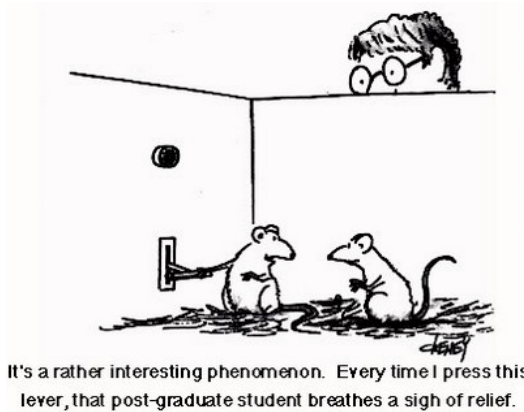# Instrumental Conditioning I: Control learning

It's a rather interesting phenomenon. Every time I press this lever, that post-graduate student breathes a sigh of relief.
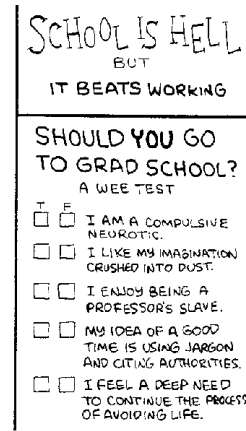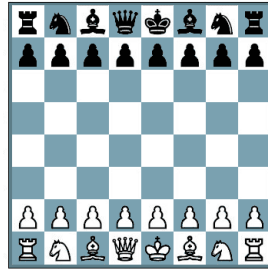
PSY/NEU338: Animal learning and decision making: Psychological, computational and neural perspectives

---

a "simple" decision making task
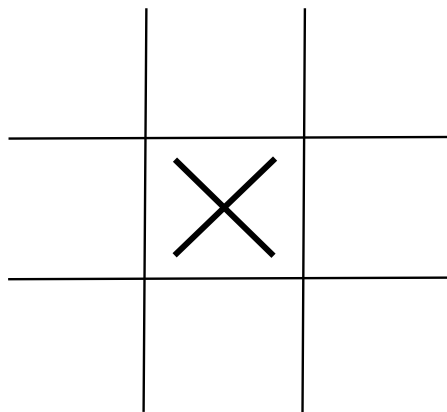
prediction helps decisions

predictions are for control

If we can predict what situations are associated with rewards we can try to bring those about through our actions

outline

- Thorndike: S-R learning
- Basic properties of instrumental conditioning
- Skinner: behaviorism, schedules of reinforcement

Edward Thorndike (1874-1949)

- Background: Darwin, attempts to show that animals are intelligent
- Thorndike was the first to show this systematically (not just anecdotes)
- Age 23: submitted PhD thesis on "Animal intelligence: an experimental study of the associative processes in animals"
- Tested hungry cats (also chicks, dogs) in "puzzle boxes"
- operational definition for learning: time to escape
- gradual learning curves, did not look like 'insight' but rather trial and error

---
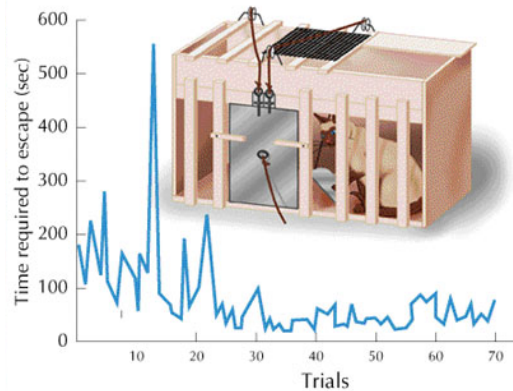
THORNDIKE'S PUZZLE BOXES AND THE ORIGINS OF THE EXPERIMENTAL ANALYSIS OF BEHAVIOR - Paul Chance, JEAB (1999)

There were 15 of these boxes, and they were constructed mainly of wooden slats and hardware cloth. Each box contained a door that the cat could open by manipulating some device. Cats opened the door to Box I by pressing a lever (The cat that first escaped from Box I may well deserve a place in history for being the first in a long line of lever-pressing animals.) Box K, the only box depicted graphically in the dissertation, required the performance of three distinct responses: The cat had to depress a treadle, pull on a string, and push a bar up or down before the door would finally fall open.

At first the cat's behavior appeared to be almost random, one might even say chaotic. Gradually, however, it became more orderly, more deliberate, more efficient. "The cat that is clawing all over the box in her impulsive struggle will probably claw the string or loop or button so as to open the door. And gradually . . . After many trials, the cat will, when put in the box, immediately claw the button or loop in a definite way".
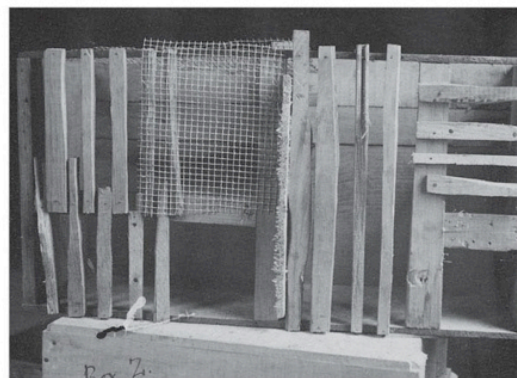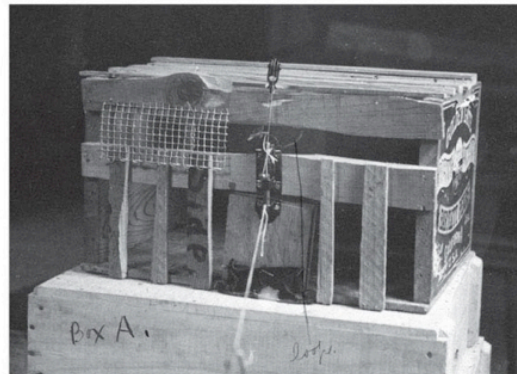
Thorndike: The Law of Effect

Of several responses made to the same situation, those which are accompanied or closely followed by satisfaction to the animal will, other things being equal, be more firmly connected with the situation, so that, when it recurs, they will be more likely to recur; those which are accompanied or closely followed by discomfort to the animal will, other things being equal, have their connections with that situation weakened, so that, when it recurs, they will be less likely to occur. The greater the satisfaction or discomfort, the greater the strengthening or weakening of the bond.
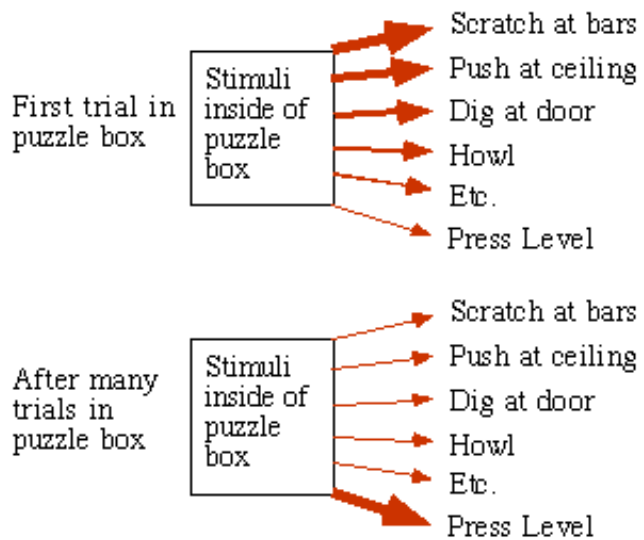
what is the role of the reinforcer?

- the reinforcer "stamps in" the association between the situation and some actions
- not needed after training: behavior becomes habitual
- automatic process once there is a goal (motivation)
- no need to assume more intelligence (imitation etc.) or causal learning/insight learning
- also: generalization, discrimination ("I must feed those cats")
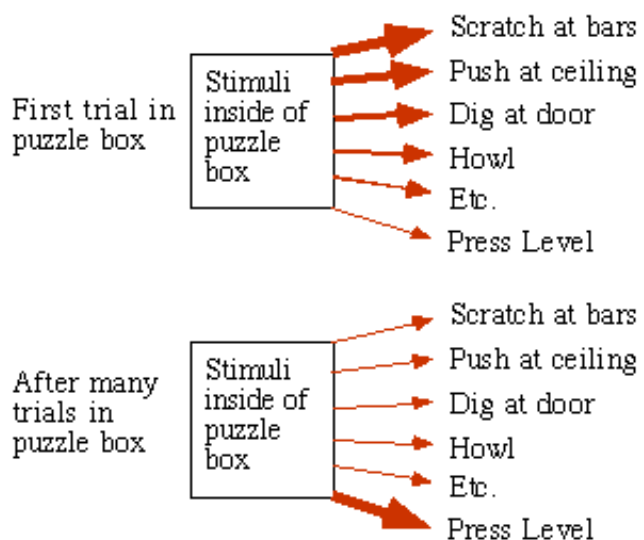
instrumental/operant conditioning

- Origin of the name: operation on the environment, actions are instrumental in achieving outcomes
- Commonly studied responses: lever pressing, key pecking, chain pulling, maze navigation

	↑	↓
appetitive	**positive *reinforcement* (reward)**	**omission**
aversive	**negative *reinforcement* (escape/avoidance)**	**punishment**

Any US can be used to *increase* or *decrease* a response! Different from Pavlovian conditioning: nature of the US does not automatically affect behavior

- Loud buzz in some cars when ignition key is turned on; driver must put on safety belt in order to eliminate irritating buzz (Gredler, 1992) the buzz is a negative reinforcer for putting on the seat-belt.

- Feigning a stomach ache in order to avoiding school (Gredler, 1992) school as negative reinforcer for feigning stomach aches.

- Rushing home in the winter to get out of the cold (Weiten, 1992). Fanning oneself to escape from the heat (Zimbardo, 1992). Cold weather as negative reinforcer for walking home (the colder the faster you walk..), and heat as negative reinforcer for fanning.

- Cleaning the house to get rid of disgusting mess (Weiten, 1992), or cleaning the house to get rid of your mother's nagging (Bootzin, et al , 1991; Leahy & Harris, 1989). Nagging/Mess as negative reinforcer to cleaning.

- Studying for an exam to avoid getting a poor grade (Bootzin & Acocella, 1980). Low grade as a negative reinforcer for studying (but.. a high grade is a positive reinforcer for studying at the same time)

- Taking aspirin to relieve headache (Bootzin & Acocella, 1980; Buskist & Gerbing, 1990; Gerow, 1992). Good example: headache as negative reinforcer to taking medication.

- Running from the building when the fire alarm sounds (Domjan & Burkhard, 1993). Fire alarm as negative reinforcer for leaving building.

- Smoking/Drinking in order to reduce a negative emotional state (Baron, 1992). Negative emotional state as negative reinforcer to smoking/drinking.

- Changes in sexual behavior (e.g., wearing condoms) to avoid AIDS (Gerow, 1992).

(some) determinants of responding

- drive (motivation) - affects both learning and performance
- reward magnitude (+ contrast effects)
- delay to reward: 2 alternative theories
  - distractors
  - lower value
- continuous reinforcement (CRF) versus partial reinforcement (PRF) - very common in life
  - random order?
  - strictly alternating?

summary so far...

- Instrumental conditioning as a form of adaptive control over the environment: animal behaves so as to bring about good things and avoid bad things
- Thorndike - a theoretician (S-R, habits, law of effect, law of readiness)
- did not study the whole spectrum of phenomena (eg., extinction)
- so far: all "trial based" experiments, but life is not always divided into trials...

outline

- Thorndike: S-R
- Basic properties of instrumental conditioning
- Skinner: behaviorism, schedules of reinforcement

Behaviorism (1913-?)

- John Watson (1913): Psychology as the Behaviorist views it (Behaviorist manifesto)
- "Psychology as the behaviorist views it is a purely objective experimental branch of natural science. Its theoretical goal is the prediction and control of behavior. Introspection forms no essential part of its methods, nor is the scientific value of its data dependent upon the readiness with which they lend themselves to interpretation in terms of consciousness. The behaviorist, in his efforts to get a unitary scheme of animal response, recognizes no dividing line between man and brute. The behavior of man, with all of its refinement and complexity, forms only a part of the behaviorist's total scheme of investigation." (1913, p. 158)

Behaviorism (1913-?)

- Many were excited. Others were not (Angell, Watson's doctoral mentor, wrote that "Watson should be spanked.")
- Why? What Watson's manifesto actually meant:
    - (i) psychology must be a science (an implicit assumption, unquestioned at the time by advocates and critics alike)
    - (ii) a fundamental principle of science is that its data must come from publicly observable phenomena
    - (iii) what had been taken to be the subject matter of psychology, namely consciousness, does not satisfy that principle because it cannot be observed publicly
    - (iv) the methods to which psychology must resort for studying consciousness, namely introspection, are not scientific methods
    - (v) therefore, the psychology of the time was not a science.

BF Skinner: Free operant training

- behaviorist (1904-1990)
  (student of Watson, the father of behaviorism)
- claimed that psychology should only study what can be measured; descriptive rather than theoretical science; black box
- believed that other than a few reflexes, all behavior is learned
- shaping of behavior through reinforcement: understanding behavior = controlling it
- against Thorndike's S-R: why hypothesize an S when it is not clear what it is? What is clear is the response and the outcome
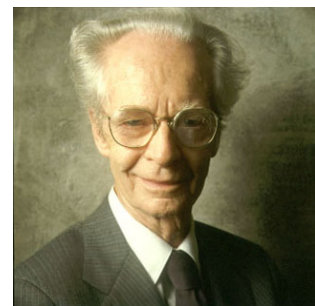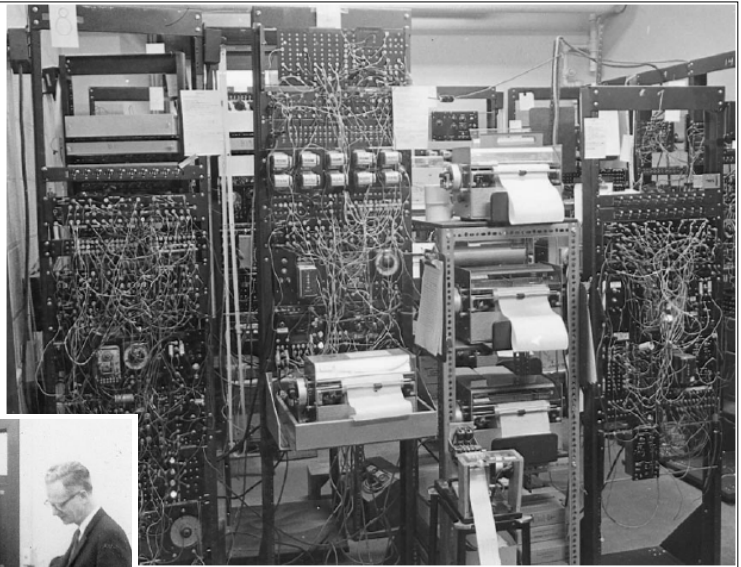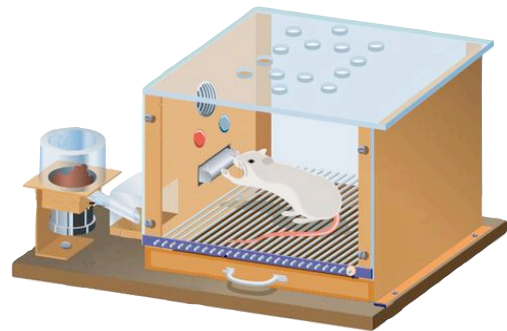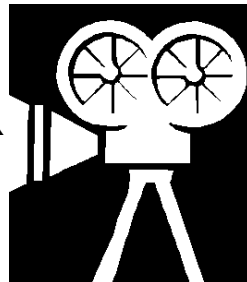
Fig. 6. Pigeon staff meeting. Seated from left to right are Mike Harrison, Bea Barrett, George Reynolds, and Charles Catania; walking behind George Reynolds is Roger Kelleher. Skinner is standing at the right, and the seated figure behind him is probably Dick Herrnstein.

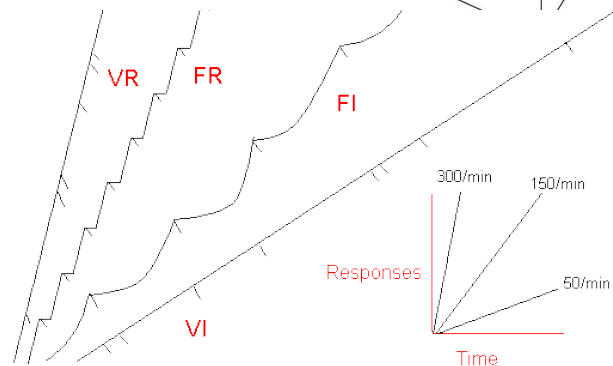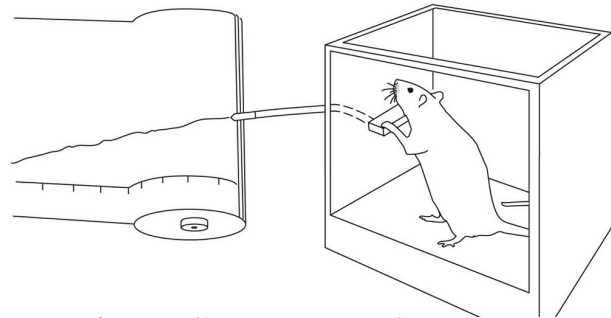# Free operant schedules

HR and LR movies

Schedules of reinforcement

Determine when a
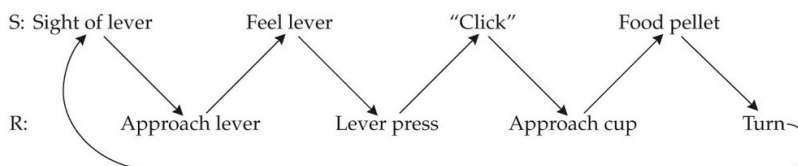reinforcer is given.
Two main types:

- Ratio schedules
  - Fixed, variable, random
- Interval schedules
  - Fixed, variable, random
- What does behavior
  look like in these?
- Examples?
- Also: DRH, DRL, S^D

VR   FR

FI

VI

300/min   150/min

Responses   50/min

Time

in practice: training

- pretraining
- shaping (superstitious behaviors)
- always start with CRF/FR1
- can use Pavlovian contingencies to help training (how?)
- which schedules are easier to train, interval or ratio? why?
- which schedules generate faster behavior? why?
- complicated to analyze behavior in these schedules from
  theoretical standpoint: many have given up (but we won't!)

S: Sight of lever   Feel lever   "Click"   Food pellet

R: Approach lever   Lever press   Approach cup   Turn

why Skinner (radical behaviorism) was wrong

- using 'intervening variables' helps make descriptions simpler (that is the whole idea behind computational modeling)
- looking into the black box helps understand behavior
- not everything that we care about is directly observable (electricity; but it is measurable)
- we care about what is in the black box itself!

summary so far...

- Instrumental conditioning as a form of adaptive control over the environment: animal behaves such as to bring about good things and avoid bad things
- Skinner - an empiricist (and a very good one!)
- concentrated on real-life-like free operant behavior (no trial structure)
- unfortunately also did not study the whole spectrum of phenomena (eg., extinction)
- whether you like it or not some aspects of behaviorism are here to stay