

Updating Great Expectations, Not Bad Ones Old wine in a new bottle?

Two Dimensions of Value: Dopamine Neurons Represent Reward But Not Aversiveness

Christopher D. Fiorillo

2 AUGUST 2013 VOL 341 SCIENCE www.sciencemag.org

Ritwik
Tea-talk: 22st August, 2013

Take Homes

- Are rewards (appetitive) and punishments (aversive)
 - Along a continuum of a single dimension ?
 - Two distinct categories: two discrete dimensions ?
- Potential confounds of past results:
 - (1) “short latency activation of DA neurons to airpuff due to sensory intensity not aversiveness”.
 - (2) “to characterize any single neuron both rewards and punishments must be provided in close proximity to one another”: highly aversive stimulus will interfere with reward; but mild punishment like airpuff may not be negative enough
 - (3) punishment must be calibrated against reward to determine its magnitude
- DA neurons fire/suppressed phasically to (unpredicted) reward and its omission. What about to punishments and omission?
- **Phasic DA appetitive RPE, not aversive RPE. [but we knew that already!]; however...**

Mid-brain DA neurons compute TD error

$$V_t = V_t + \alpha \delta_t$$

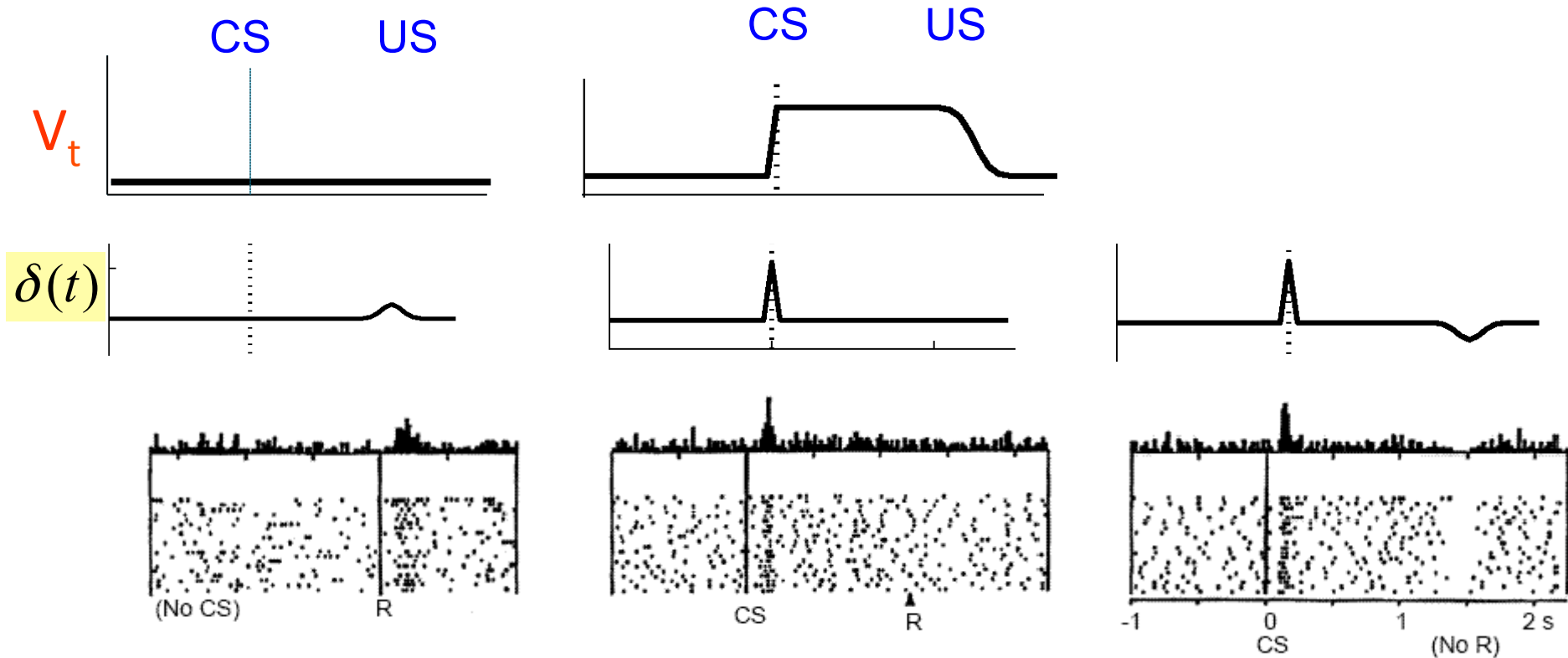
Reward Prediction Error $\delta_t = r_t + V_{t+1} - V_t$

Schultz, Dayan & Montague (1997)

Before

After

Reward Omission

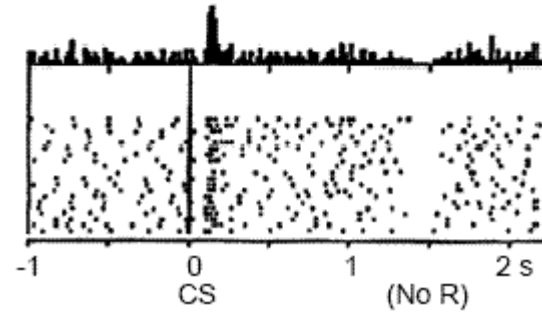
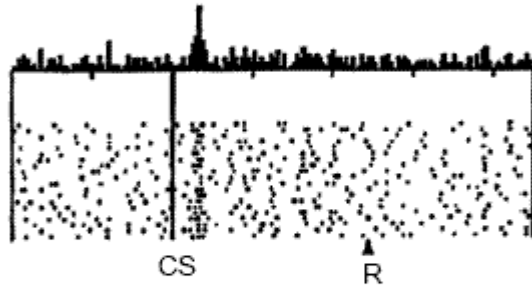


no prediction

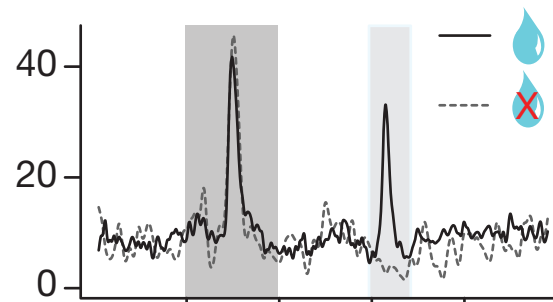
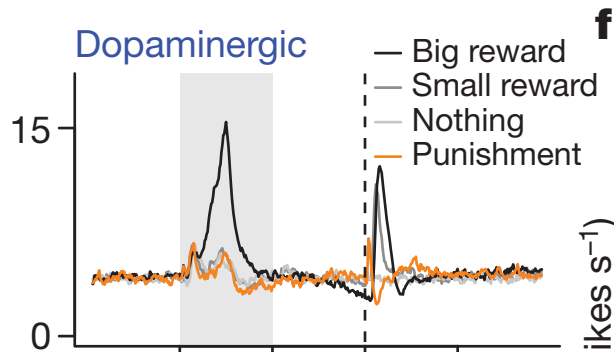
prediction, reward

prediction, no reward

Mid-brain DA neurons compute appetitive TD error



Schultz, Dayan & Montague (1997)

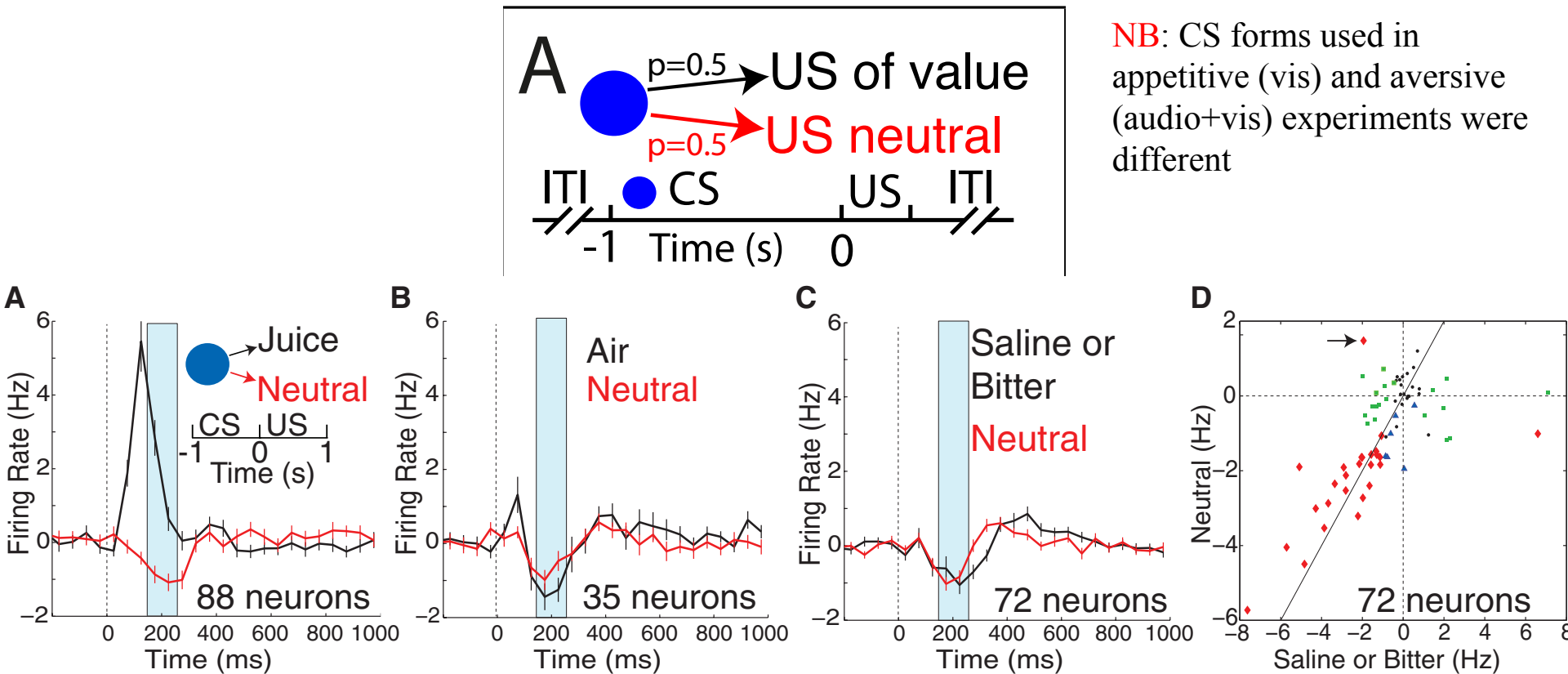


Cohen et al (2012)

prediction, reward

prediction, no reward

DA neurons not activated by omitting aversive UCS



NB: CS forms used in appetitive (vis) and aversive (audio+vis) experiments were different

Fig. 1. Dopamine neurons are not activated by omission of an expected aversive stimulus. Monkeys were conditioned with audiovisual Pavlovian stimuli to expect a stimulus (after a 1.0-s delay) that was either neutral sound or had appetitive or aversive value [(A) inset and fig. S1A]. (A) Juice (black) and its absence (red) caused an increase and decrease in average firing rate, respectively, across a population of 88 neurons. Neuronal discrimination of value was best at 150 to 250 ms after stimulus onset (shaded region) (16). All peri-stimulus time histograms (bin size, 50 ms) are averages across all recorded neurons, some of which were unresponsive. (B) Both air (black) and its absence (red) caused suppression. Unlike (A), (C), and (D), data are only from monkey F.

(C) Both saline or bitter (black) and its absence (red) caused suppression. (D) Firing rates (150 to 250 ms, baseline rates subtracted) of each neuron to saline (or bitter) and neutral outcomes. The arrow indicates a single neuron in which the neutral stimulus caused activation, which is consistent with the single-dimensional hypothesis. Symbols indicate results of *t* tests: activation or suppression to saline or bitter (green squares), to the neutral stimulus (blue triangles), both (red diamonds), or neither (black circles). The diagonal line indicates identity. Pearson correlation $r = 0.63$; $P < 10^{-8}$. Of these 72 neurons, 8, 2, and 62 were from the ventral tegmental area, retrorubral field, and substantia nigra, respectively. 35 were from the dorsal tier, and 37 were from the ventral tier.

No difference in phasic suppression for predicted and unpredicted aversive UCS

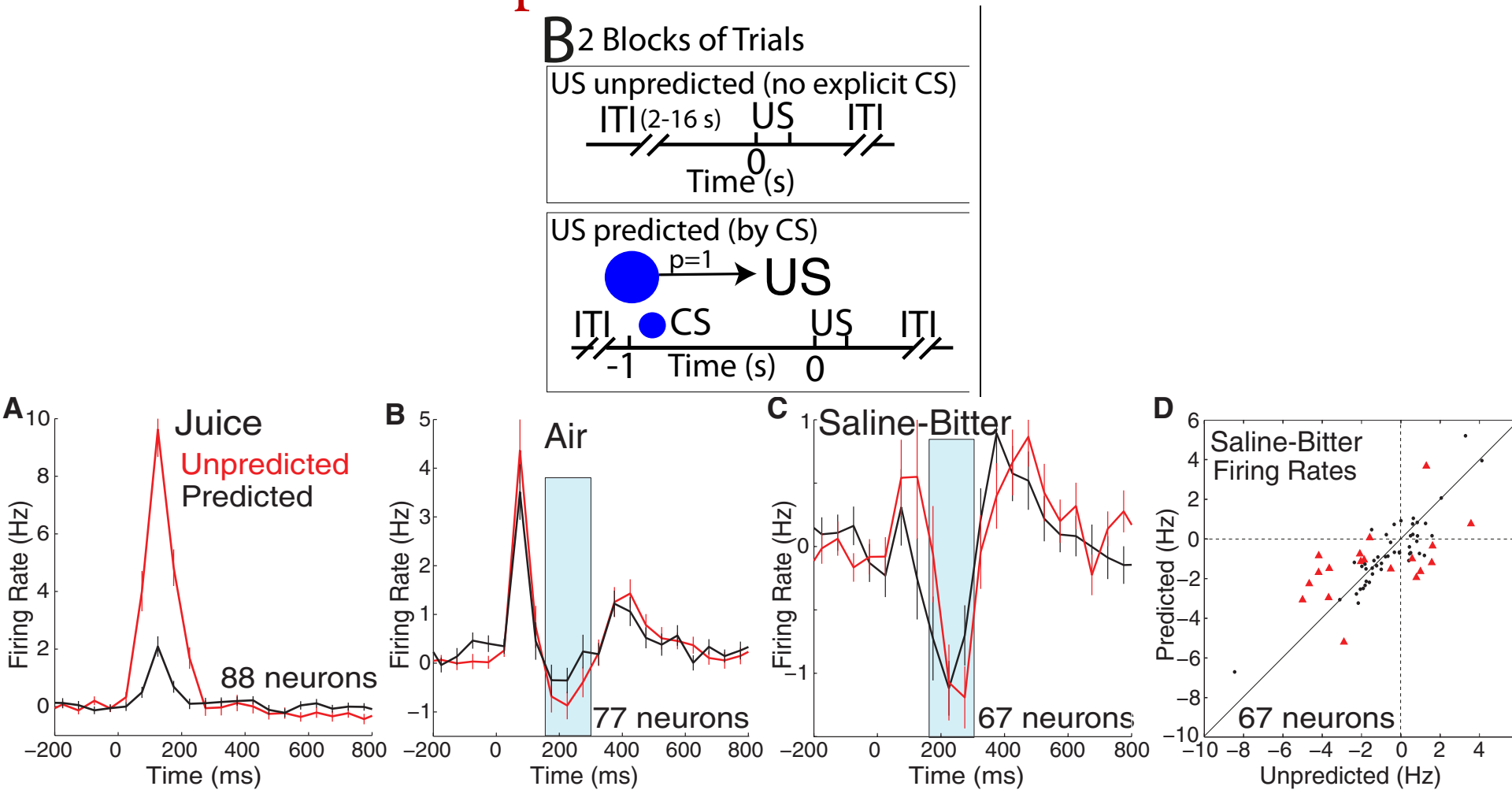


Fig. 2. Suppression by aversive stimuli is insensitive to prediction. Predicted stimuli occurred 1.0 s after a CS, whereas “unpredicted” stimuli were delivered once every 2 to 16 s with no CS (fig. S1B). There are differences in scales of y axes across (A) to (D). (A) Unpredicted (red) but not predicted (black) juice reward caused strong activation. (B) Prediction only marginally diminished the sensory-related activation and subsequent

suppression, to air. (C) Prediction of saline or bitter had little or no effect on suppression. (D) Firing rate [150 to 300 ms; shaded region in (C)] for each neuron to predicted and unpredicted saline or bitter (with baseline rates subtracted). The diagonal indicates identity. Red triangles indicate significant difference between responses to predicted and unpredicted saline-bitter.

Learning (reduced) Value for saline-bitter “not

airpuff”

NB: no cue preceding CS in C and D

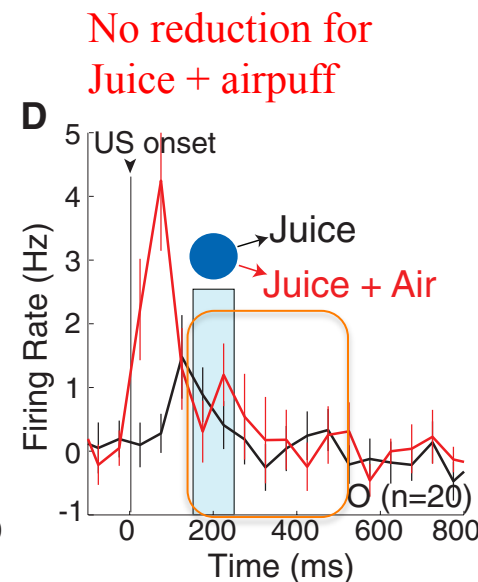
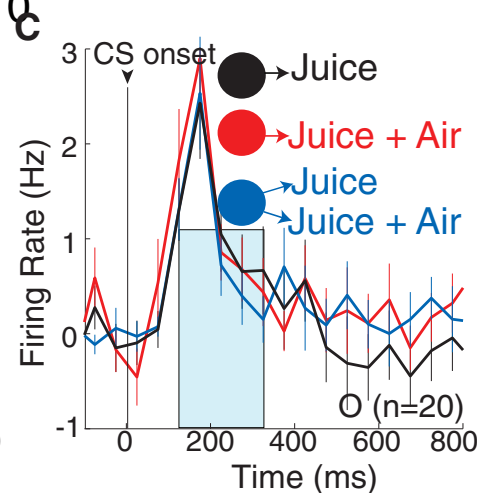
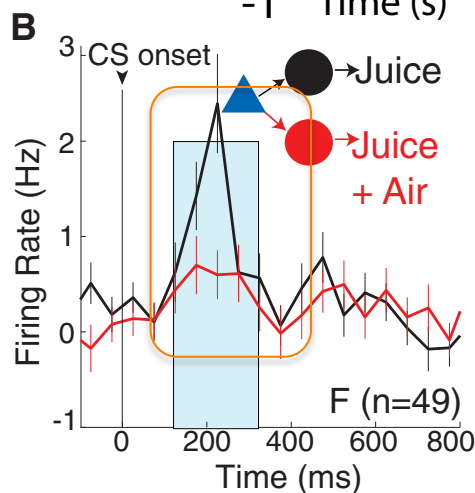
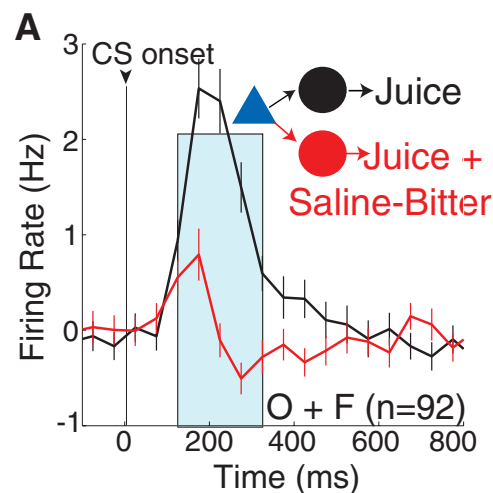
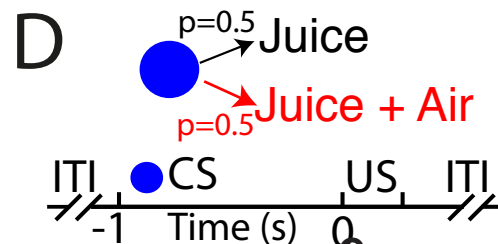
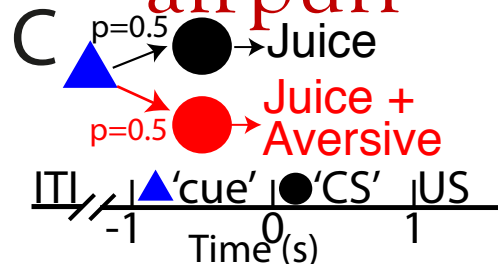


Fig. 3. In the context of juice, dopamine neurons are highly sensitive to saline and bitter, but not air. One CS predicted juice alone (180 μ l), and another predicted simultaneous delivery of juice plus an aversive stimulus (insets and fig. S1, C and D). **(A)** Prediction of saline (or bitter) suppressed activation to CS onset in 92 neurons from monkeys O and F. **(B)** Prediction of air caused only a modest suppression of activation to CS onset in monkey F. **(C)** Prediction of air

in monkey O had no effect. Unlike (A) and (B), no cue predicted CS onset. **(D)** After the “blue” CS in (C), firing rates did not discriminate delivery of air plus juice from juice alone during the period of 150 to 250 ms after unconditional stimulus (US) onset, in which reward value is best discriminated. The short latency activation to air (40 to 100 ms) is due to its high sensory intensity and was more prominent in monkey O than in monkey F (16).

Take Homes

- C. Fiorillo claims
 - Phasic DA appetitive RPE, not aversive RPE.
 - Four types of neurons: R_{ON} , R_{OFF} , A_{ON} , A_{OFF}
 - R_{ON} is DA.
 - Other 3 may be other neuromodulators
- BUT: R_{OFF} (omission of R) is also DA
- Subset/spatially segregated region of VTA DA neurons which respond to aversiveness (Ungless).

[R7's] Outstanding questions

- Phasic DA appetitive RPE, not aversive RPE. [but we knew that already!]; however...
- Unpredicted vs predicted juice+ bitter (is r reduced directly)?
- Punishments are not the only costs (or R7 would not have a thesis!)
- What are the natural statistics of rewards and costs? How are they represented?
- What makes a reward +ve, cost -ve? Is the distribution over motivational UCSs integrated to yield a $[|r|, \text{sign}(r)]$ representation? Implementation?
- Punishments (or other costs) need not involve TD based learning
- DA neurons could carry (not compute) the punishments (not punishment pred errors) computed by upstream (eg. Habenula) neurons