LEARNING IN AN UNCERTAIN ENVIRONMENT II

Peter Bossaerts

CONTENT

I. Reinforcement learning… of (reward) risk
   - Link with outlier risk

II. Learning of stochastic affectively neutral stimuli
   - Specifically: Bayesian learning

MESSAGES - METHODS

- Why “contrasts” in imaging can be misleading
- Why knowing the biology behind behavior is important
- How “converging evidence” works
- Understanding models beyond the mathematics
- Experiments are not simply about “treatment” vs “control”

LEARNING OF

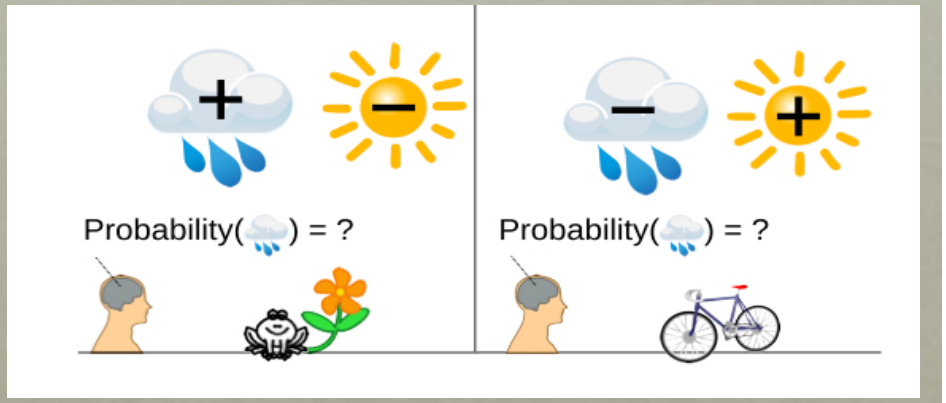
- … stochastic affectively neutral stimuli

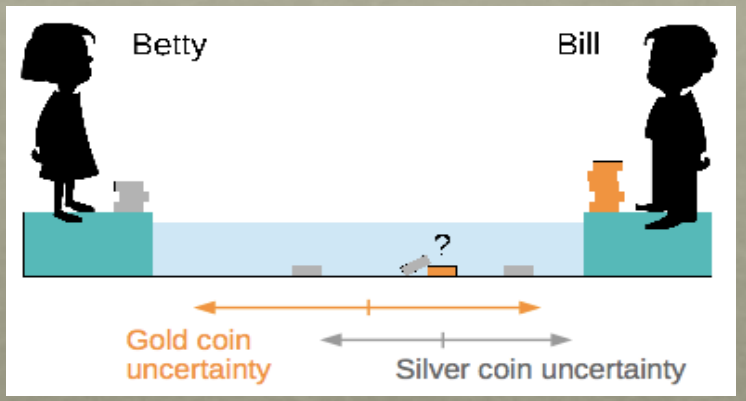
MOTIVATION

- In one important way, there has been a subtle bias in the study of belief formation in decision neuroscience:
  - *(Almost always) no dissociation between belief formation and valuation*
  - … because beliefs directly concern rewards (or losses)
- Dominance of "reinforcement learning" (Peter Dayan: Q-Learning; TD models and Dopamine; model-free, model-based approaches)
- (Perhaps bias originates in concern for animal model of learning)

Realistic belief formation tasks are often about stimuli/events without *a priori* affective content

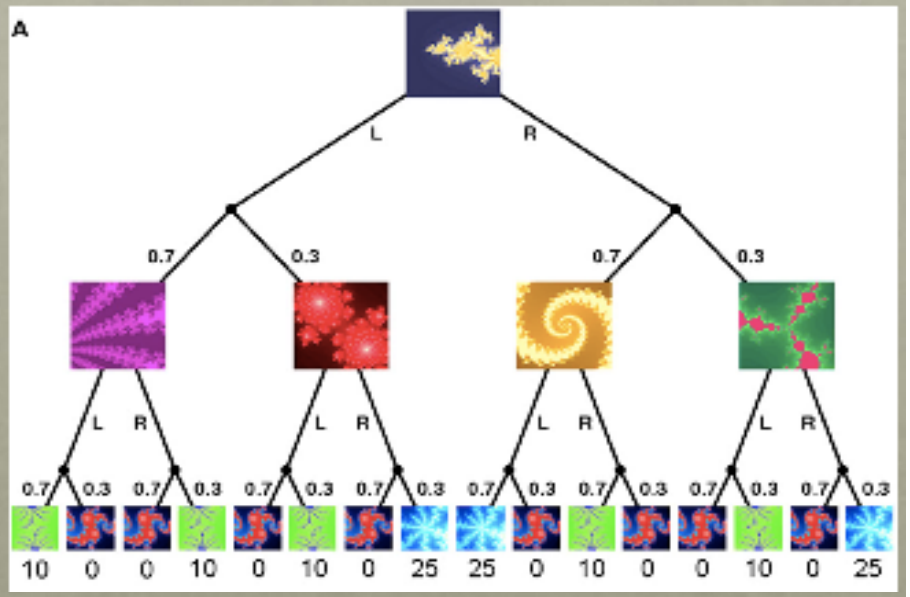
- Example 1: predicting rain before knowing whether rain is good (because one is to plant new flowers) or bad (when biking is required)
- Example 2: predicting where a coin landed (as opposed to being paid for where the coin landed)





SOME RECENT WORK PARTIALLY INVOLVING LEARNING OF PROBABILITIES OF STIMULI DEVOID OF VALUE

- … includes *Gläscher, Daw,* ***Dayan****, O'Doherty*, Neuron 2010
- Model-based reinforcement learning
- Learning of structure of stimuli involved assessing transition *probabilities* in 2-stage binomial tree
- Separation beliefs-rewards: *Third stimulus predicted particular reward, but association was not immediately known*



HERE:

- Belief formation unrelated to (model-based) reinforcement learning
- While stimuli/events (about which beliefs are formed) are not always strictly neutral (no valuation attached to them),
  1. During learning, one does not know value
  2. Even if one does, value association may change
- Probabilities about multiple (>>2) events
- Investigate the subjective ("Bayesian") nature of beliefs, not just objective ("frequentist") one

REMARK

- Belief formation *separate from* valuation is an important concept in decision theory
- Indeed, the ability to form beliefs unaffected by affective content of stimuli/events is part of rational choice
- Called "probabilistic sophistication" (*Machina and Schmeidler,* Econometrica 1992)
- Violated in, a.o., *ambiguity aversion* (where beliefs depend on payoffs: if a stimulus/event is associated with large rewards, assign low probability to it)

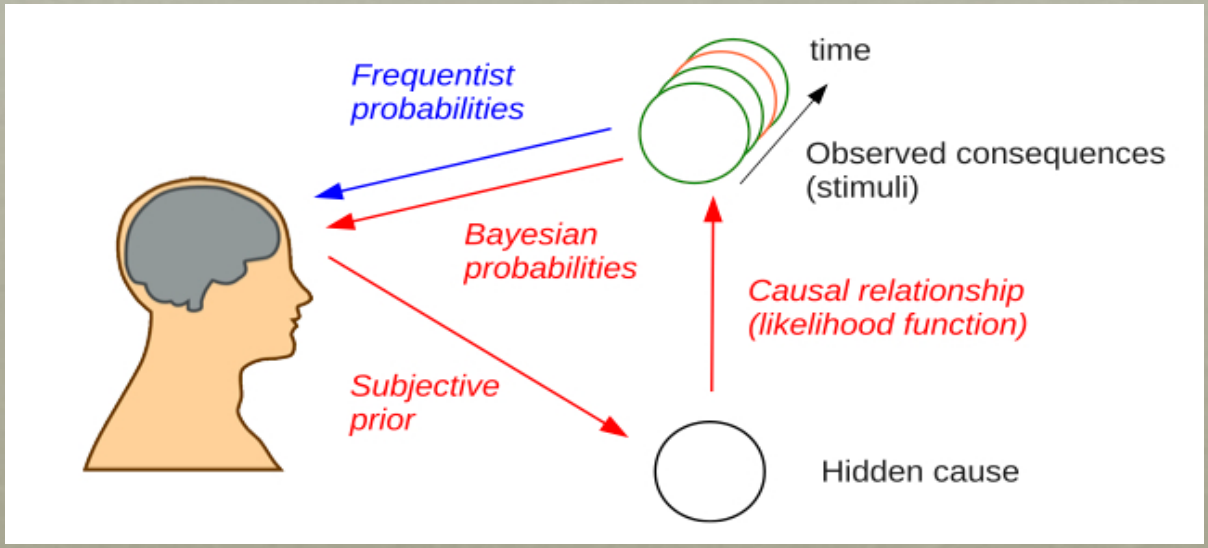
SOME QUESTIONS ONE MAY ASK ABOUT NEUROBIOLOGICAL FOUNDATIONS OF BELIEF FORMATION (SEPARATE FROM VALUATION)

1. How are beliefs encoded?
2. How is the evidence (frequencies) tracked?
3. How is prior information (prior evidence; subjective beliefs) encoded?
4. How is the prior merged with the evidence?
5. Specifically, does it involve Bayes' rule?

START WITH LAST QUESTION

- It is often asserted that human belief formation is *not* Bayesian
- (Incidentally, this contrasts with evidence from perceptual motor tasks – *Bayesian Brain*, Doya, Pouget, Rao, eds.)
- Specifically, updating is "too slow" (conservative)
- Is this conclusion warranted?

BAYESIAN BELIEF UPDATING

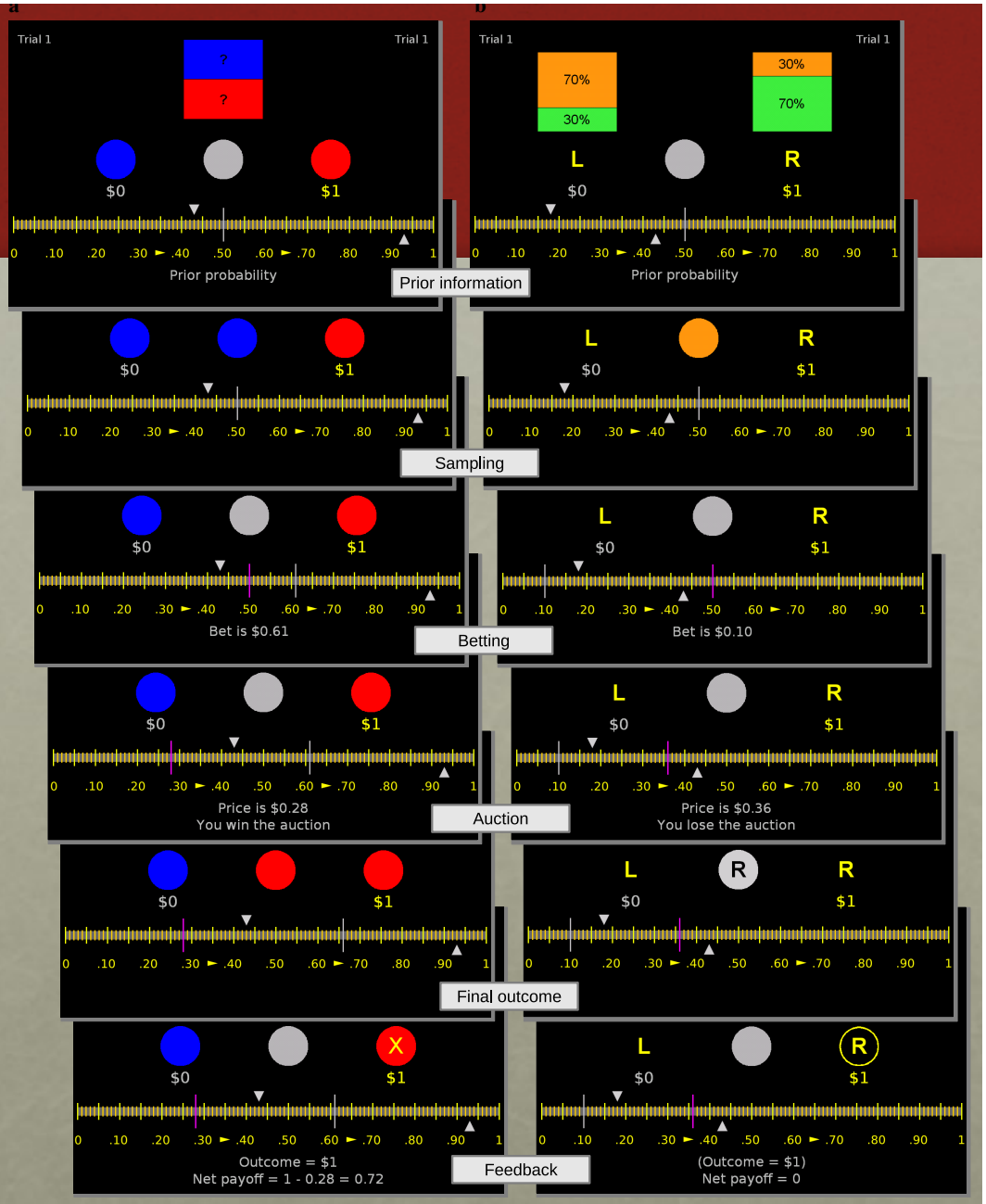


- Involves: Prior, Evidence (And Likelihood), Posterior
- To ascertain that updating follows Bayes' law, need to be sure that prior is controlled…

EXAMPLE

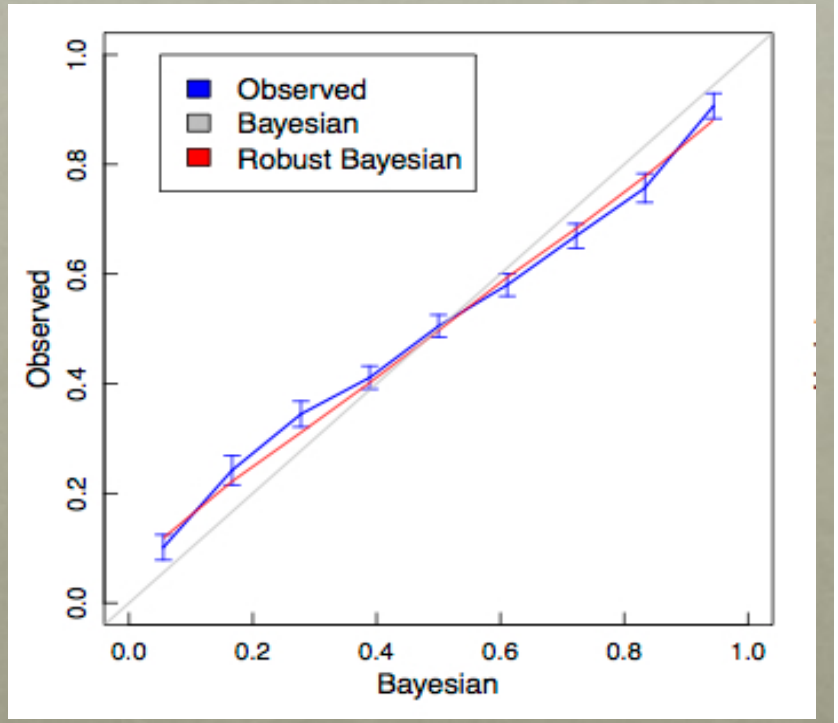
Right:

- Two urns (Left/Right); balls repeatedly drawn and displayed in middle
- From which urn do the draws come?
- Prior indicated with arrows
- After every draw, post a bet that it is the Right urn
- (Bet resolved with “second price auction”)



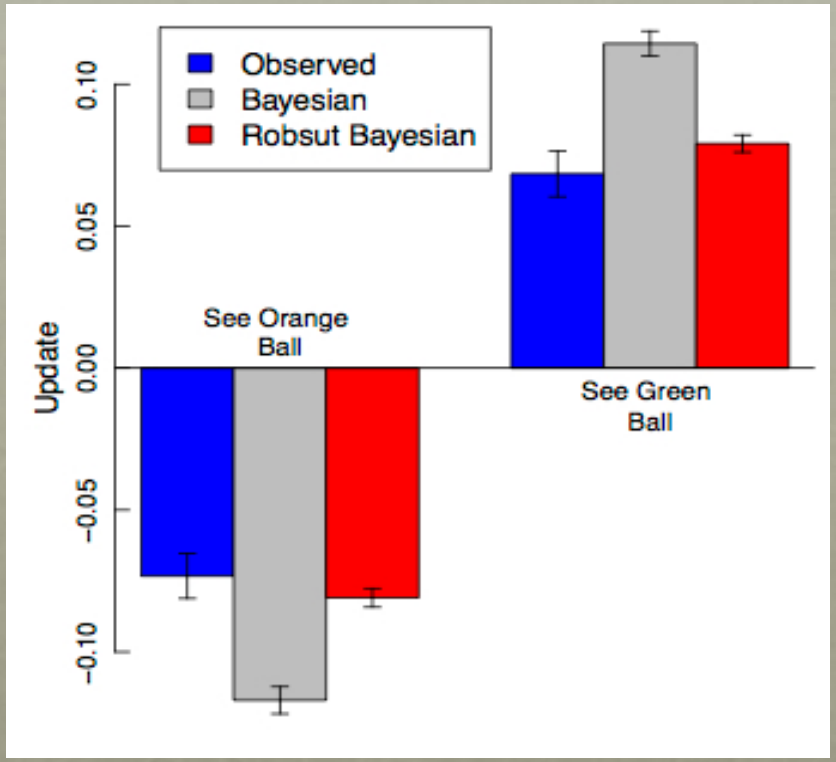
Assuming that subjects took indicated prior to be literally true...

- Compare blue line (observed beliefs) against grey one (diagonal)
- Belief updating is too conservative relative to Bayesian beliefs

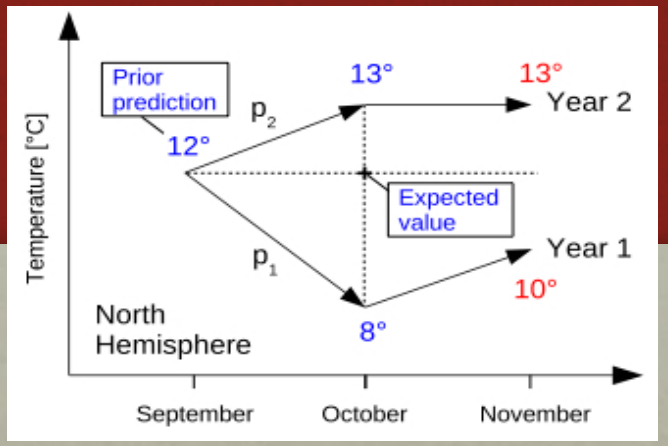


ASSUMING THAT SUBJECTS DOUBT VERACITY OF THE WAY DRAWING HAPPENS

- … while using a standard *robust prior* for in case the experimenter (we) had been cheating
- Now belief updating is Bayesian (see also previous slide; red line)
- Beliefs are *subjective*



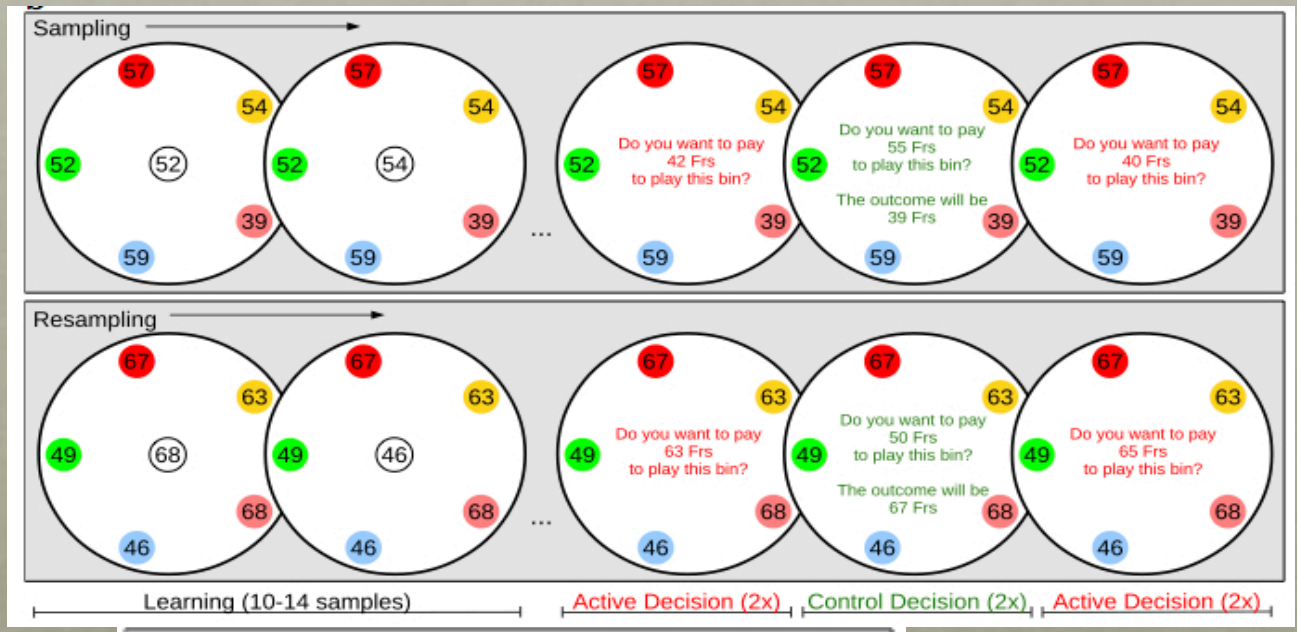
MAIN MESSAGE

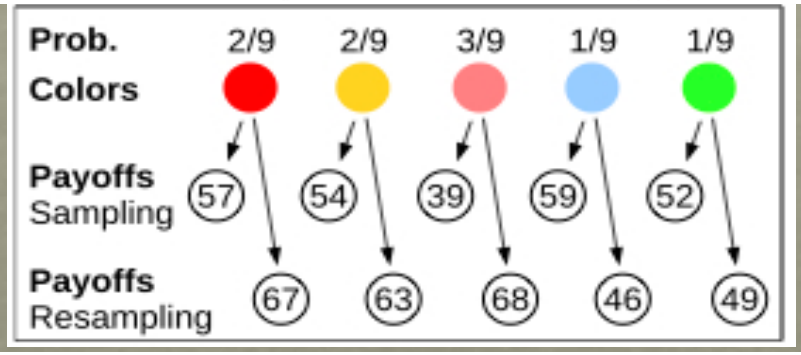


- We may want to use *general properties* of Bayesian updating that hold for a family of priors (rather than just the one the experimenter induces)
- E.g. martingale property: belief updates are not predictable (like a random walk)
- *(work with Mathieu d'Acremont, Leon Guerrero and Wolfram Schultz, re-submitted Cognitive Science)*

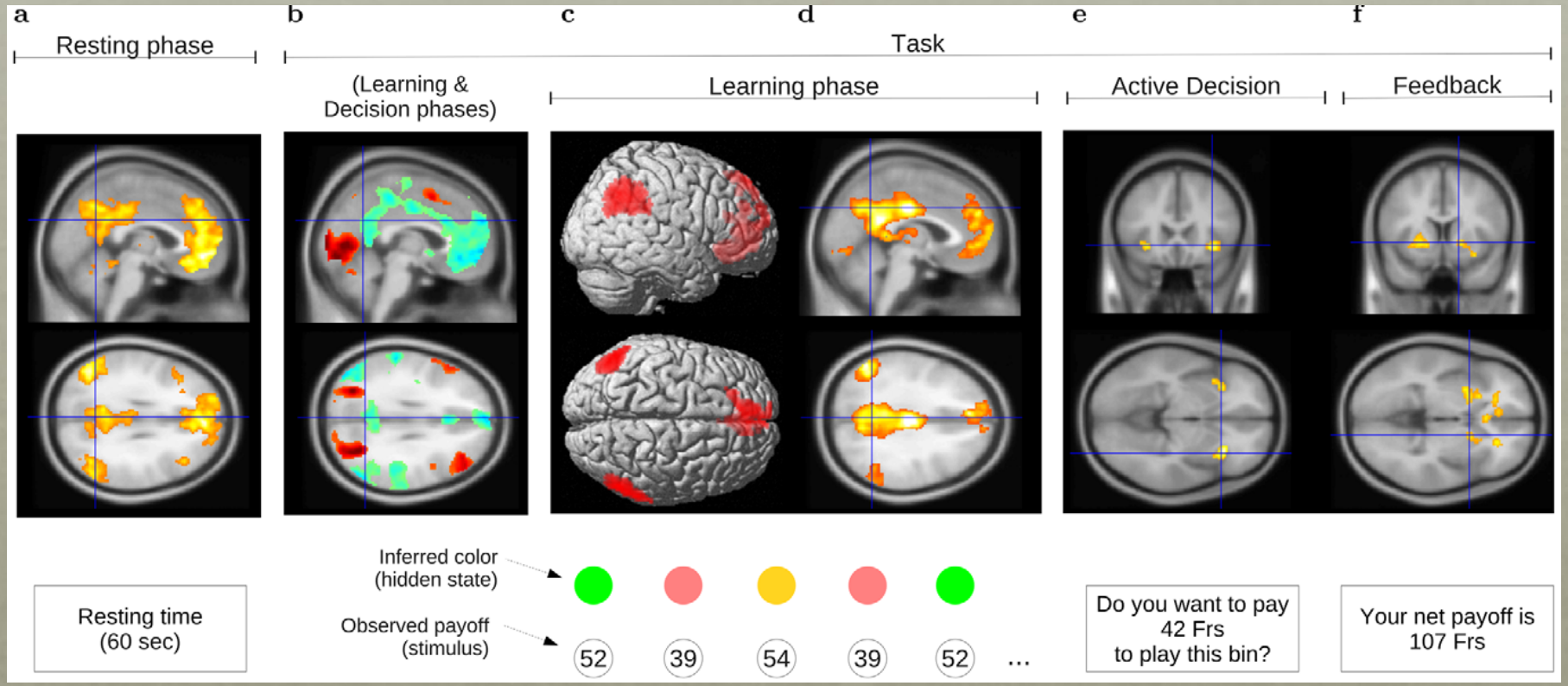
FIRST QUESTION: HOW ARE BELIEFS ENCODED?

- Learn composition (colors) of bin through sampling, then decide to buy into gamble at posted prices once rewards are associated with colors
- Change color-reward association, resample (same bin)
- *(With Mathieu d'Acremont and Eleonora Fornari, PLoS 2013)*





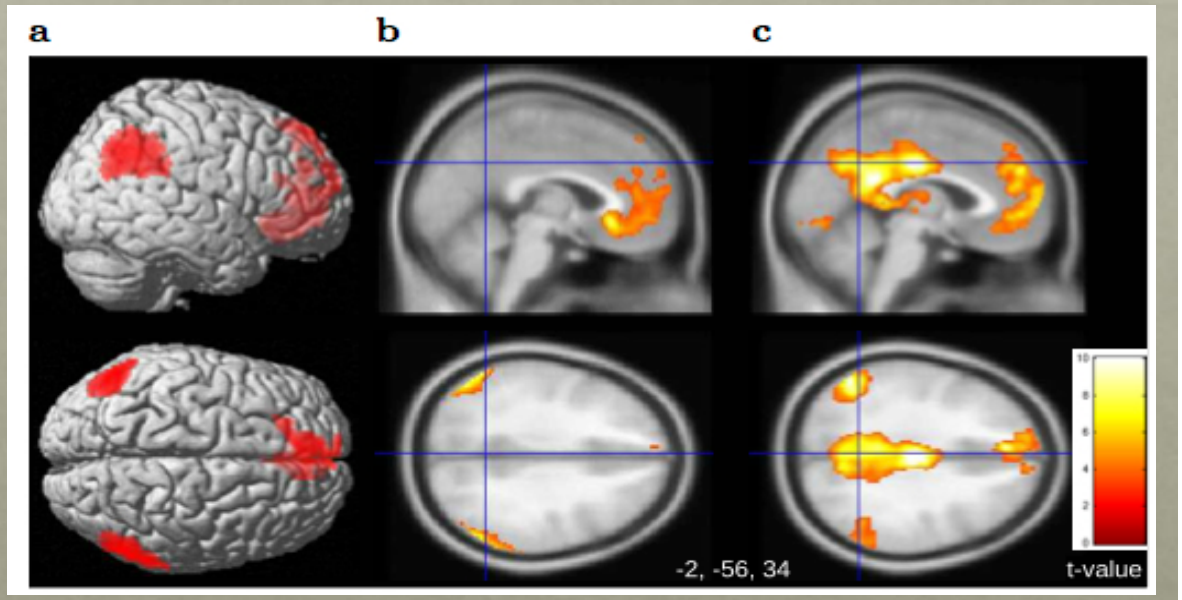
IMAGING



- Default mode network de-activated during learning/(decision)
- … yet frequencies encoded there

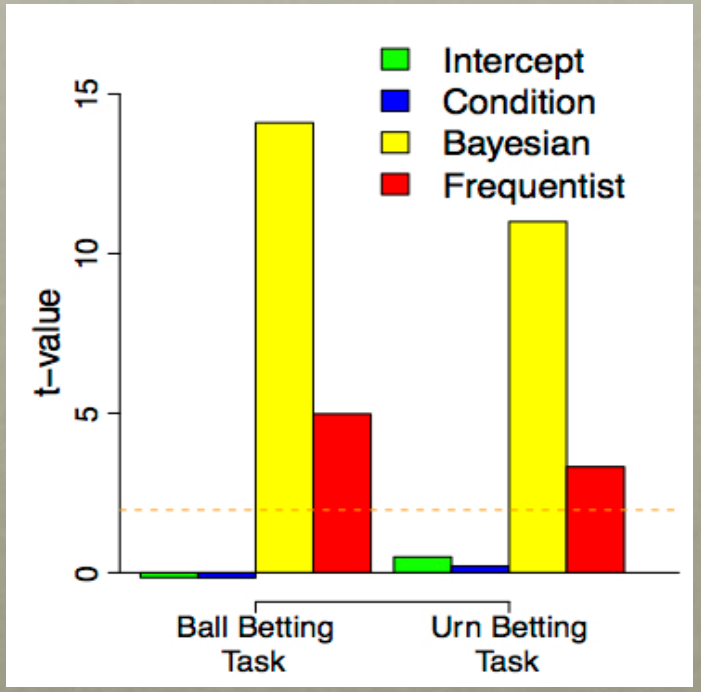
THE DEFAULT MODE NETWORK

- **While de-activated on average, this network is engaged in learning probabilities**
- Because of episodic memory retrieval induced by external stimuli?
- c: connectivity analysis
- Remark: de-activation does NOT mean "less involved"

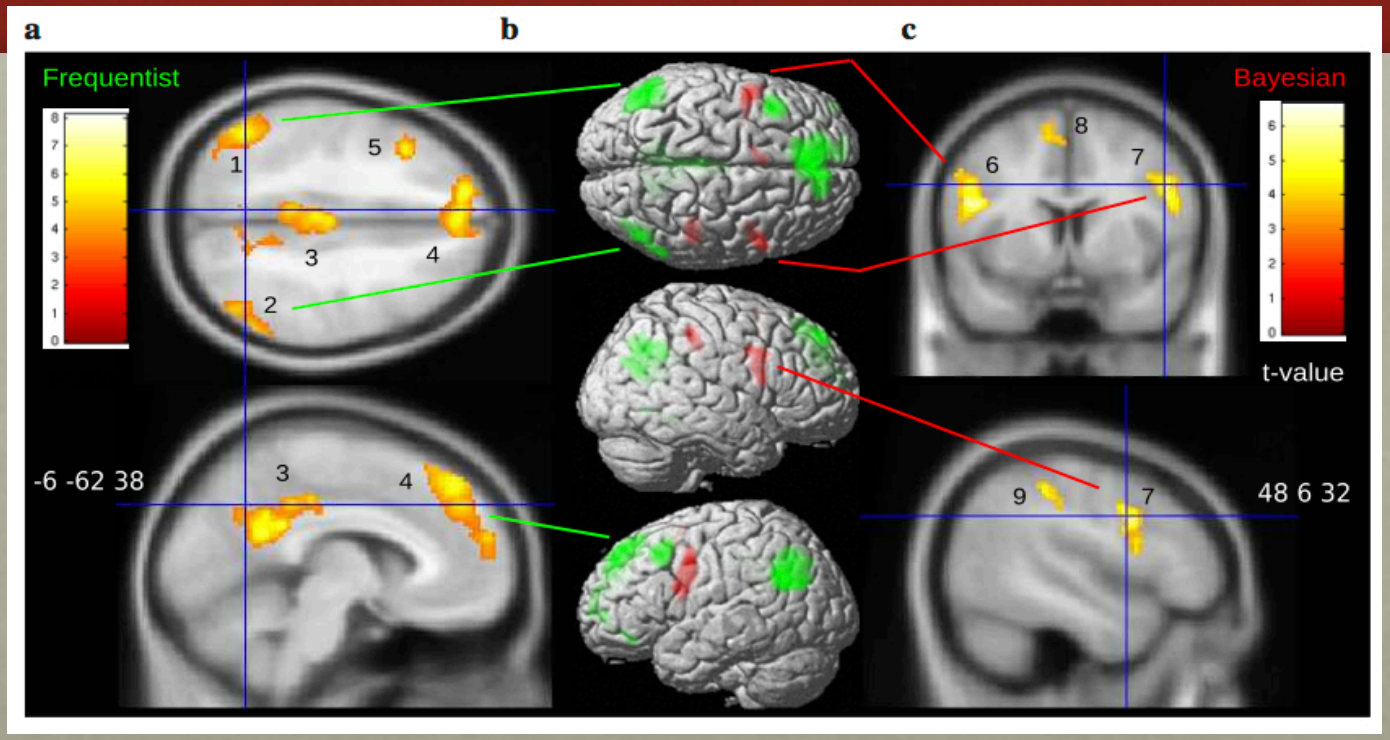


QUESTION 2-3-4: PRIOR BELIEFS, FREQUENCIES, COMBINATION (POSTERIOR)

- Task 1: Which urn did draws come from? "Urn Betting Task" (as before)
- New Task: What is color of next ball? "Ball Betting Task"
- Model beliefs revealed in bets as (i) Bayesian* (yellow), (ii) Frequentist (i.e., ignoring prior information) (red)
  - **Assuming correct priors!*
  - *(With Mathieu d'Acremont and Wolfram Schultz, J Neuroscience, resubmitted)*



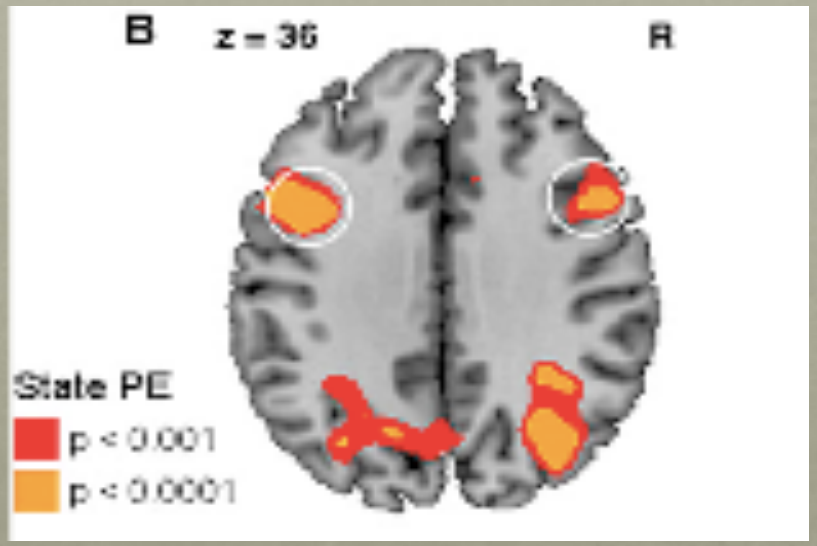
IMAGING



- **Stimulus (color of ball) probabilities** (unrelated to value of game to subject):
  - Frequencies ("objective" evidence) are encoded in default mode network (Left)
  - Bayesian "subjective" *im*probabilities are encoded in lateral precentral gyrus (extending from posterior IFG): Right; *semantic memory? – "I strongly believed not to see this stimulus"*

BACK TO GLÄSCHER, DAW, DAYAN, O'DOHERTY

- Activation correlating with "prediction error" of state probabilities (their wording), which is equivalent to *im*probability of observed state transition (our wording)
- "Dorsal posterior inferior frontal gyrus" is exactly as in our setting
- [Posterior parietal: should be interpreted as "objective evidence" (frequencies)?]



INTRIGUING…

- Evidently, the human brain tracks frequencies ("objective" evidence) *separately* from Bayesian beliefs (which merge frequencies with "subjective" prior)
- This is very useful: if prior somehow changes, can re-compute beliefs from the SAME (objective) evidence rather than starting from scratch!
  - (Epstein-Schneider in economics: need this to resolve ambiguity)
  - (Psychotherapy based on "re-living the past"?)

MESSAGES - REJOINDER

- Why "contrasts" in imaging can be misleading
  - See default mode network and tracking of frequencies
- Why knowing the biology behind behavior is important
  - How can we understand that therapies involving "reliving past" work?
- (How "converging evidence" works)
- Understanding models beyond the mathematics
  - Bayesian updating is about martingales
  - Bayesian updating is about merging prior belief and factual information
- Experiments are not simply about "treatment" vs "control"
  - Changes in (Bayesian) probabilities modulate brain activation