

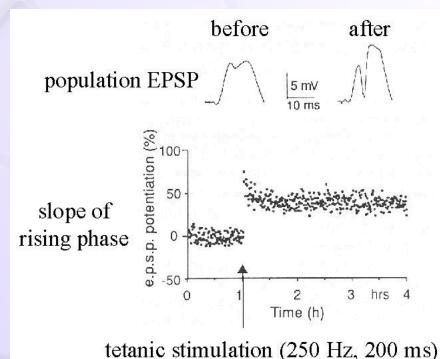
Learning, Motivation & Emotion

Randall C. O'Reilly

Learning <- Motivation <- Emotion

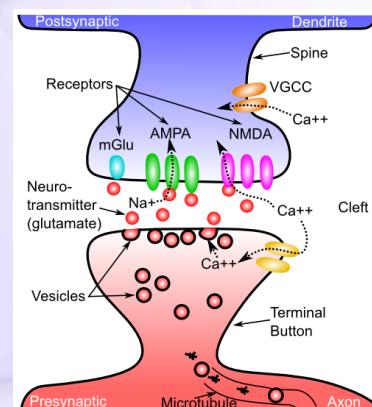
- Learning is the most important feature of the human brain: we learn almost *everything*!
 - Happens in **synapses** everywhere
 - Powers current deep AI models
- Motivation drives what we learn
 - Intrinsic motivation most powerful: goals!
- Emotion drives Motivation, Learning
 - Dopamine drives learning in basal ganglia
 - Classical / Instrumental conditioning

Synapses Change Strength (in response to patterns of activity)



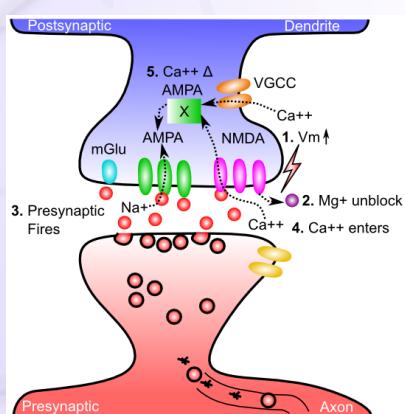
3

What Changes??



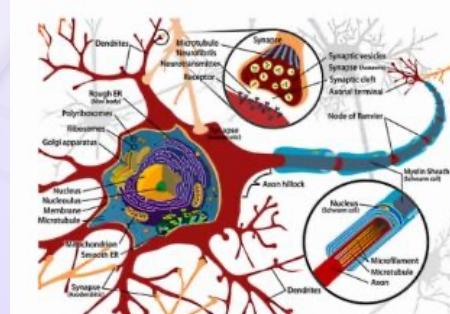
4

Gettin' AMPA'd



5

HEBBIAN LEARNING

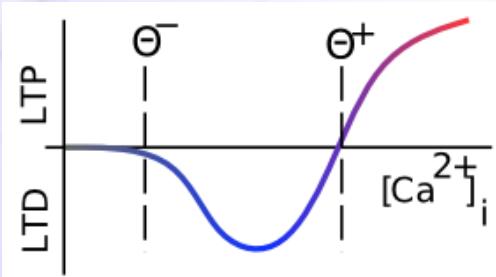


CELLS THAT FIRE TOGETHER,
WIRE TOGETHER

© BYODS & UNIVERSES

6

Which Way?

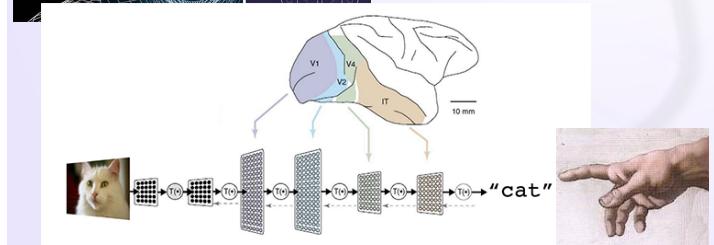


Low Ca = "long term depression" – synapse gets weaker
High Ca = "long term potentiation" – synapse gets stronger

7

Neocortical Learning

What works:
error backpropagation



Not biological: error goes opposite direction from spikes
Not psychological (where is that hand when you need it!?)



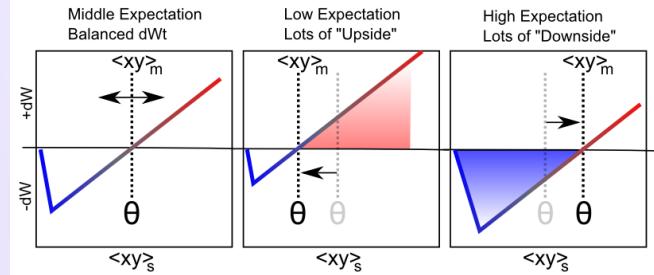
Prediction is very difficult, especially about the future.

Robert Storm Petersen (1882-1949)
Danish cartoonist, writer, animator,
illustrator, painter and humorist

Key idea: We learn by constantly generating hypotheses or predictions about *what will happen next!*

Error = outcome - prediction

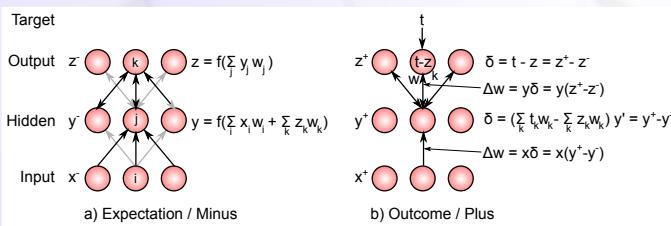
Dynamic Ca^{2+} Curve = Error



$dW = \text{Outcome} - \text{Prediction} = \langle xy \rangle_s - \langle xy \rangle_m$

10

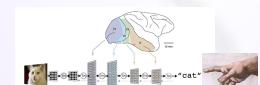
Bidirectional Connections Propagate Error



Math! It works!

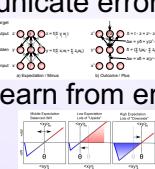
Summary of Neocortical Learning

Error Backpropagation works



Bidirectional connections communicate error

Dynamic Ca^{2+} curve: synapses learn from error



Predictive learning: constant source of error
(and Thalamus has circuits that could do it!)

Learning Rules Across the Brain

Area	Learning Signal		
	Reward	Error	Self Org
Primitive Basal Ganglia	+++	- - -	- - -
Cerebellum	- - -	+++	- - -
Advanced Hippocampus	+	+	+++
Neocortex	++	+++	++

+= has to some extent ... +++ = defining characteristic – definitely has
-= not likely to have ... --- = definitely does not have

13

Learning happens where it's used
(memory => processing)

Basal ganglia: learning what actions (not) to use

- based on reward / punishment (operant)

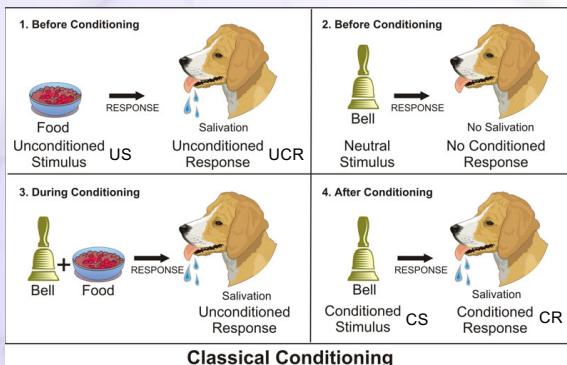
Cerebellum: learning to perfect actions

- based on error signals (e.g., feeling awkward)

Neocortex: learning how to see, hear, speak, reach, act, socialize... *everything!*

Hippocampus: learning snapshots of *everything*
(*explicit, declarative learning in Hippo, Cortex*)

Classical Conditioning



CS associated with US, thinking of US drives CR

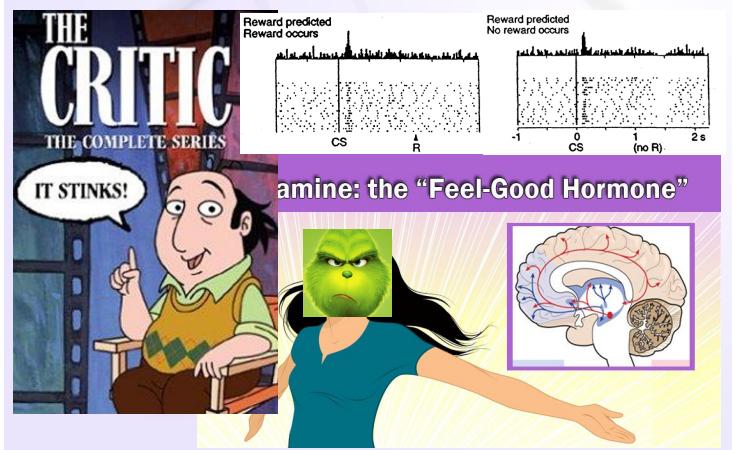
“Real World” Conditioning

The Office: <https://vimeo.com/35754924>

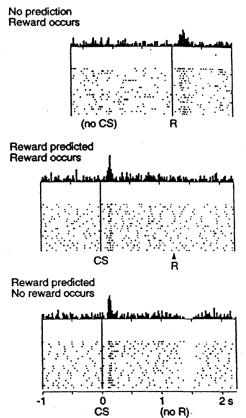
Dopamine Drives Conditioning



Dopamine is a Nasty Critic!



Reinforcement Learning: Dopamine



CS = Tone

R = Juice drop

Classical conditioning happens in **dopamine**

19

Dopamine Lessons

- Dopamine = Outcome – Expectation
- Should you just always have low expectations, so even low outcomes seem good??
- I try hard to avoid hearing *anything* about movies

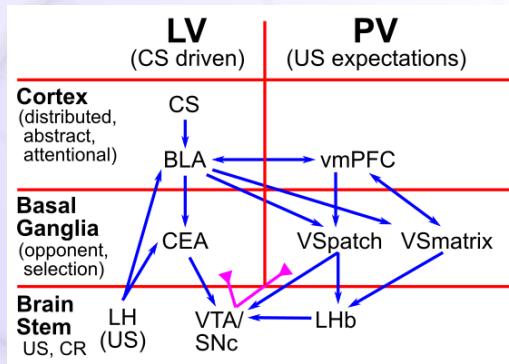
Partial Reinforcement!

Keeping your dopamine in the zone..

Dopamine learns to expect anything reliable and “cancels” it out



Biology of Classical Conditioning



Acquisition & Extinction

Acquisition: initial learning of CS -> US Assoc

– **Second order:** CS1 -> CS2 -> US

Extinction: learning that CS !-> US anymore

- This is NEW learning, not UN-learning!
- **Context-specific exceptions** to prior learning
- **Spontaneous recovery** of extinguished learning
- **Renewal:** Learn in A, extinguish in B, A -> renewed
- **Reinstatement:** expose to US, CS -> US reinst.

Limits of Classical Conditioning

Biological Preparedness: built-in pathways for CS's and US's

- Food can cause nausea, lights / tones shock, but not the other way around!

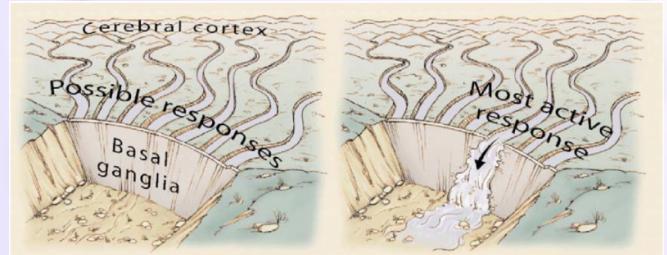
Operant / Instrumental Conditioning

Thorndike's Law of Effect:

- Actions -> Good stuff are "stamped in"
- Actions -> Bad stuff are "stamped out"

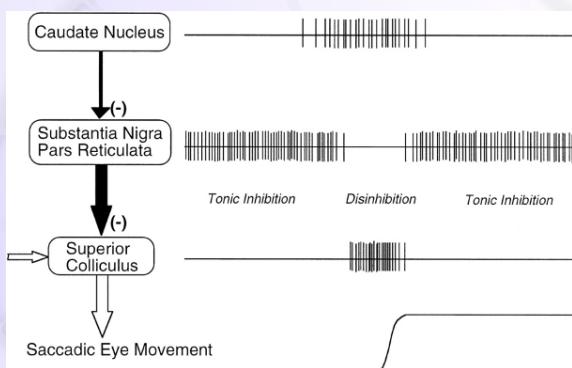
Dopamine = Good (bursts) vs. Bad (dips/pauses) drives learning in Basal Ganglia in accord with Law of Effect!

Basal Ganglia and Action Selection



26

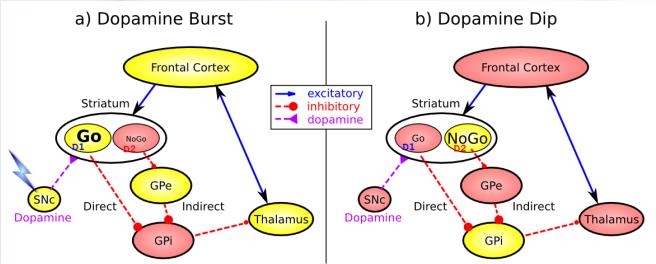
Release from Inhibition



27

Basal Ganglia Operant Learning

(Frank, 2005...; O'Reilly & Frank 2006)



Dopamine burst = do more of what you just did (Law of Effect)
Dopamine dip = do less of what you just did (bad outcome!)

-> Classical conditioning drives operant conditioning!!

28

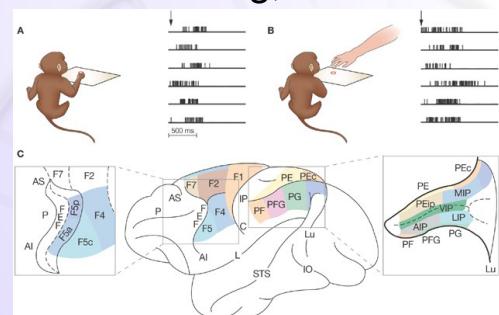
Operant Tricks

Secondary Reinforcer (e.g., \$\$): something associated with actual **Primary Reinforcer**

Shaping (by successive approximation) – it's how you get here:



Imitation Learning, Mirror Neurons



Innate, animal imitation very limited

Does this mean when we watch violent media, we act more violent??

Motivation

Behaviorism is all about **external** motivation

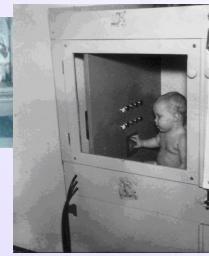
- Reward / punishment

People are all about **internal** motivation: goals!

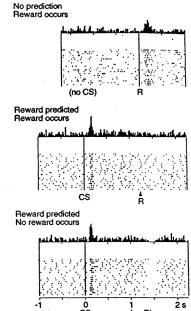
Behaviorist Motivations: Get rewards, avoid punishments..



Pavlov



Skinner



Dopamine

All **extrinsic**

Drives

Drive: desire to reduce unpleasant state from need (**drive reduction theory**: Hull, 1943)

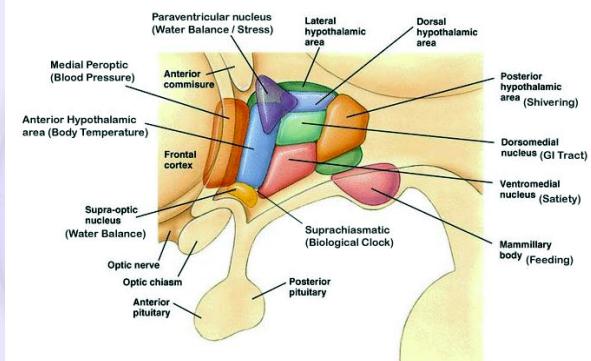
Homeostasis: maintain target levels of blood sugar, water, etc. Drive = restore target level.

Maintain **optimal arousal level?** Not everything is a drive..



Basic Drives: Hypothalamus

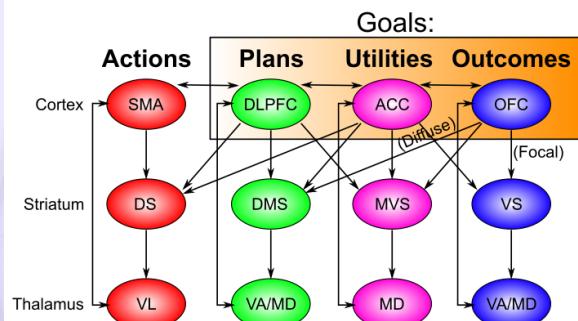
Nuclei of the Hypothalamus



Needs



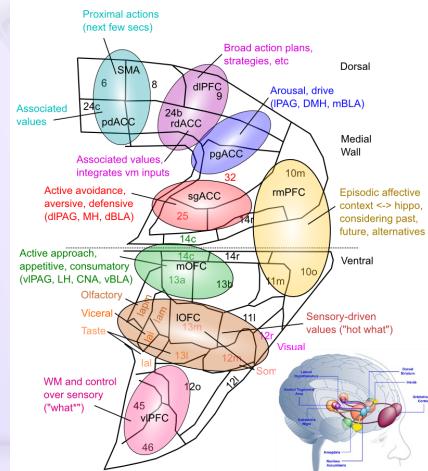
Distributed Goal Network



Striatum: helps select, maintain coordinated reps throughout network (BG gated WM)

Map of Goals in vmPFC

- Driven by subcortical connectivity
- High-dimensional, multi-factorial representation
- Consistent with fMRI, MDD in sgACC, etc



Extrinsic undermines Intrinsic?

Providing **extrinsic** rewards undermines **intrinsic** motivation!

- e.g., rewarding kids for homework?

Emphasizing trait makes people nervous

- "you're so smart" vs. "you worked hard!"

Motivating Phenomenology

- Why is it so hard to *start* something (packing for a trip, writing a paper, paying bills, cleaning desk...)
- But once started, it really isn't so bad..

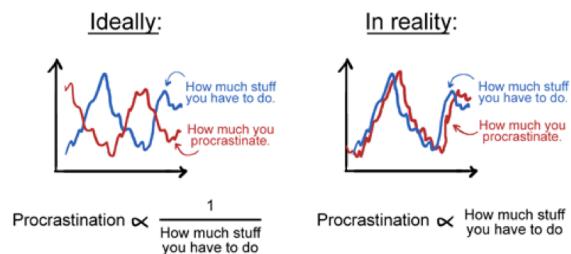
Motivating Phenomenology

- Ever find yourself playing mindless video games for far longer than you should?
- Why can't I stop myself from organizing my kid's Legos, or cleaning leaves from pool?

Piled Higher and Deeper by Jorge Cham

www.phdcomics.com

Procrastination



title: "Procrascorrelation" - originally published 10/27/2010

JORGE CHAM © 2010

Two Phases of Mental Life

Goal selection

- Careful weighing of costs / benefits to select goal
- Multiple constraint satisfaction of needs, "drives", opportunities, risks, costs, effort, etc..

Goal engaged

- Selected goal robustly held – hard to give up..
- Continuous evaluation of proximity to goal
- Dopamine bursts, dips as function of changes
- Costs are significantly downplayed (but learned)

Strong dissociations in value functions

Applied to Phenomenology

- Why is it so hard to *start* something (packing for a trip, writing a paper, paying bills, cleaning desk...)
- But once started, it really isn't so bad..

- *Goal selection process carefully weighs costs / benefits, considering many different possible goals*

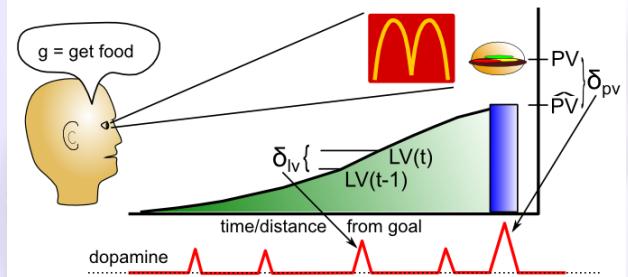
Applied to Phenomenology

- Ever find yourself playing mindless video games for far longer than you should?
- Why can't I stop myself from organizing my kid's Legos, or cleaning leaves from pool?

- *Goal is engaged: incremental progress drives dopamine – video games engineered to deliver*
- *Costs, alternatives are downplayed*



Dopamine = progress toward goal



LV = phasic dopamine driven by engaged goal
PV = was goal achieved or not; time to select new

Approach vs. Avoid

Approach = left hemisphere = dominant

Avoid = right hemisphere = subordinate

We approach positive outcomes, and avoid negative ones.

Anger is negative: is it approach or avoid?

Need to turn negative into a positive in order to act!

Work Motivation

If motivation is purely homeostatic, why are so many people motivated to work?

- Extreme: **workaholic** (esp USA, Japan)

Work Motivation

Industrial and Organizational Psychology (I/O)

Traits: stable personality factors – conscientious, honest, lazy, aggressive, sociable, shy, etc..

- Workaholic: neuroticism, perfectionism, conscientiousness

Perceived self efficacy: If we are good at something, we reinforce that, it becomes part of our self-image.. CCC = Control! (**competence**)

Goal-setting theory: specific and difficult goals are good for motivating employees..

Challenge Problem: Grad School

Grad students work long hours for little \$ and a low-probability shot at becoming a professor.. Why?

Social Motivation

Need to Affiliate: we like to have people around (at parties and speeches, but not at beaches..)

Women affiliate under stress, men do opposite

Need to Belong: positive mutual in-group interactions of an enduring nature: key factor in overall happiness.

Loneliness: not as much belonging as wanted

Intimacy: self-disclosure, partner responsiveness

The Motivational Perfect Storm

Imagine something that taps directly into multiple built-in motivational pathways simultaneously.. Something so powerful that it:

- Leverages strong social drives for belonging, affiliation, in-group promotion, out-group aggression
- Provides frequent, unpredictable bits of “progress” in flow of news, funny, entertaining information, etc..

Emotion

Emotion is short lived psychological state that:

- Activates specific emotional **neural systems**
 - (amygdala, ventral striatum, ventral-medial PFC)
- Has characteristic subjective **feeling**
- Triggers **physiological changes** (heart rate etc)
- Also **cognitive changes** (thinking is altered)
- And **behavioral responses** (behavior is affected)

Mood is long lasting, less intense, not triggered

Affect is simple “map” for categorizing emotions

- **valence** (good / bad)
- **arousal** (high / low)

Valence vs. Arousal



Basic Emotions

Fear, Anger, Sadness, **Happiness**, Disgust, Contempt, **Surprise/Interest**

Why are these basic? What causes them? What benefits do they provide? Why only 1 positive?

Where are Love and Hate?

Why not Pain? Hunger? Lust?

The Face of Emotion



These facial expressions are universal.
What are they!?

News Flash: “Not” Face Discovered!



C. F. Benitez-Quiroga, R. B. Wilbur, A. M. Martinez (2016), Cognition, 150, 77-84

Emotion for Communication

It is important to communicate emotion – hence facial expressions.. And emoticons :)

Display rules are cultural conventions about when and how to display emotions: intensify, de-intensify, mask, neutralize

Emotion -> Motivation

Emotions drive **motivation**, learning: without them everything is *neutral* and you can never decide!

Fear: avoid bad things; **Anger:** attack bad things

Disgust: don't eat rotten things

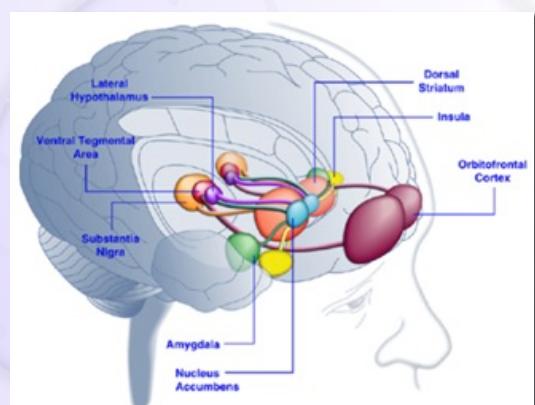
Contempt: punish bad people

Happiness: approach, enjoy good things

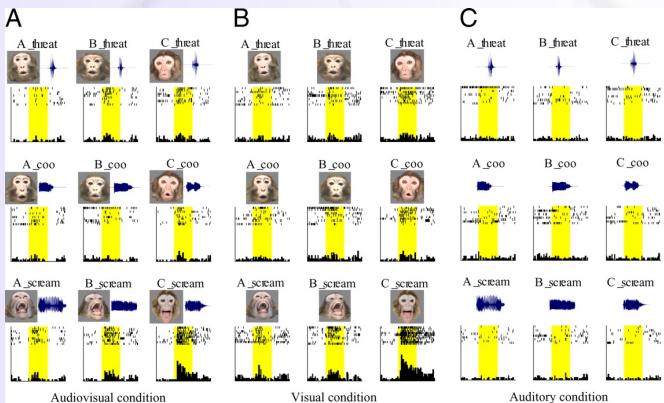
Surprise/interest: learn about unexpected things

What is the point of **Sadness**!?

The emotional brain



Amygdala Encodes Different Emotions



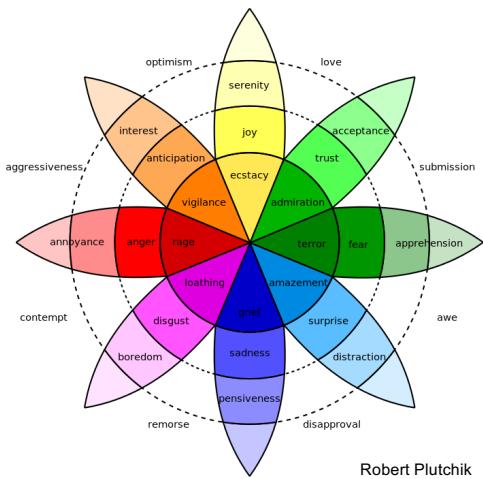
Negative Emotions

Happy families are all alike; every unhappy family is unhappy in its own way.

- Leo Tolstoy, Anna Karenina

We have many more negative emotions than positive ones!

So we should learn to **accept** this fact!



The Two Faces of Happiness

Happiness = instantaneous emotional feeling

Happiness = long-term life satisfaction

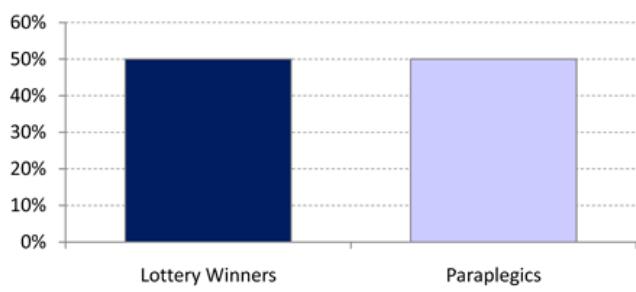
Good news: everybody's happy-ish!

Bad news: some people aren't!

Parents are low on instantaneous, high on long-term..

Surprise!?

Happiness One Year Later



Money != Happiness

- Above ~\$50k, more \$!= more happiness
- So what does = more happiness?
 - Immediate social network, relationships
 - Standing in community / social recognition
 - Goal achievement
 - **Dopamine driven by social feedback and progress toward goals!!!**
 - And a major (50%) contribution of genes..

Anger is not a liquid

- You don't have to vent it! Leads to rehearsal..

Figure out the underlying source of frustration

Deal with it

And / or learn to accept the situation

Stress Management..

Stress is mostly “in your head” – does all this stuff really matter that much, or perhaps not??

To reduce stress, become a Boulderite / Davisite:

- Exercise
- Meditation
- Optimism
- Sunshine, fresh air, etc..

Stress THEN Sick

- Stress itself does not make you sick..
- It is the depletion period AFTER the stressor is over that makes you sick!

