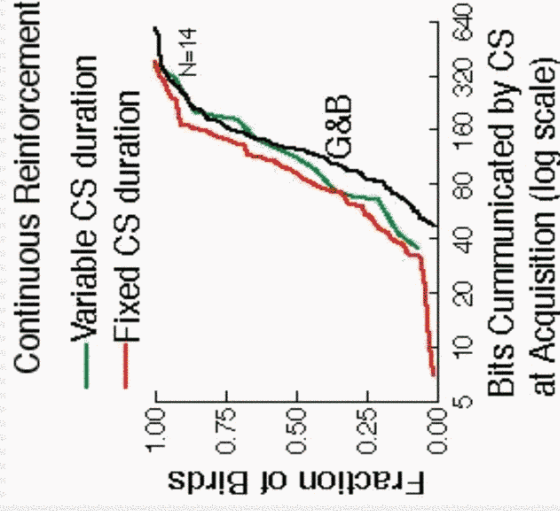
Accumulated Bits at Acquisition

N_a is the number of reinforced CS presentations prior to the onset of conditioned behavior

$$N_a \log_2 \left(\frac{\lambda_{CS}}{\lambda_B} \right)$$

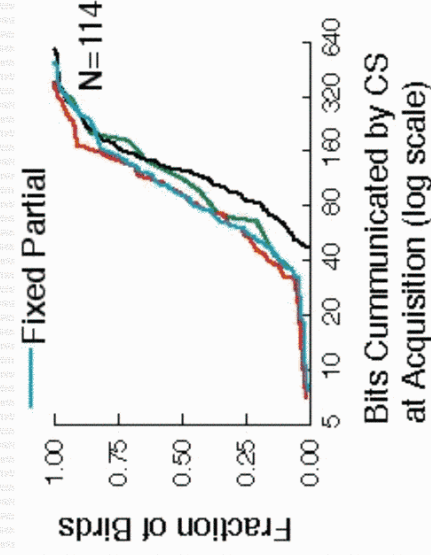
Continuous Reinforcement

- When there are no USs during the intertrial intervals, it does not seem to matter whether CS predicts US at fixed or variable latency



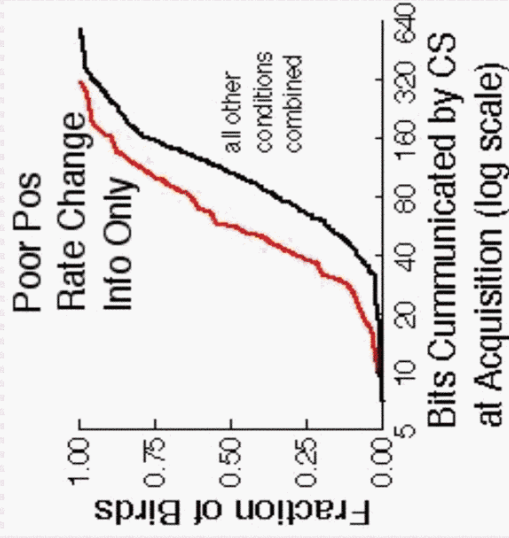
Partial Reinforcement

- Partial reinforcement has no effect on λ_{CS} / λ_B hence, it should have no effect on reinf to acq, hence no effect on bits cumulative bits conveyed at acquisition



Poor Positive Protocols

- As in previous analyses, we ignore the contribution from the fixed CS-US latency
- In terms of bits accumulated, birds take off too soon



Contribution from Fixed CS-US Interval

$$H = \frac{1}{2} \log_2 \left[2\pi e \frac{\sigma^2}{(\Delta\tau)^2} \right]$$

Entropy of a Gaussian

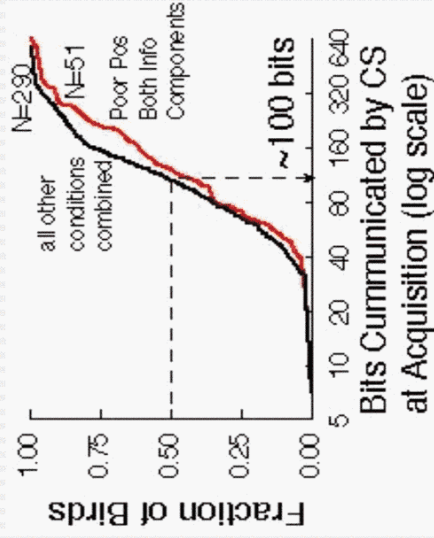
$$\text{Additional info available} = \frac{1}{2} \log_2 \left(\frac{e}{2\pi} \right) - \log_2 w$$

where w = Weber fraction

$$\text{Cumulative Bits at Acq} = N_a \log_2 \left(\frac{\lambda_{CS}}{\lambda_B} \right) + 2N_a$$

Poor Positive Recidivus

- If we assume that the additional information from the fixed CS-US latency matters only when there are USs during the intertrial intervals, then the poor positive distribution fits with all the others



Conclusion

- For all the protocol variations for which we have data, the cumulative bits conveyed by the CS at the point of acquisition in the median bird is about 100 and the distributions about this median are the same

Collaborators & Support

- The late John Gibbon
- Peter Balsam
- Stephen Fairhurst
- RO1 MH68073 Time and Associative Learning