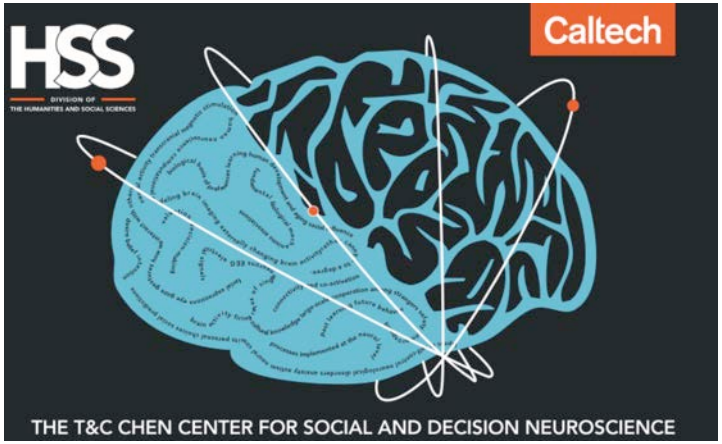
The *lingua franca* of neuroeconomics

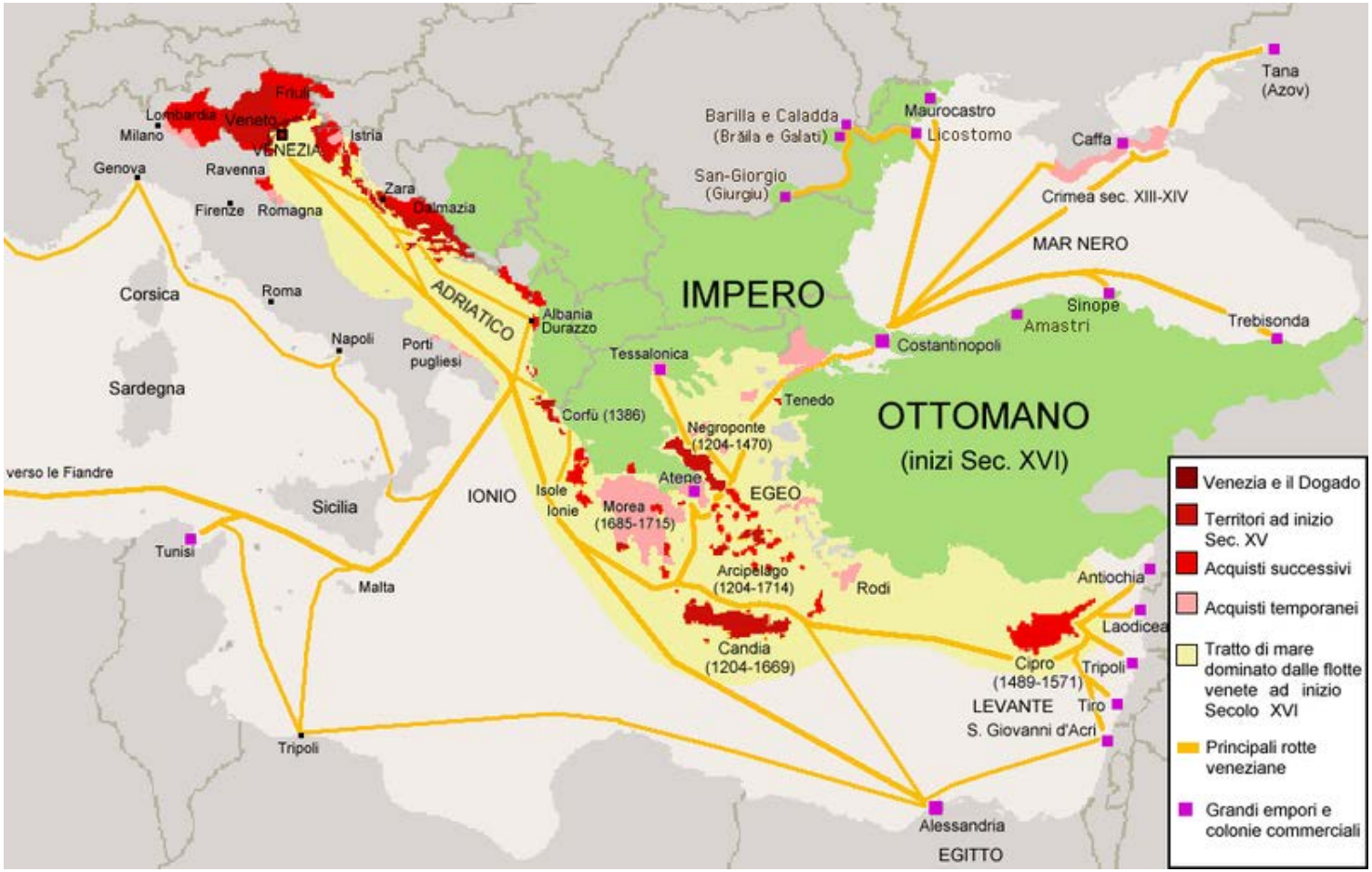
Colin F. Camerer

Caltech

Director, T&C Chen Center for Social and Decision Neuroscience



Republic of Venice (15-16th C)
needs a *lingua franca* (=trade language")



Levallois NatRevNeuro 2012

Outline

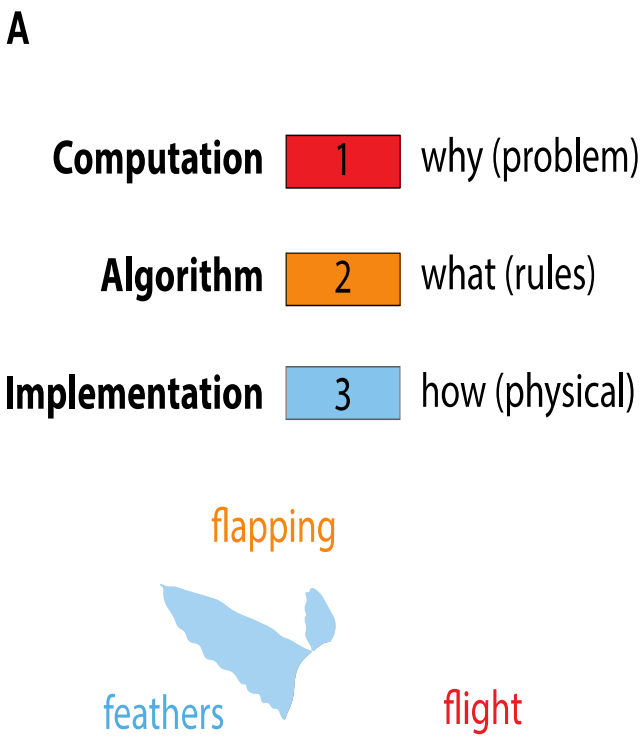
- What is (micro)economics trying to explain?
- Marr's 3-level model: Why? What? How?
- Examples:
  - Imagination and reward discounting
    - Future valuation ~~"time preference"~~
  - Habit
  - Biological non-fungibility (pain of paying)

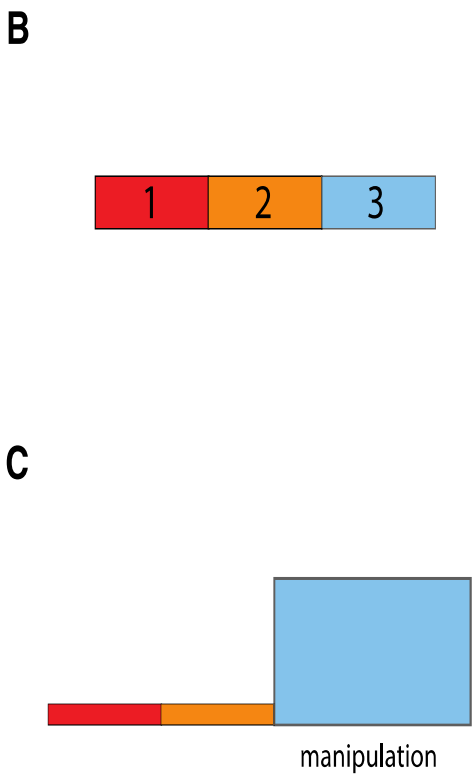
Neuroeconomics 2.0

- Neuroecon 1.0: [“incremental”]
  - Can brain processes help explain puzzles in economics?
  - “anomalies” Does brain evidence suggest generalizations of standard theory?
- Neuroecon 2.0 [“radical”]
  - *Do not start* with standard theory
  - What properties should neurally-plausible theories have?

What is microeconomics trying to explain? My answer

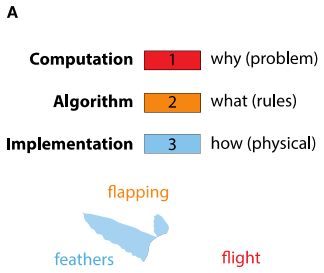
- Predict consequential human choices and [policy] what variables *change* choices
  - "nonchoice" methods are allowed; but must prove their value by predicting
  - Because economists are masters at creating new theories….
    - …Choice data do not provide enough constraint
    - Neural data, attention, etc. must be used to test predictions + explicit mechanism*





mentation. Step 2: implementation level work feeds back to inform the algorithmic level.
(C) An epistemological bias toward manipulation-based view of understanding induced by technology (black filled arrow).

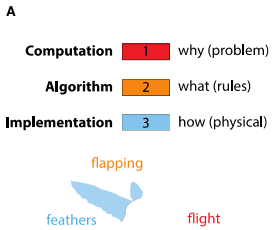
Krakauer+ Neuron 2017 (following Marr 82)



three levels

Neuroeconomic view

- Why? (evolved for fitness+ upgrades)
- What? (computational algorithms)
- How? (mechanism)
  - Computational plausibility
  - Detailed neuroeconomics



three levels

Traditional view

- Avoid market exploitation, or fitness*

↕

- Max u(x)

$$u_t + \sum_{\tau=t+1}^{T} \delta^{\tau-t} u_\tau$$

- How? (mechanism)

Neuroeconomic view

- Why? (evolved for fitness+ upgrades)

↕

- Algorithm (computational models)

↕

- How? (mechanism)
  - Computational plausibility
  - Detailed neuroeconomics

*evolutionary theory: Robson-Samuelson+, Firms: profit-maximization

HABIT

A neurally-inspired model of habit in consumer choice

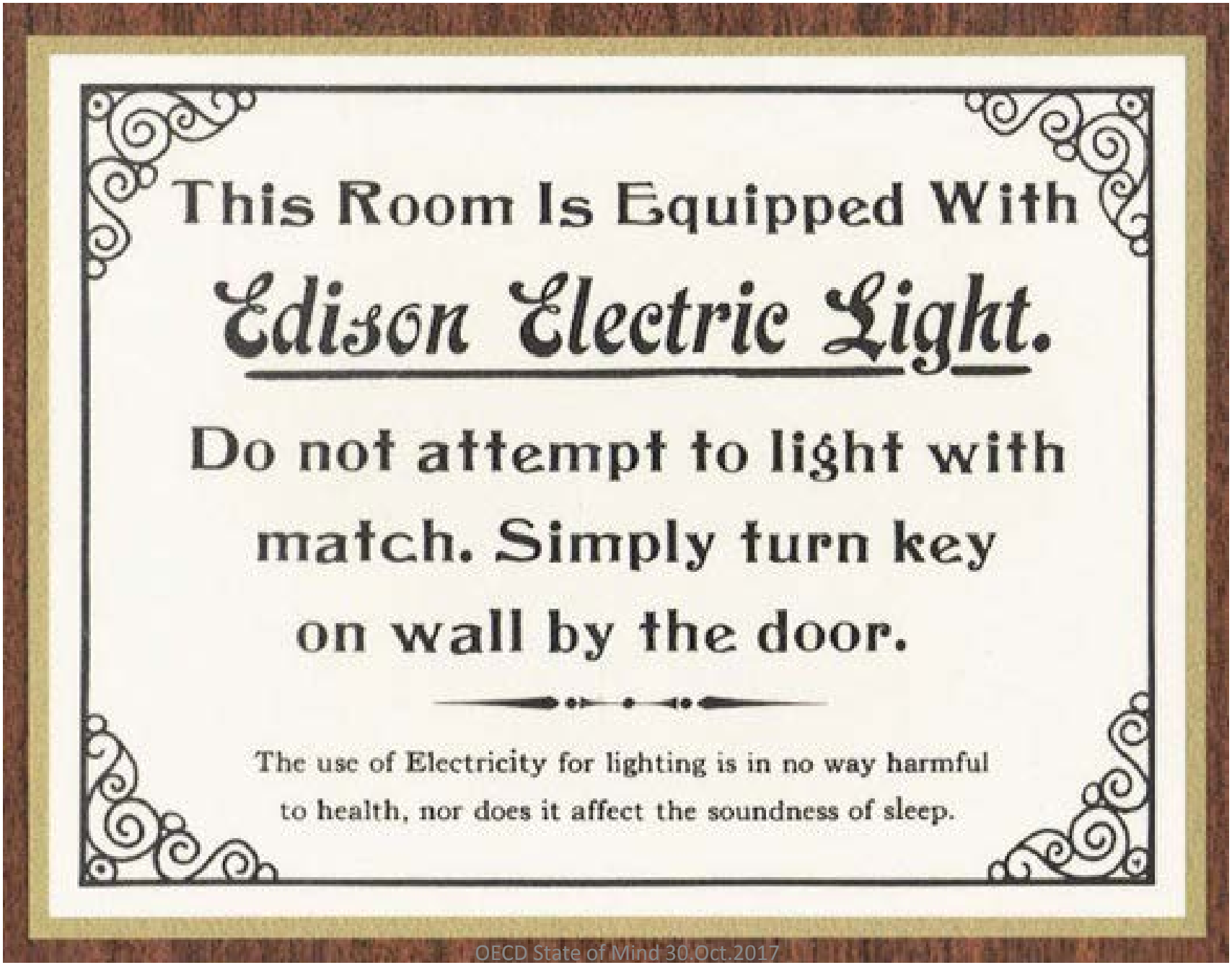
Colin F. Camerer

Caltech

Director, T&C Chen Center for Social and Decision Neuroscience

Peter Landry, Ryan Webb

U Toronto
Rotmann


This Room Is Equipped With
Edison Electric Light.
Do not attempt to light with match. Simply turn key on wall by the door.
The use of Electricity for lighting is in no way harmful to health, nor does it affect the soundness of sleep.

the shoe rack

Simple consumer theory: constrained utility-maximization

- Maximize utility given preferences u(.) and information θ, subject to prices p_i and income constraint y

$$\text{Max } u(x_1, x_2, \ldots x_n \mid \theta)$$

$$\text{subject to } \Sigma_i\, p_i x_i \leq y \qquad \text{(income)}$$

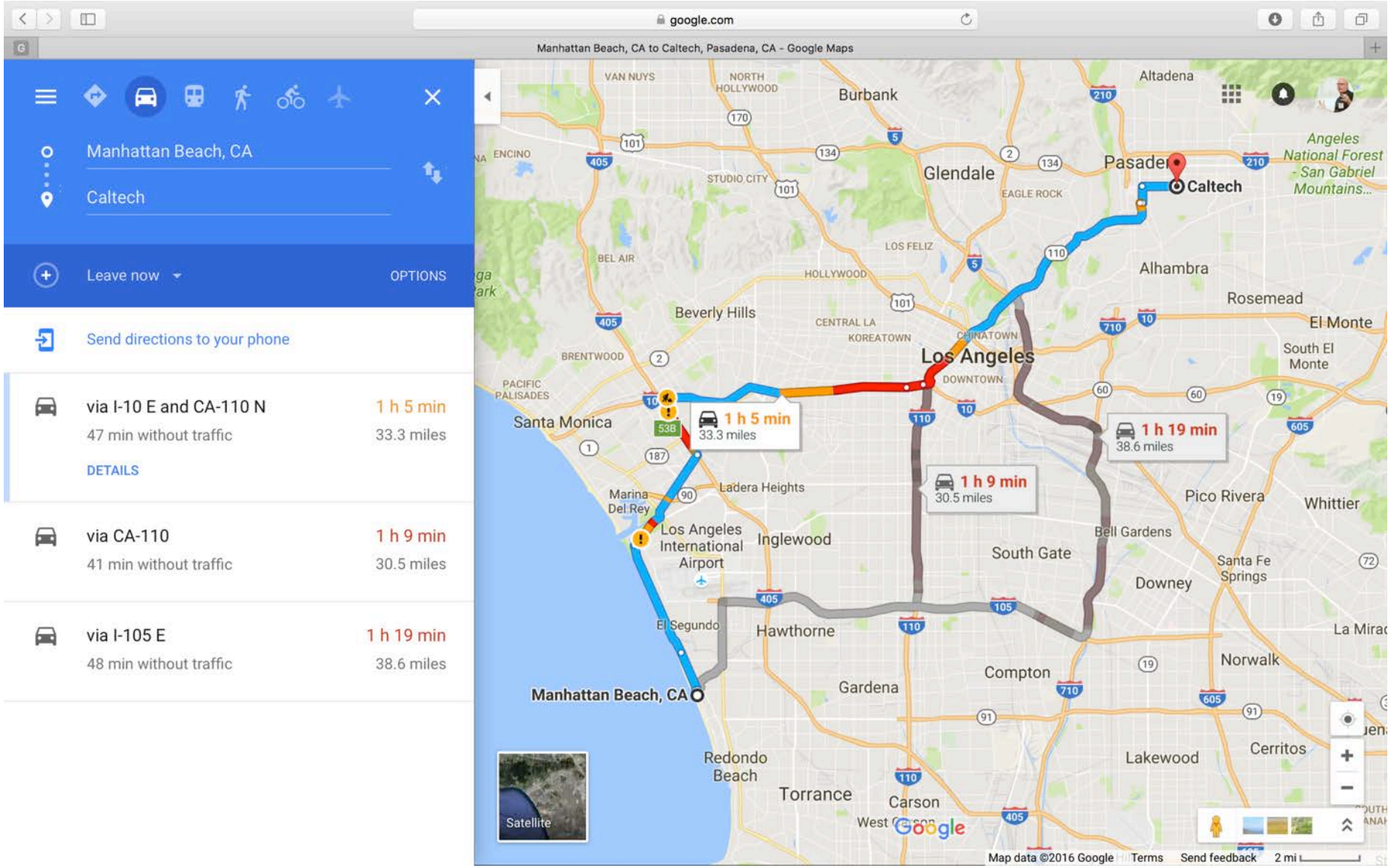
0. Preview: A neuroeconomic theory of consumer choice

- Model-free (MF) habit
  - Face a state (choice set, prices, θ, location)
  - Make the same choice as last time
  - fast, low-effort…but inflexible, backward-looking
- Model-directed (MD)
  - Evaluate actions by expected reward of consequences using a mental representation ("model")
  - Max or softmax response
  - Slow, high-effort…but flexible, forward-looking

How does the “model-directed” system work?

- Depends on internal mental representation
- Model maps actions to consequences + values
  - Decision tree
  - Attributes are listed and weighted

- What does model-directed look like?


google.com
Manhattan Beach, CA to Caltech, Pasadena, CA - Google Maps
Manhattan Beach, CA
Caltech
Leave now
OPTIONS
Send directions to your phone
via I-10 E and CA-110 N
1 h 5 min
47 min without traffic
33.3 miles
DETAILS
via CA-110
1 h 9 min
41 min without traffic
30.5 miles
via I-105 E
1 h 19 min
48 min without traffic
38.6 miles
1 h 5 min
33.3 miles
1 h 19 min
38.6 miles
1 h 9 min
30.5 miles
Manhattan Beach, CA
Caltech
Los Angeles
Santa Monica
Pasadena
Glendale
Burbank
Altadena
Alhambra
Rosemead
El Monte
South El Monte
Beverly Hills
Inglewood
Hawthorne
Gardena
Compton
South Gate
Downey
Norwalk
Pico Rivera
Whittier
Bell Gardens
Lakewood
Cerritos
Torrance
Carson
Redondo Beach
El Segundo
Los Angeles International Airport
Marina Del Rey
Ladera Heights
Satellite
Map data ©2016 Google
Terms
Send feedback
2 mi

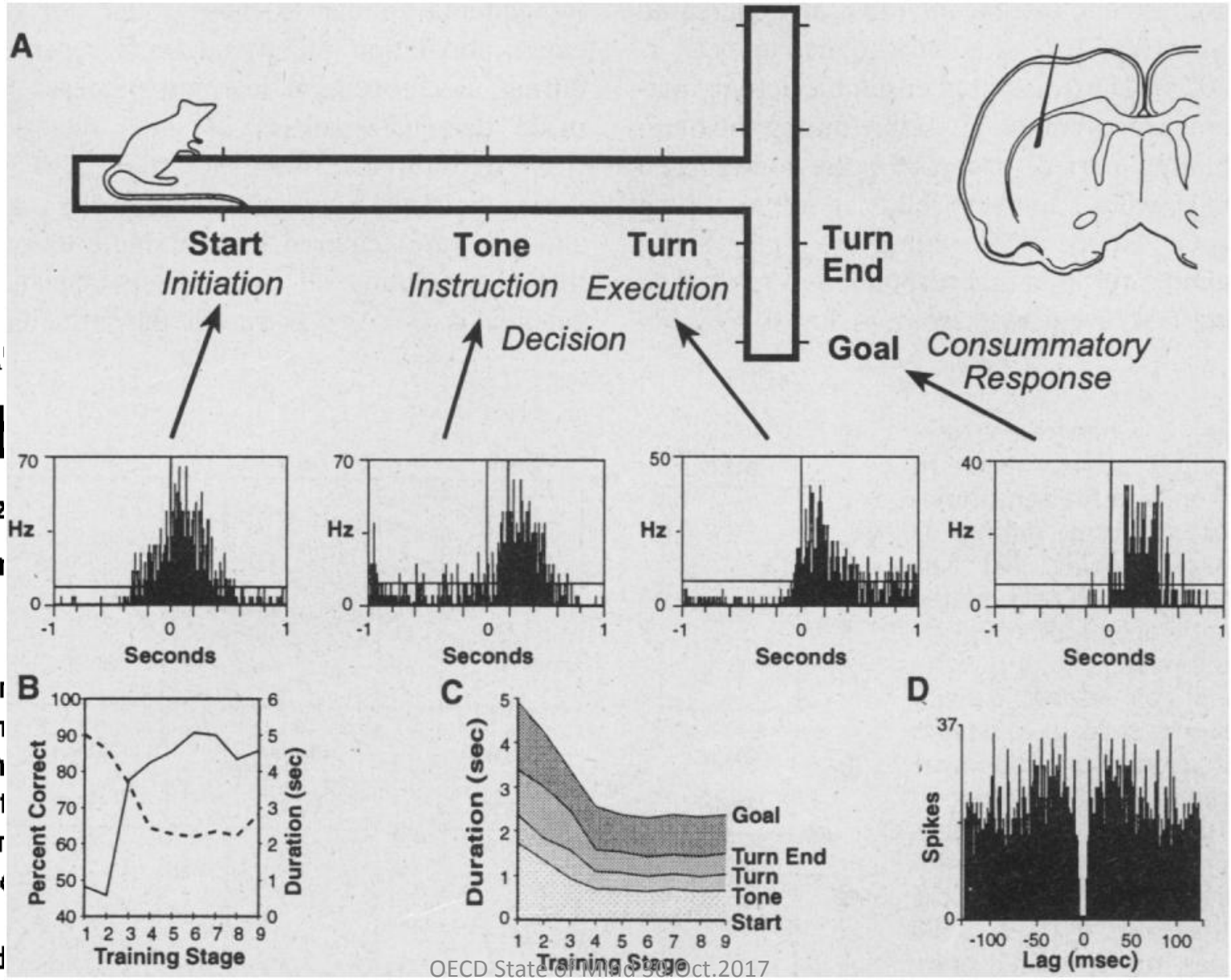
scope

- Canonical examples:
  - regular decisions, every day or week
  - food
  - commuting
  - morning/bedtime rituals
- Do NOT consider habits induced by
  - learning-by-doing and switching costs
  - biologically addictive substances

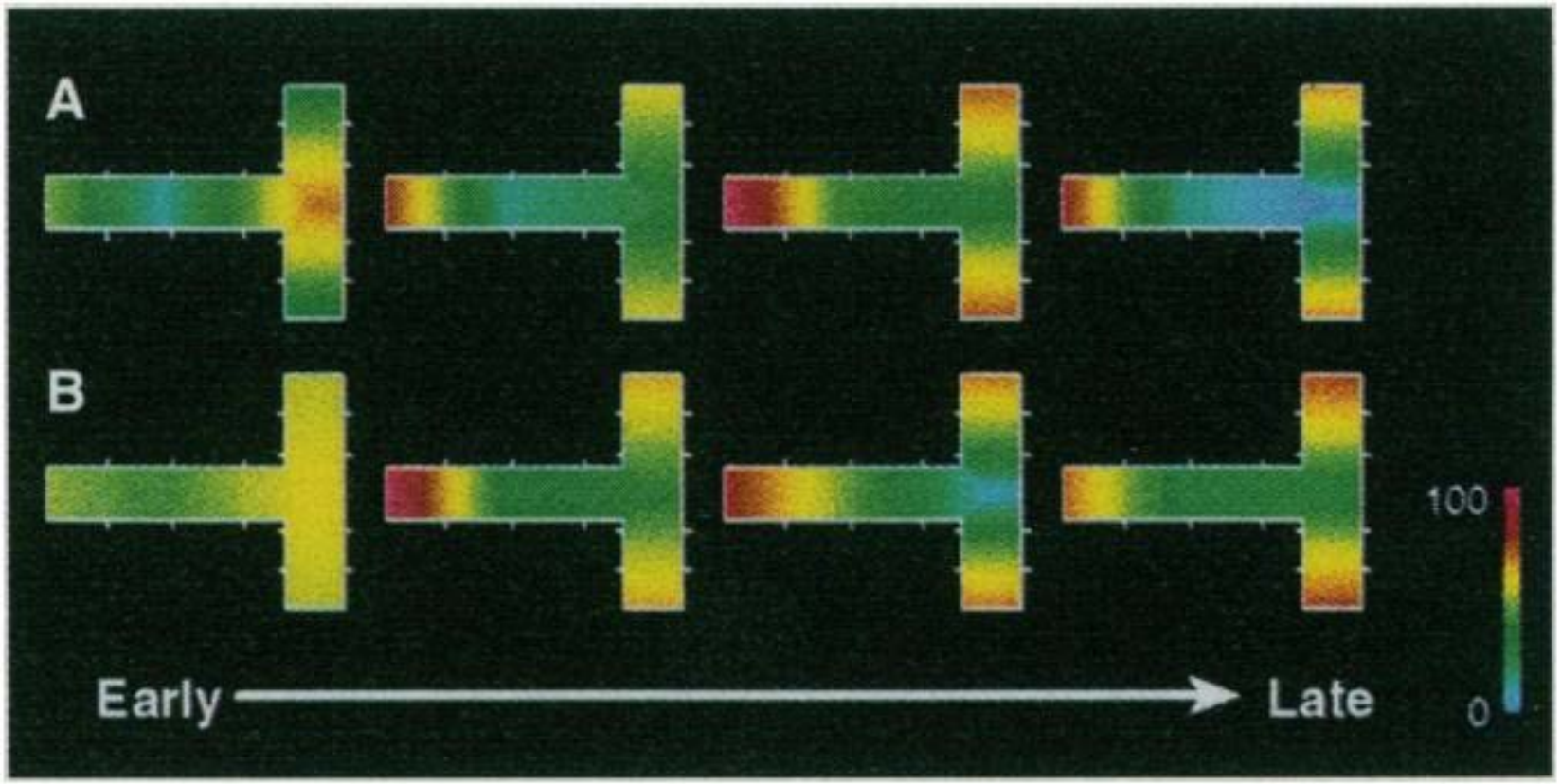
two central questions

- How *exactly* do habit and MB systems work?
  - psych, neuro (c 1980s)
  - computational specification (c 1990s)
  - modern neuroscience (c2000s)
- How do people *shift* between systems?
  - reliability signal
  - Habits yield low reward repeatedly → unreliability goes up → transition to MD

Habits as action sequences (Jog+ Sci '99)



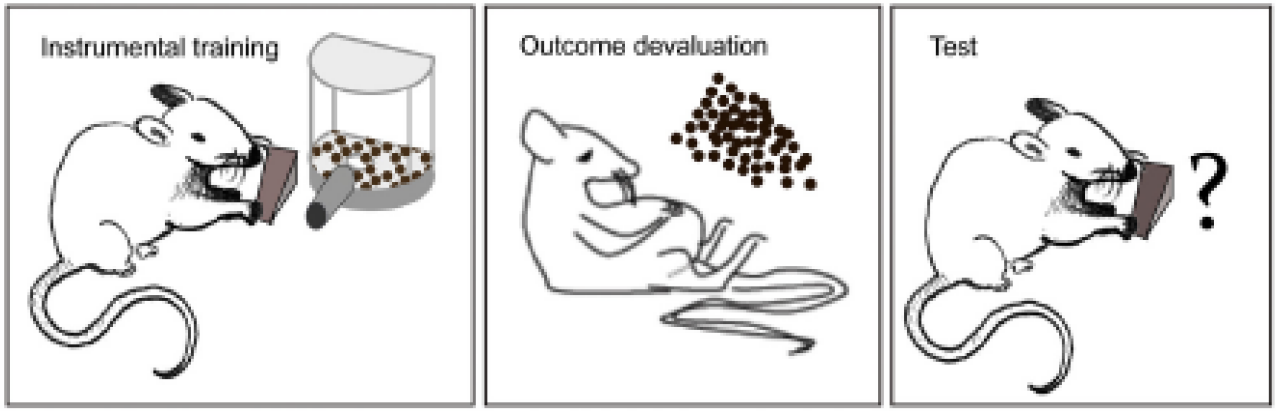
‘reorganization’ of neural firing over 9 training days

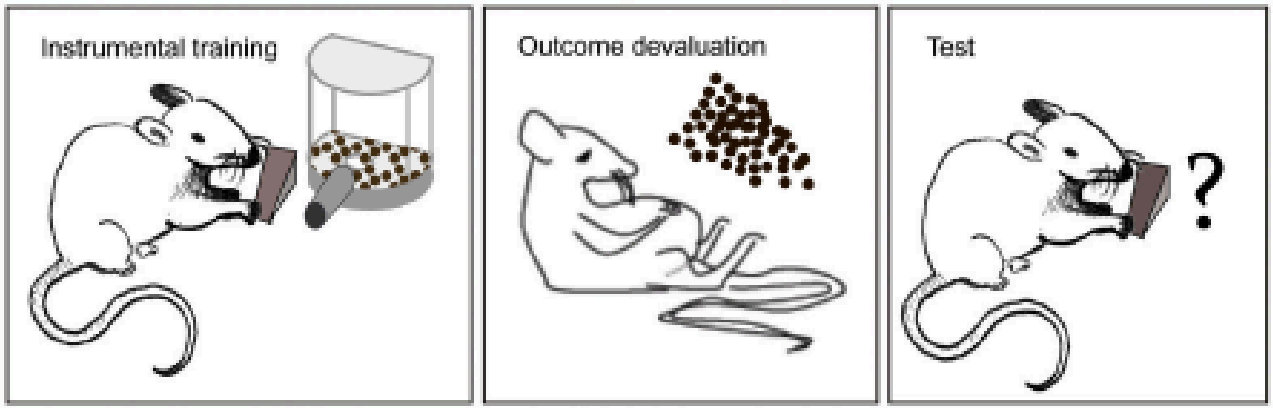


and addictive states (*32*). This view is com-

Adams, Dickinson et al 1980s

Marker of habit is persistent choice despite *devaluation* of rewards


Instrumental training
Outcome devaluation
Test
?



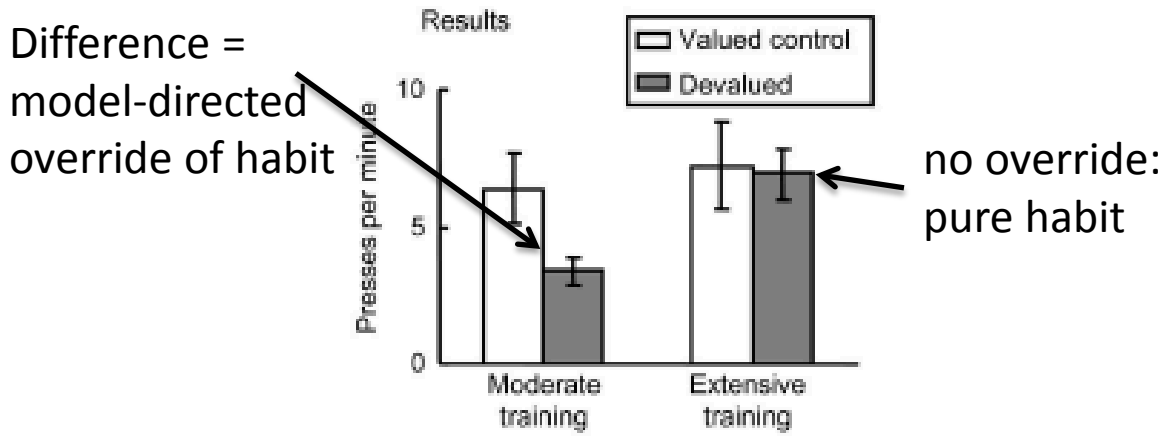
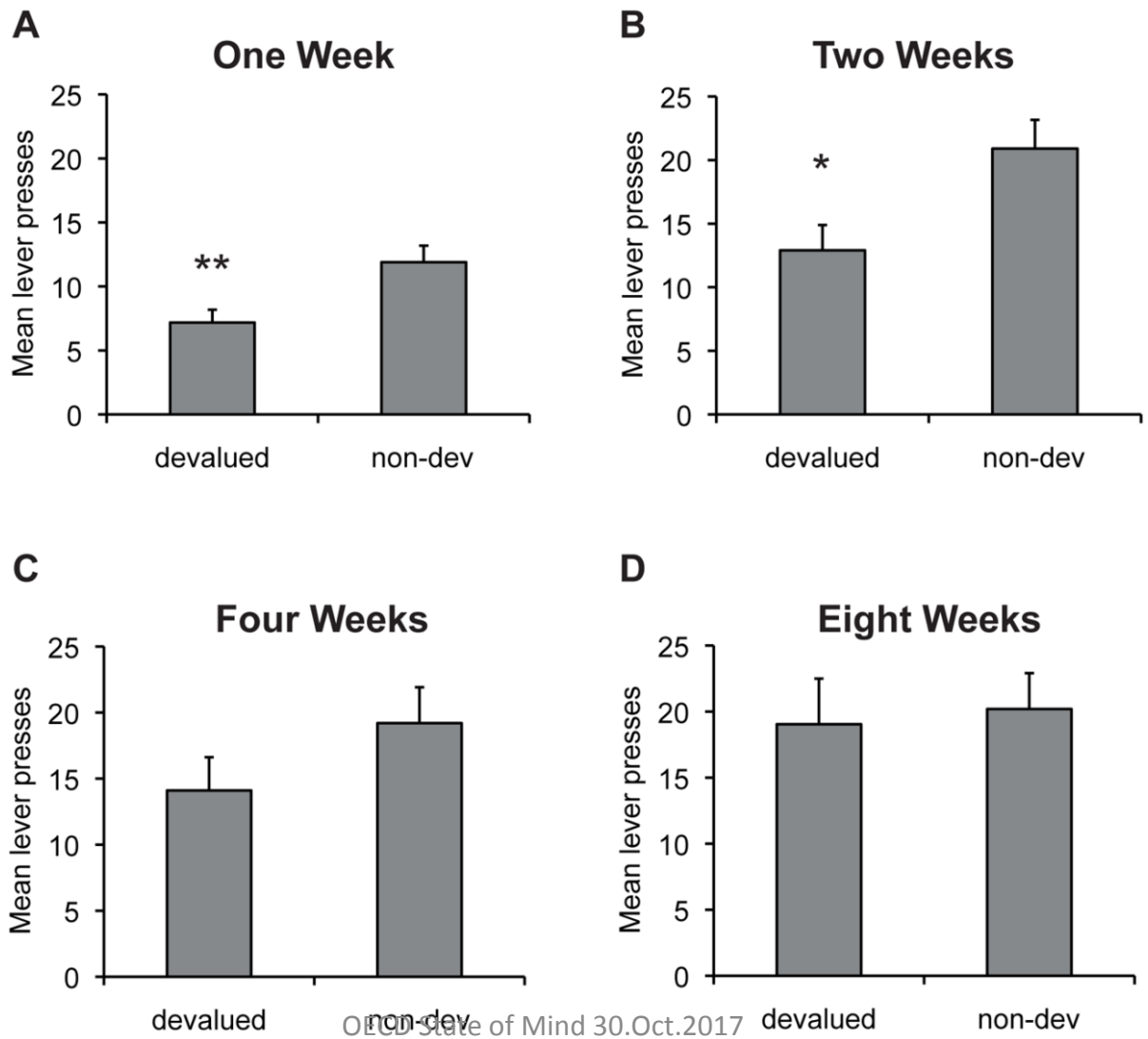


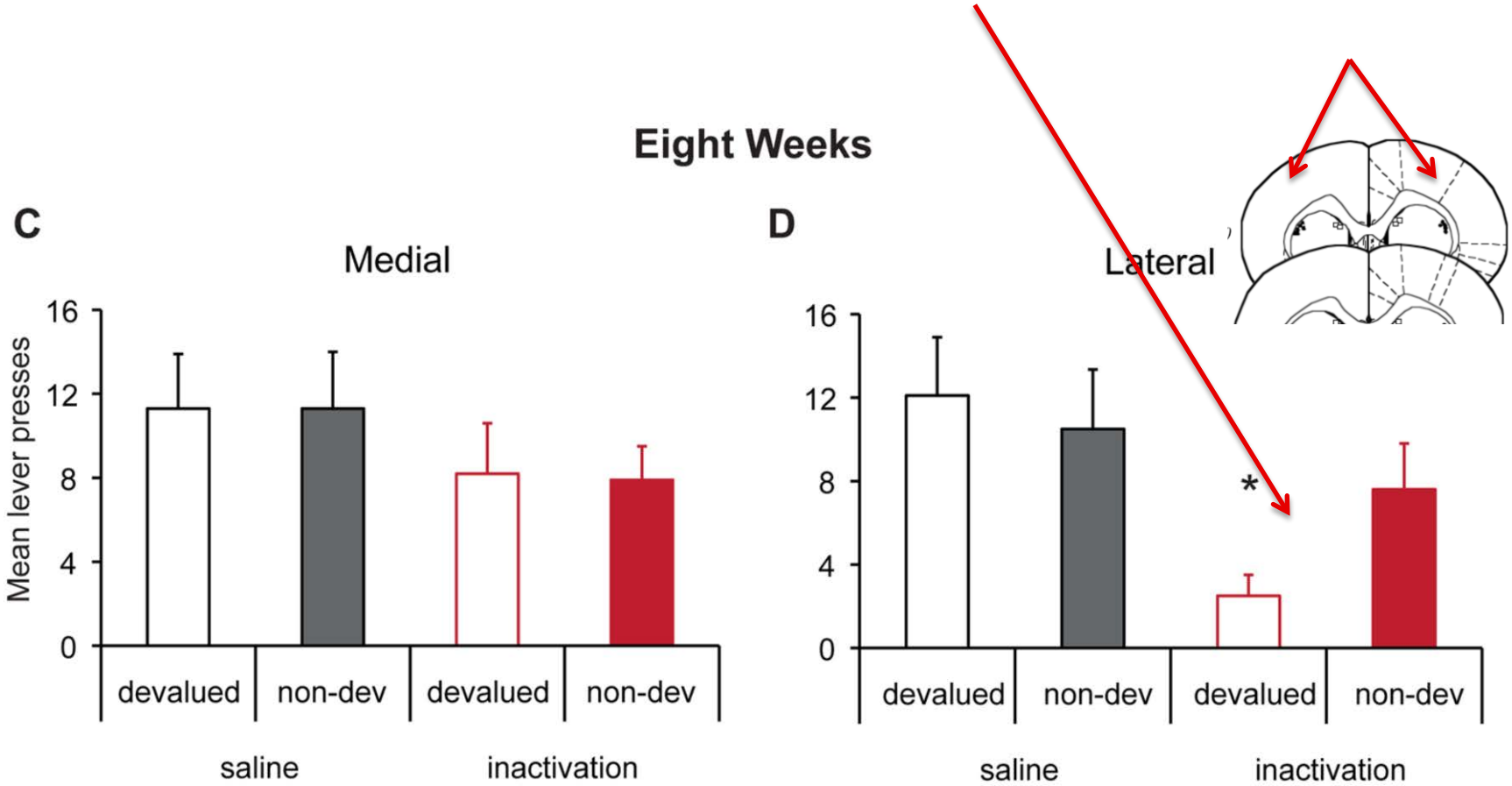
FIGURE 21.1 Distinguishing habitual from goal-directed instrumental learning using outcome devaluation. Left: Rats are first trained, when hungry, to lever press for food. Center: the food is devalued, e.g., by feeding the animal to satiety. Right: animals are tested to assess whether they will maintain lever pressing for the devalued outcome, compared to control animals who still value the outcome. The test is conducted in extinction (without food delivery) to ensure that any changes in behavior relate to the animal's internal representation of the outcome, rather than learning from new experience with it during the test phase. *Drawings by Sam Constantino.* Bottom: both devaluation-sensitive (goal-directed) and devaluation-insensitive (habitual) responses are observed under different circumstances. In this graph, the effect of the amount of instrumental training is illustrated (*data replotted from Holland, 2004*): goal-directed behavior dominates early, but gives rise to habitual (devaluation-insensitive) behavior following overtraining.

Development of ethanol habit in rats

(Corbit+ BioPsych '12)



Inactivation of dorso*lateral* striatum erases habit, restores devaluation sensitivity



concepts of habit in economics:

- reference-dependent preferences in macro-finance (Campbell, Cochrane JPE 1999) (Rozen Ecma 2010)

nation is *external*, as in Ab
formulation or Duesenbe
individual's habit level de
ption rather than on the

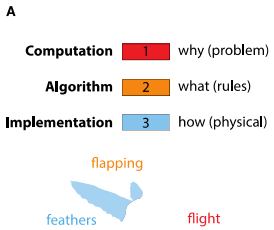
$$U_h(c) = \sum_{t=0}^{\infty} \delta^t u\left(c_t - \sum_{k=1}^{\infty} \lambda_k h_k^{(t)}\right)$$

- Adjacent complementarity (Pollak JPE 70; Becker-Murphy JPE 88; Crawford RES '10)

$$\max_{\mathbf{q}_t^c, \mathbf{q}_t^a} \sum_{t=1}^{\infty} \beta^{t-1} u\left(\mathbf{q}_t^c, \mathbf{q}_t^a, \mathbf{q}_{t-1}^a\right)$$

2. A computational model

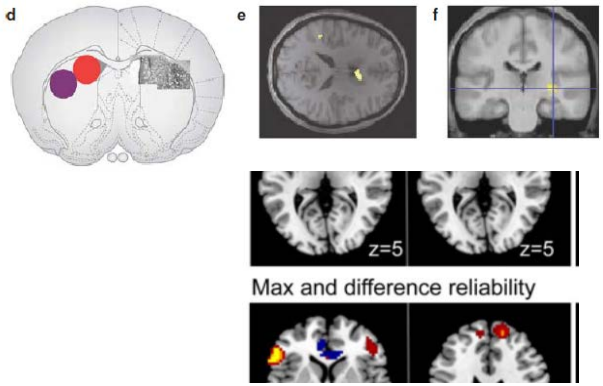
- Model-based mode maximizes utility
- habit mode computes updated reward value
  - compute “doubt stock” (reliability) based on absolute prediction error
- Choice is reliably rewarding? → habit
- Rewards unreliable? → model-based



three levels of habits

- Why? (function)
- What? (algorithm)
- How? (mechanism)

- Manages tradeoff:
  - optimal when rewards reliable
  - Saves scarce cognition
- RL learning of reward, reliability signal, shift MF-MB



NB: economic theories do not address Why? or How?

Habit system tracks reward perception from subjective value u(x)

λ_r =learning rate

→ 1 = fast

→0 = slow

$$r_t(x) = \begin{cases} (1-\lambda_r) r_{t-1}(x) + \lambda_r u_{t-1}(x) & \text{if } c_{t-1} = x \\ r_{t-1}(x) & \text{if } c_{t-1} \neq x, \end{cases}$$

Reliability or “doubt stock” cumulates absolute reward prediction error

$$d_t(x) = \begin{cases} (1-\lambda_d)d_{t-1}(x) + \overbrace{\left|r_t(x) - u_t(x)\right|}^{\text{absolute value of RPE}} & \text{if } c_{t-1} = x \\ (1-\lambda_d)d_{t-1}(x) + 1 & \text{if } c_{t-1} \neq x, \end{cases}$$

Choosing

$$c_t = \begin{cases} c_{t-1} & \text{if } \; d_t(c_{t-1}) < \sigma \\ \arg\max_{x \in \{A,B\}} u_t(x) & \text{otherwise.} \end{cases}$$

- Reflects desired properties of transitions between systems:
  - Pref.-to-habit: requres repeatedly choosing the same option
  - Habit-to-pref.: requires sufficiently large "surprise" (positive or negative)

3. predicting price elasticities

- We consider the effect of a change in the utility of each good at time τ

$$u_\ell(x) = \begin{cases} \bar{u}^x & \text{if } t < \tau, \\ \bar{u}^x - \Delta^x & \text{if } t = \tau. \end{cases}$$

- Interpret $\Delta^x > 0$ as price increase, $\Delta^x < 0$ as price decrease for good x

pre-habit maximization until time τ^*

Inexperience $\Rightarrow$ Preference-Based Choice

Proposition 1 $\tau < \tau^* \equiv \lfloor \ln(\sigma\lambda)/\ln(1-\lambda) \rfloor$ *implies, for all* Δ^A, Δ^B*:*

$$Q^A(\Delta^A, \Delta^B) = \tfrac{1}{2}\left(1 - \Delta^A + \Delta^B\right)$$

$$\eta^A(\Delta^A|\Delta^B) = -\frac{\bar{p}}{1-\Delta^B} < 0, \quad \eta_c^A(\Delta^B|\Delta^A) = \frac{\bar{p}}{1-\Delta^A} > 0$$

$$c_\tau = \arg\max_{x \in \{A,B\}} u_\tau(x), \quad \textit{for all consumers}$$

Zone of Perfect Inelasticity

Corollary 1 *Given* $\tau \geq \tau^*$, $-\Delta^* < \max\{\Delta^A, \Delta^B\} < \Delta^*$ *implies* $\eta^A(\Delta^A|\Delta^B) = \eta_c^A(\Delta^B|\Delta^A) = 0$.

$$\Delta^*_{(\tau)} \equiv \sigma - \frac{(1-\lambda)^\tau}{\lambda} > 0$$

Range of price changes → inelastic grows over time

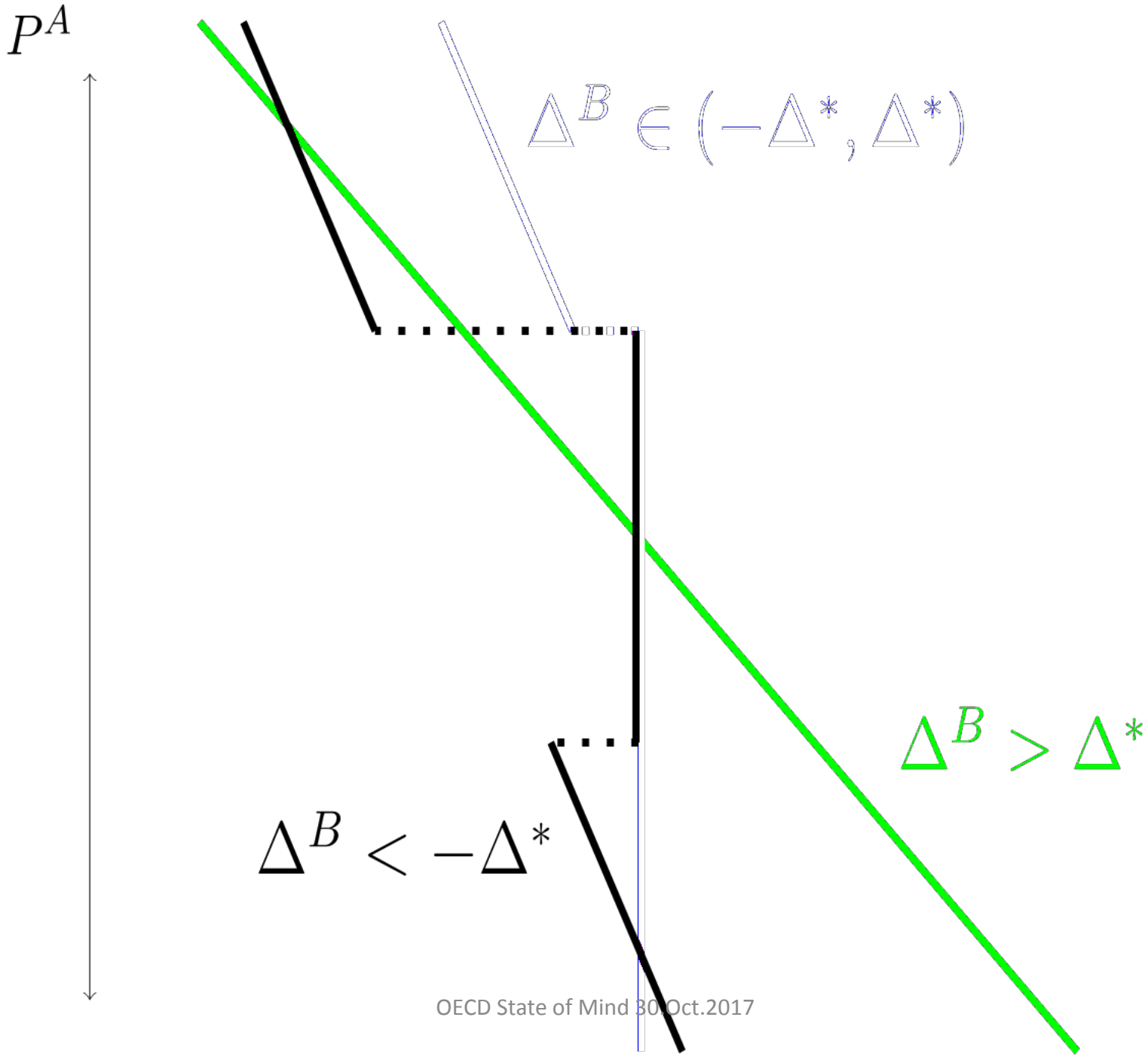

P^A
$\Delta^B \in (-\Delta^*, \Delta^*)$
$\Delta^B > \Delta^*$
$\Delta^B < -\Delta^*$

Table 7. Review Results from Goodwin (1992)

Dependent variable	Elasticity with respect to fuel price	
	short	long
Vehicle kilometres	-0.16 (n=4)	-0.32 (n=6)
Fuel Consumption	-0.27 (n=57)	-0.73 (n=53)

(Results originally identified as 'ambiguous' or 'unspecified' are not included in this table)

Table 8. Review Results from Espey (1998)

Dependent variable	Elasticity with respect to fuel price	
	Short	long
Fuel consumption	-0.26 (n=277)	-0.58 (n=363)

Long-run vs short-run: Suggestion, but no parametric definition

- "Habits that are difficult to break also cause differences between the short-run and long-run elasticity....*The short run is simply a period of time before people have made all their adjustments or changed their habits*; the long run is a period of time long enough for people to make such adjustments or change their habits." (Taylor, *Principles of Microeconomics* '95)

Long-Run vs. Short-Run Elasticities

Proposition 6 *If* $|\max\{\Delta^A, \Delta^B\}| \in (\tilde{\Delta}, \Delta^*)$, *then* $|\tilde{\eta}^A(\Delta^A|\Delta^B)| > |\eta^A(\Delta^A|\Delta^B)| = 0$ *and* $|\tilde{\eta}_c^A(\Delta^B|\Delta^A)| > |\eta_c^A(\Delta^B|\Delta^A)| = 0$ *where* $\tilde{\Delta} \equiv \sigma\lambda(1-\lambda)^{\left(\frac{(1-\lambda)^{\tau-1}-(1-\lambda)}{\lambda}\right)}$. *Otherwise,* $|\tilde{\eta}^A(\Delta^A|\Delta^B)| = |\eta^A(\Delta^A|\Delta^B)|$, $|\tilde{\eta}_c^A(\Delta^B|\Delta^A)| = |\eta_c^A(\Delta^B|\Delta^A)|$.

- Long-run demand generally more elastic than short-run demand

When is short vs. long-run transition?

- If change in price of A is Δ^A

  elasticity increases after τ periods where

  $\Delta_L < \Delta^A < \Delta_H$ satisfies

$$\Delta_L \equiv \sigma\lambda(1-\lambda)^{[(1-\lambda)^{\tau-1}-(1-\lambda)\ ]/\lambda}$$
$$\Delta_H \equiv \sigma - (1-\lambda)^{\tau}/\lambda$$

e.g. λ =.3

Δ_L *goes from* $.22\sigma$ *to* $.75\sigma$ *as* $\tau \rightarrow \infty$

Δ_H *goes from* σ-2 *to* σ *as* $\tau \rightarrow \infty$

Period at which “long run” begins

σ	λ	**period**	Δ^A
0.08	0.3	14	5.5%
	0.25	16	4.8%
	0.2	21	4.0%
	0.15	31	3.0%
	0.1	49	2.1%
	0.075	69	1.5%

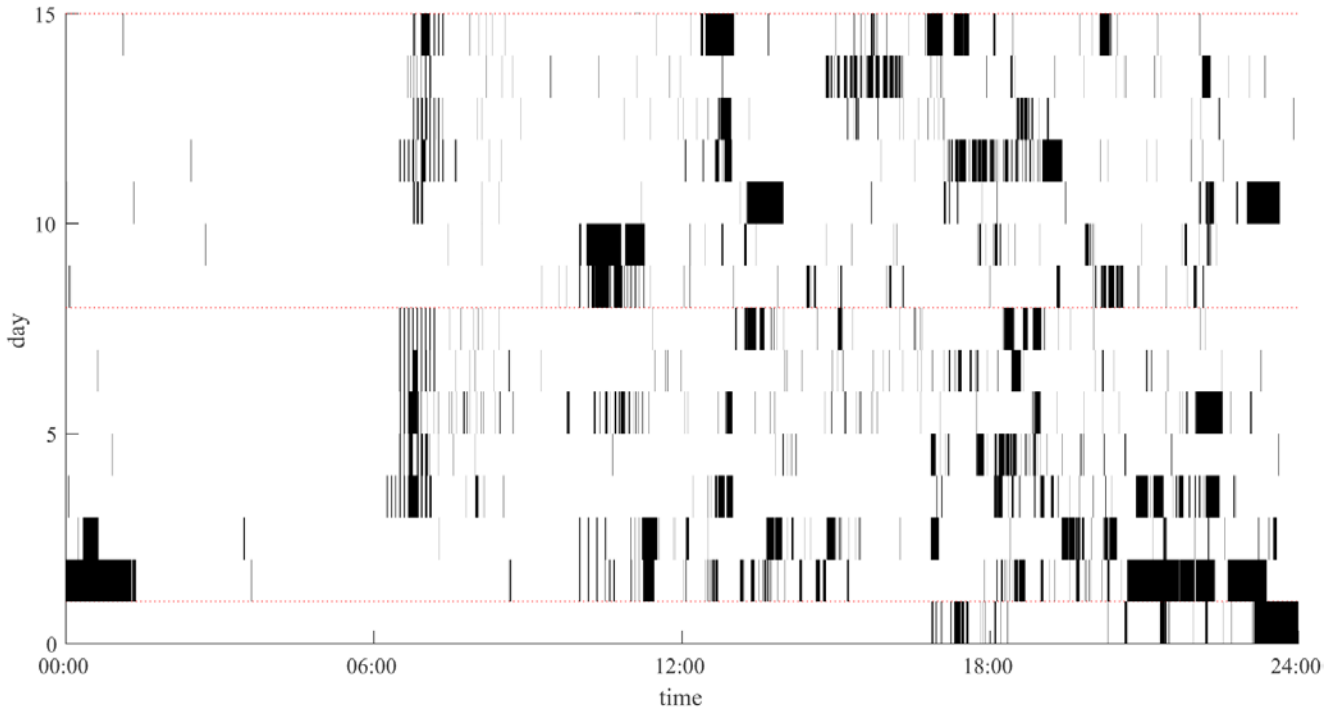
4. Other empirics

- Cognitive and empirical markers of habit
  - High inferred search costs
  - Effect of product unavailability
  - Drug overdose

impaired memory for habitual choice: my Amazon spouse

Cellphone use:
Actual checks (84.7) >> estimated (37.20) actual duration (5.05 hrs) > estimated (4.12 hours)

(Andrews PlosOne 2015)



day alarm clock times (and snoozing) are clearly evident.

doi:10.1371/journal.pone.0139004.g001

Other evidence for habit-model approach

- Change after product unavailability (London Tube strike)
- High search cost shadow prices
- Pavlovian ‘habitual’ cues cause drug overdose

Interruption of commuting habit

- London 48-hour *partial* strike (Larcom+ '16)
- great quasi-experiment
  - Natural control group (no-strike lines)
  - Tube map not drawn to scale (=search value)
  - Oyster swipe-card data

What happened?

- 5% of interrupted riders switched to a better route
- Time savings = 40secs/day
- Economic value = £380

High inferred search costs

- Most habitual
  - Consumer products $2.10/min=$126/hr
    (rfid shoppingcarts)

- Not so habitual
  - One extra online bookstore $4.14
  - Online textbooks: 50% no search, searchers 1st extra search value $1.30-19
  - Computer chips $8

(Pinna-Seiler wp 16; De Los Santos+ AER 12; Hong-Shum RAND 06; Moraga-Gonzalez JAppEconometrics 13;)

Conclusions about habit

- Neuroeconomic model of habit + model-based
- Neuro-economic
  - predictions about elasticities
- Neuro-economic
  - Concept of habit from ethology, neuroscience, computation
  - A *computational control process*, not a change in *preference*

  → many predictions about habitual cognition

Brain properties can constrain economic theory

"We do not lack models, we lack the ability to reject some of them on the basis of a well formulated strategy research. Neuroeconomics can be a valuable tool in this direction." (Rustichini, AEJ:Micro 2009 p 58)

Properties

- Structural
  - Modularity
  - Plasticity
  - Connectivity
- Functional
  - Repurposing
  - Compression, efficient coding, normalization
  - Explicit vs. implicit
  - Learning >> maximization

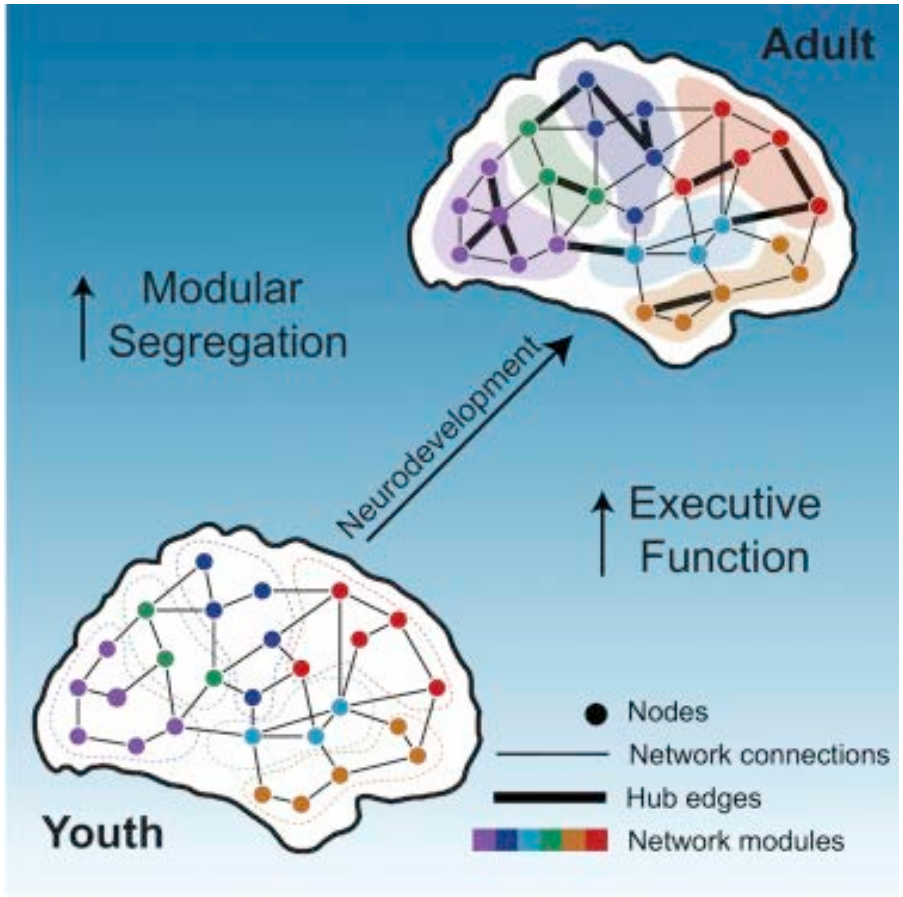
Properties (today's talk)

- Structural
  - Modularity
  - Plasticity
  - Connectivity
- Functional
  - Neural repurposing
  - Compression, efficient coding, normalization
  - Explicit vs. implicit
  - Learning >> maximization

Brain properties: **Structural**

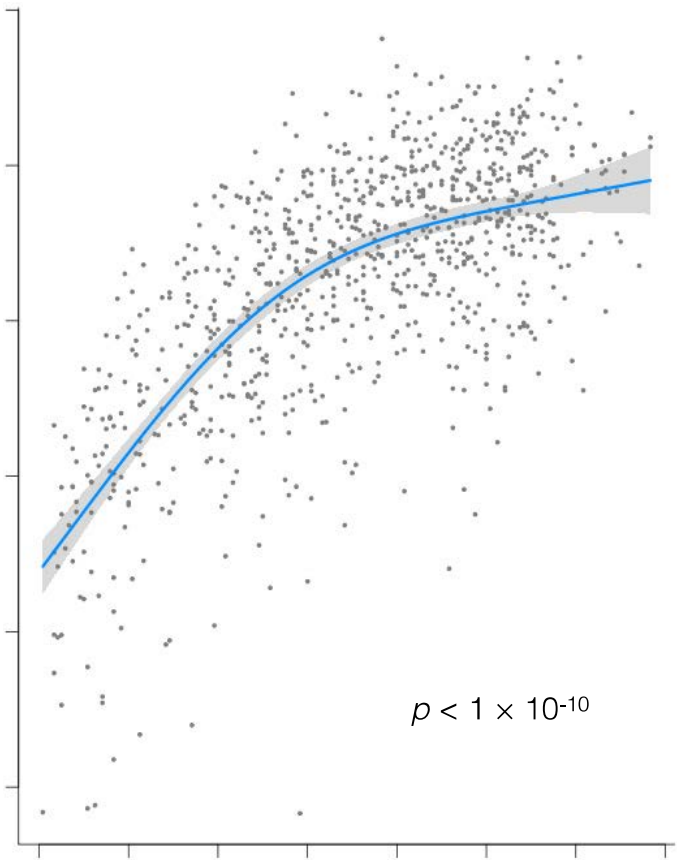
- Modularity
- Plasticity
- Connectivity

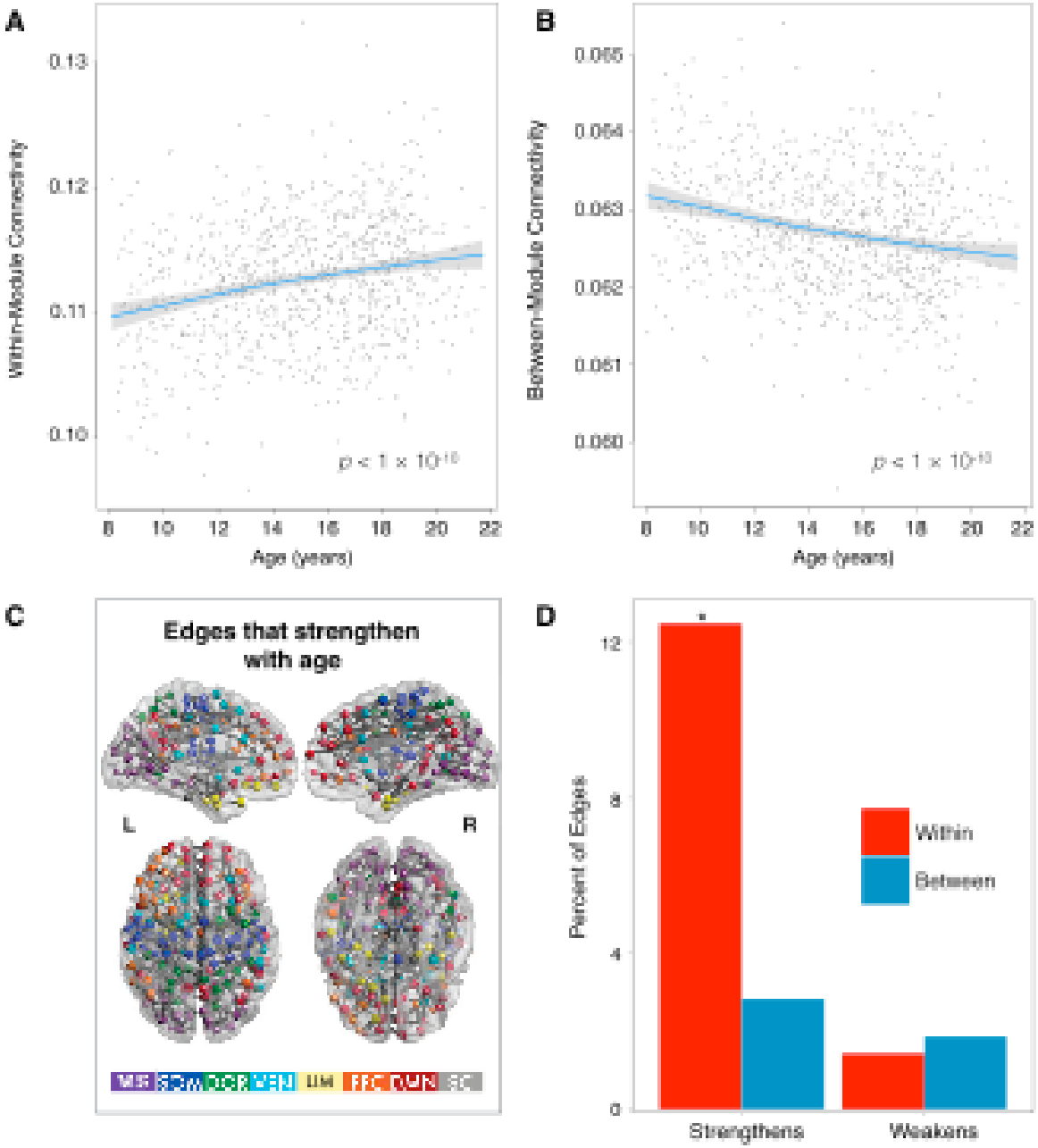
Structural: Connectivity, development, and executive function



Baum+ CurrBio 2017

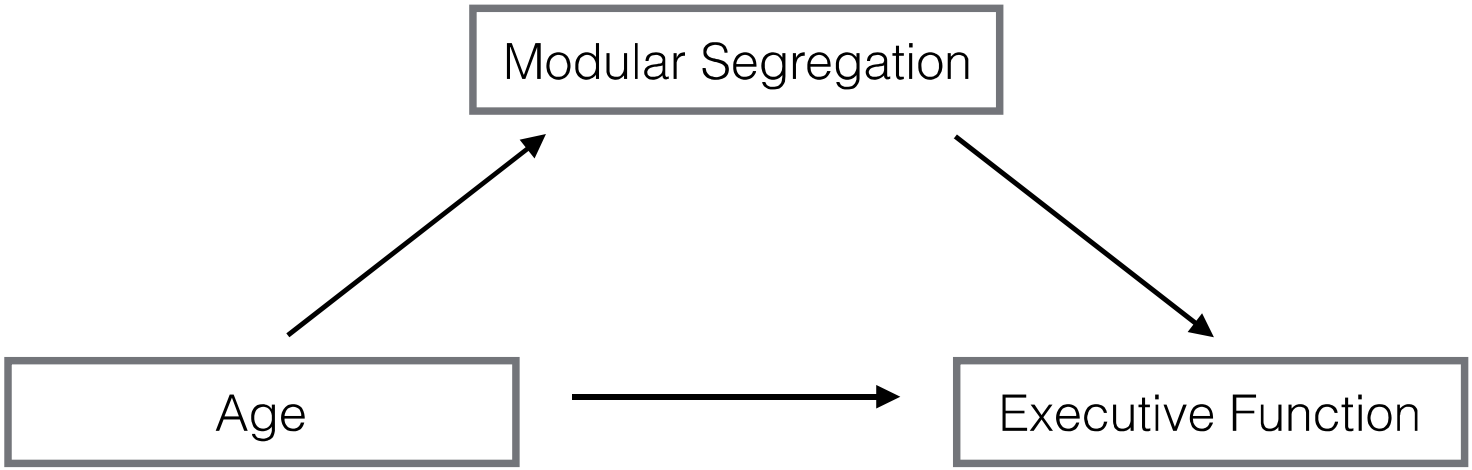
Executive function (EF) increases 8-22 years (N=880)




A
Within-Module Connectivity
0.13
0.12
0.11
0.10
$p < 1 \times 10^{-10}$
8 10 12 14 16 18 20 22
Age (years)
B
Between-Module Connectivity
0.065
0.064
0.063
0.062
0.061
0.060
$p < 1 \times 10^{-10}$
8 10 12 14 16 18 20 22
Age (years)
C
Edges that strengthen with age
L
R
D
Percent of Edges
12
8
4
0
Within
Between
Strengthens
Weakens

Brain connectivity → EF

~~"age → EF"~~ "segregation → EF"



Brain properties: **Function**

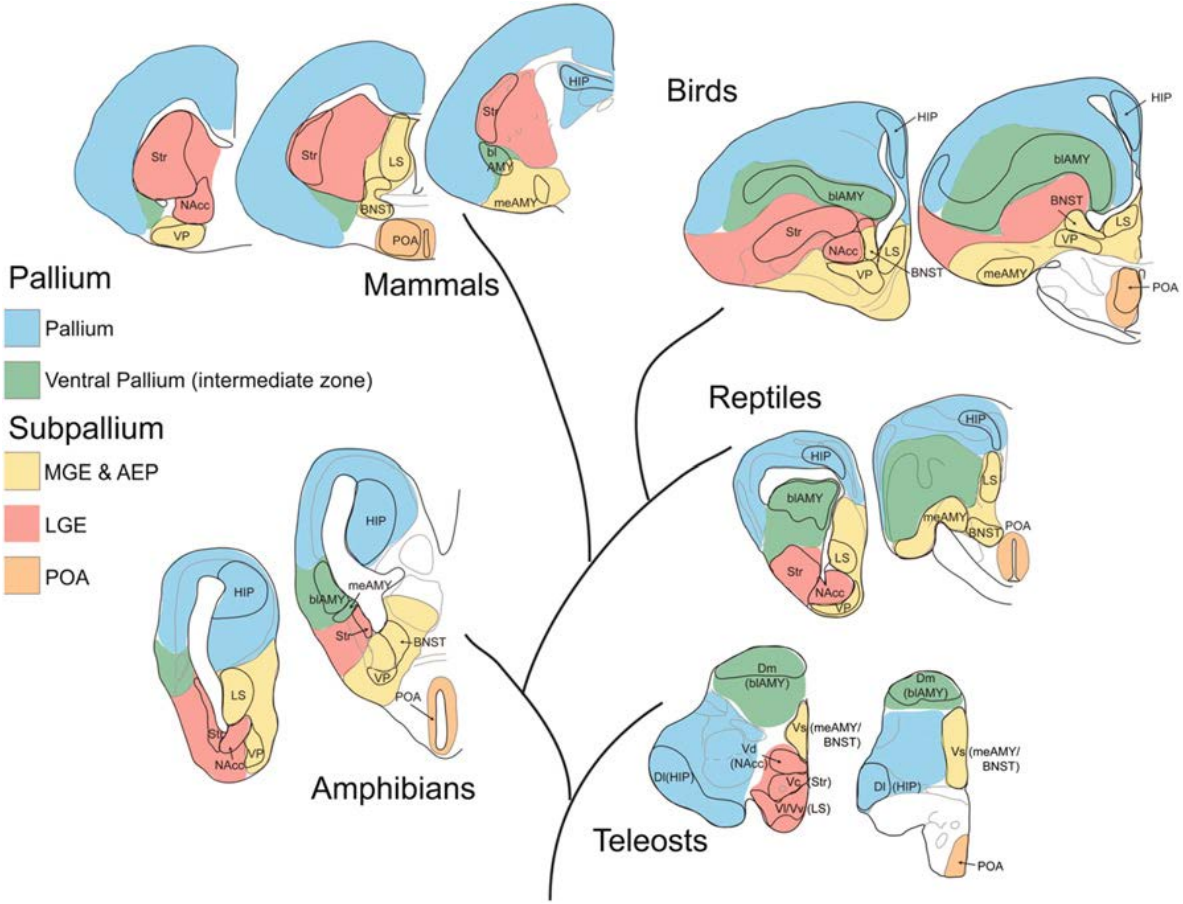
- Neural repurposing
- Learning is the fundamental "primitive" process (*not* utility-maximization*)
  - Maximization framework replaced by "When does learning result in approximate maximization?"
- Compression, adaptive coding, normalization
- Implicit vs. explicit knowledge

Functional: Repurposing

- "In many instances, and without any well-defined long-term project, the tinkerer picks up an object which happens to be in his stock and gives it an unexpected function. Out of an old car wheel, he will make a fan; from a broken table, a parasol. ***This process is not very different from what evolution performs when it turns a leg into a wing,*** or a part of a jaw into pieces of ear."
  - *(*Jacob, Francois. *The Possible and the Actual* 1994)

Economic implication: "2nd best". Will be kludges, not blank-slate optimal designs (Ely AER)

Evolutionary conservation of the social behavior network (SBN)

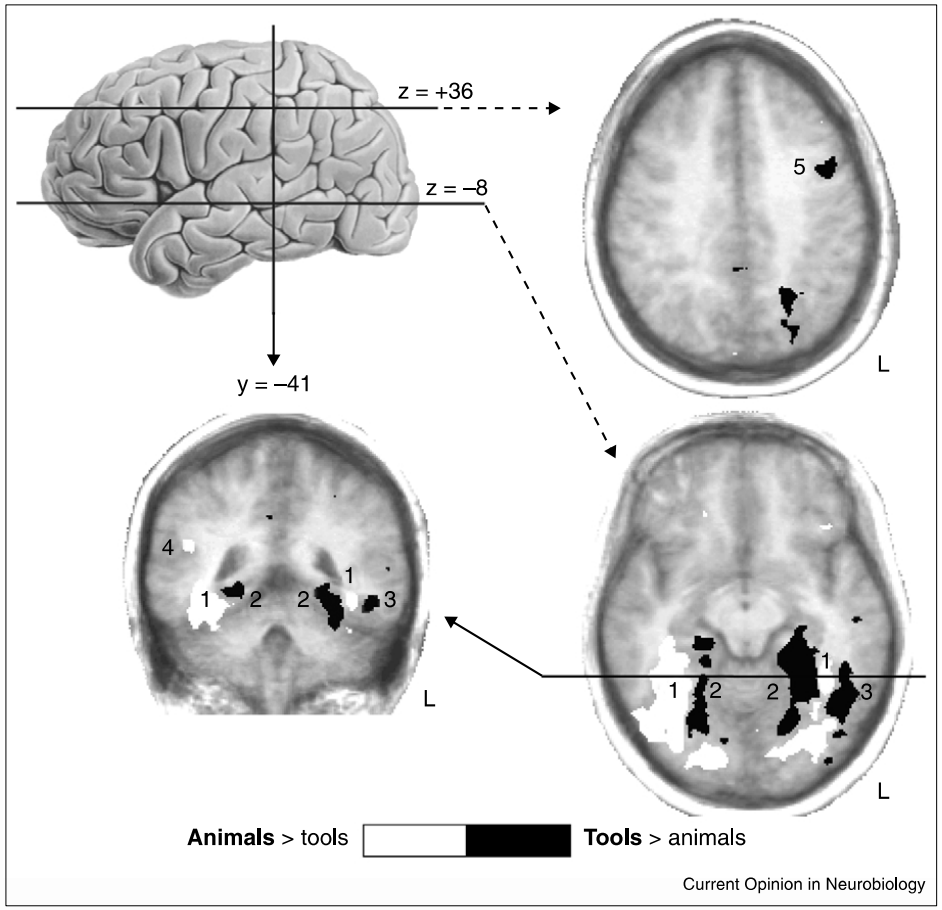


Baran+ PNAS 2017

the brain repurposes body growth encoding-- a tool is an arm

"Monkeys [were] trained to use a tool (a long rake) to pull food forward toward them, for two weeks…. the neural motor coding system is prepared to adapt to a mammal's arms getting longer during body growth. That same capacity is then ***repurposed*** to encode the location of a *tool* in motor cortex, and adjacent regions, as if it were an extension of the *arm"* (Iriki and Taoka, 2012)

Language: motor (tools), visual (animals)



TMS increases motor response to verb recognition 500msec (Papeo 09

Martin+ Sci 1995, Nature 1996

SNARC

Spatial-numerical association of response codes (SNARC) (Dehaene+ JXP:G 2003)

- “mental number line” imagined spatially

X

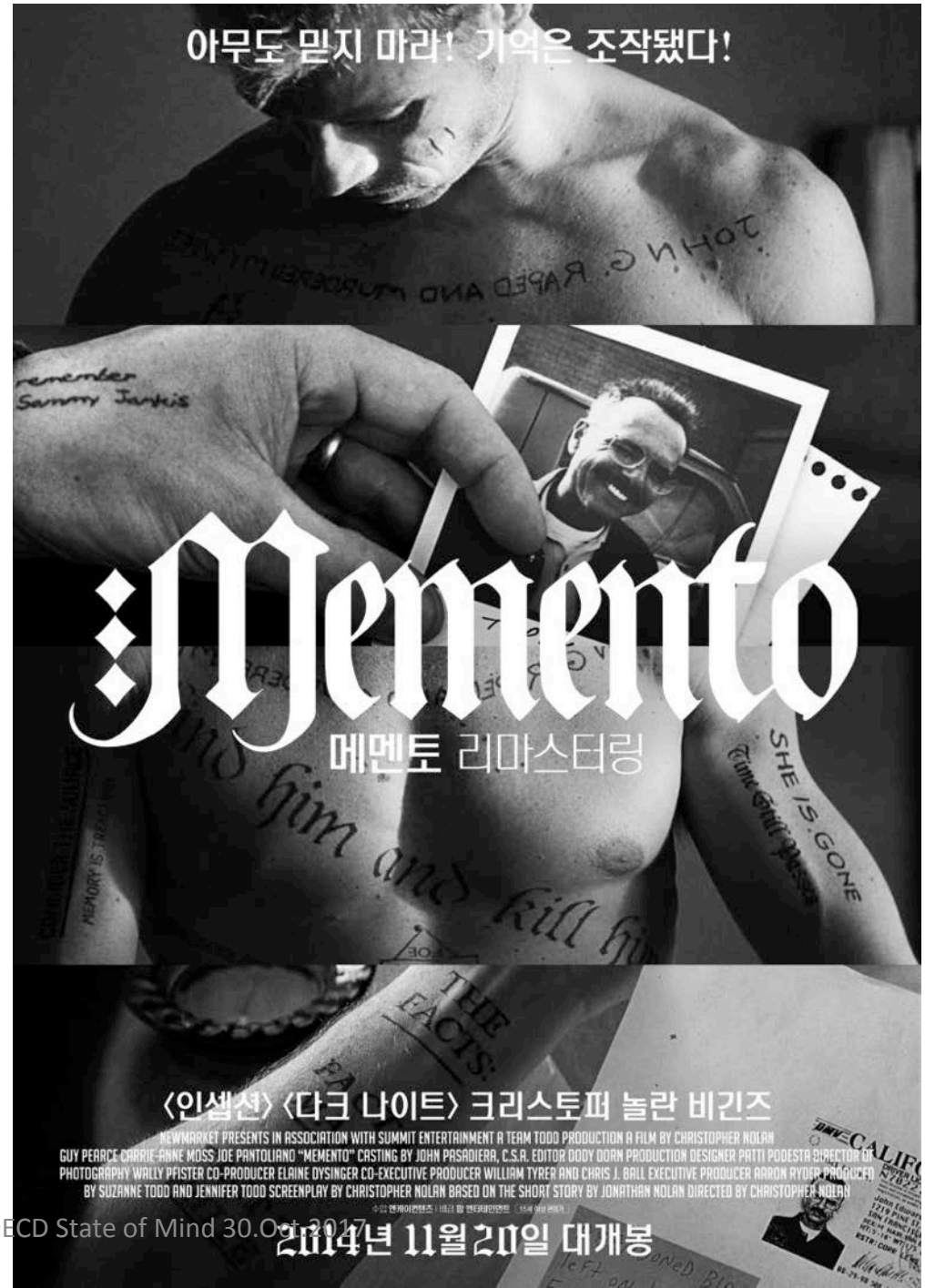
← 2

9 →

Implicit vs explicit

- An old, robust distinction (a/k/a conscious vs unconscious)
- Explicit = verbalizable and corresponds to behavior
- Implicit = not verbalizable *but* revealed by behavior

Implicit knowledge: Claparède's patient



implicit knowledge: autistic martial artist in “Chocolate”

https://en.wikipedia.org/wiki/Chocolate_(2008_film)

explicit knowledge (math PhD)

Friedman-Savage "as if" billiard player hypothesis (1948 JPE)

- It would in no way disprove or contradict the hypothesis, or weaken our confidence in it, if it should turn out that the billiard player had never studied any branch of mathematics and was utterly incapable of making the necessary calculations: *unless he was capable in some way of reaching approximately the same result as that obtained from the formulas, he would not in fact be likely to be an expert billiard player.*

Friedman-Savage “as if” billiard player hypothesis (1948 JPE)

- It would in no way disprove or contradict the hypothesis, or weaken our confidence in it, if it should turn out that the billiard player had never studied any branch of mathematics and was utterly incapable of making the necessary calculations: unless he was ***capable in some way of reaching approximately the same result*** as that obtained from the formulas, he would not in fact be likely to be an expert billiard player.

Billiard player knowledge

- Observe expert play → could be explicit or implicit
- Economics profession →
  - Implicit/explicit distinction is not important

WRONG

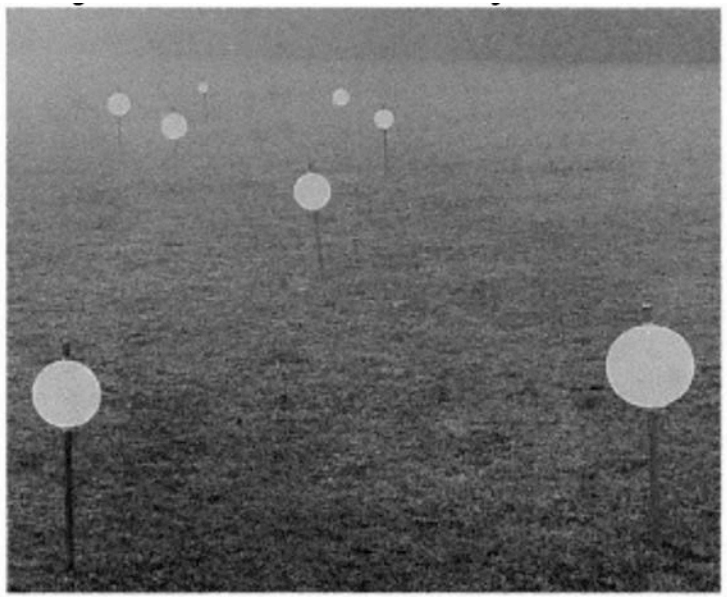
Explicit vs implicit

- Explicit knowledge more flexible to change
  - Billiards with more friction
  - Playing croquet
    - Explicits will play better

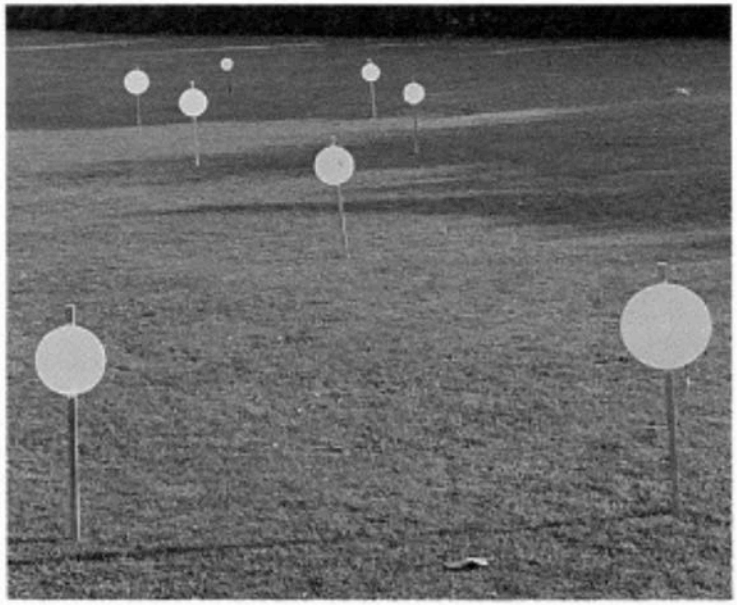
- Coaching
  - Those with implicit knowledge will coach worse
- Surveys, BDM prices (explicit) can be wrong
- Implicit easier to causally manipulate exogeneously

Example: Aerial illusion

- Judging distances
- (Implicit) upstream system gets (*d,a)*
- Reports E(distance|d,a) to (explicit) downstream system which integrates with higher-order information (e.g. measurement)
- *Explicit system cannot adjust for a*
  - → Overestimates distance for a>0 (hazy)
  - → Underestimates distance for a<0 (clear)

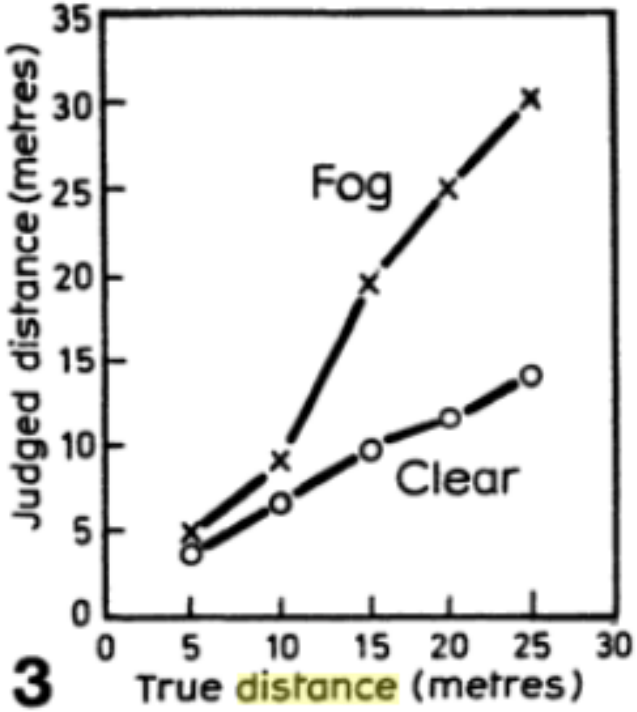


Morning Fog



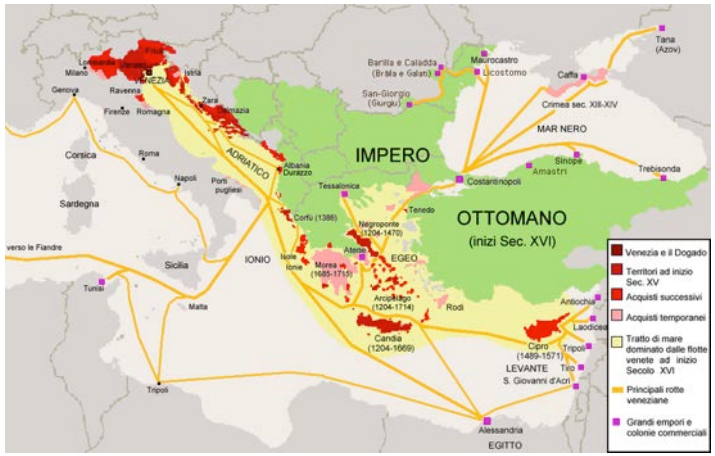
Midday Sunshine

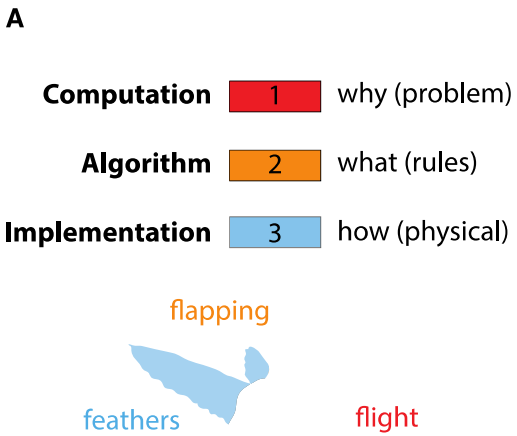
"haar" morning fog in East Coast UK (Hull)
"I had been waiting for this for some months…"



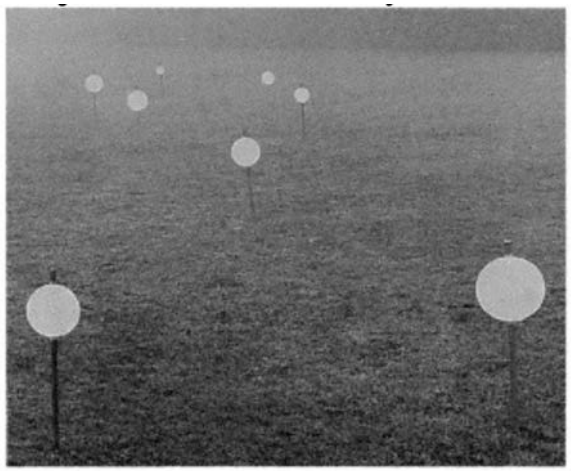
Ross, 1975 NewSci

recap





$$c_t = \begin{cases} c_{t-1} & \text{if } \; d_t(c_{t-1}) < \sigma \\ \arg\max_{x \in \{A,B\}} u_t(x) & \text{otherwise.} \end{cases}$$



Morning Fog

Thanks for inviting me!

Many methods used, *all* low marginal cost (except fMRI)

	fMRI	Lesions	TMS	EEG	Decoding	SCR	Disorders	Personality	Reference
High marginal equipment cost?	Yes	No	No	No	No	No	No	No	
High marginal labor cost?	Yes	No	No	Yes	Yes	No	No	No	
Ambiguity-aversion	√	√							Hsu et al. (2005)
Loss-gain framing*	√					√	√ (ASD)		De Martino et al. (2006, 2008), Roiser et al. (2009)
Time preference*	√							√	Peters & Buchel (2010)
Time preference			√						Figner et al. (2010)
Strategic thinking	√								Coricelli & Nagel (2009)
Private-information Bargaining*				√	√				Yun et al. (in prep)[a]
Neurally-revealed preference*	√				√				Smith et al. (2012)
Mechanism design	√				√				Krajbich et al. (2009)
Anxiety						√			Kang et al. (2012)
Emotional regulation	√					√			Sokol-Hessner et al. (2009)

Camerer, AnnRevEcon 2012

conclusion

- Neuroeconomics strives for 3-levels
  - Each level constrains and inspires level ↑ and ↓
- Neural evidence for basic economics
  - pain of paying
  - imagination and future valuation
- General properties of function and structure