

Melnikoff and Strohminger 2023

Kevin Dorst

24.223 Rationality

Biased assimilation. Bayesians explain with differing priors in...

... who's trustworthy (Henderson and Gebharter)

... where they're likely to find a flaw (Kelly and Dorst)

Challenge: belief polarization can happen even when priors are held constant.

- Randomly assigned to prosecute or defend in a mock trial.
- Didn't actually advocate; just told they'd be paid to do so.
- Result: those assigned to defend increased confidence that innocent; those assigned to prosecute increased confidence that guilty.

Proposal: Perhaps people are getting hidden information.

- *Affect* is a source of evidence.
- People update on their *affective prediction errors*—if feel better than expected about prospects of defending, update in favor of innocence.

Hypotheses:

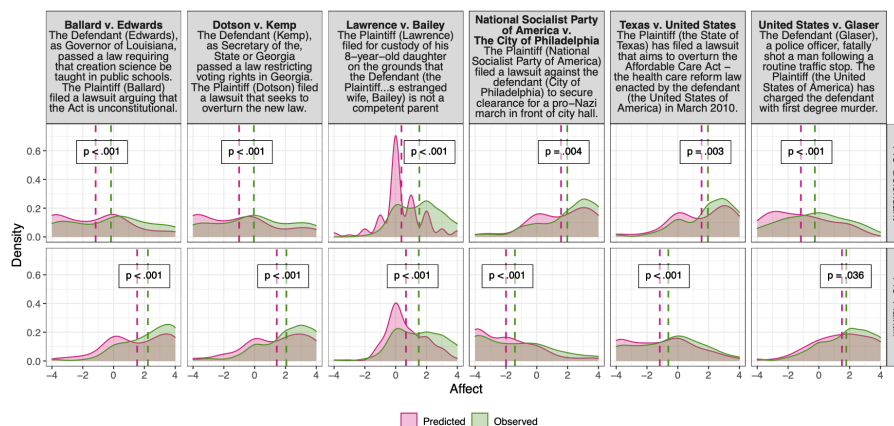
H1. *Positive affective prediction errors:* On average, people will feel better than expected.

H2. *Error-based updating:* people update their beliefs in proportion to size of the affective prediction error.

H3. *Noise hypothesis:* belief updating will be inversely proportional to subjective observation noise

Results:

Affective prediction errors:



Error-based updating:

Or randomized

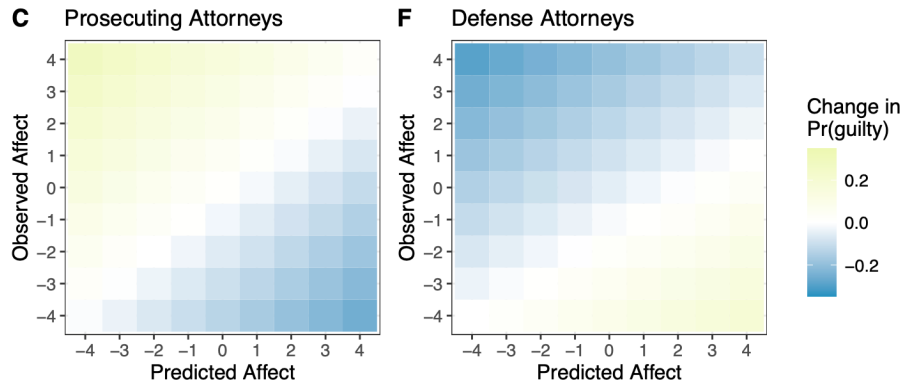
Broad vs. narrow evidence (Kelly).

Clear evidence in third-person case

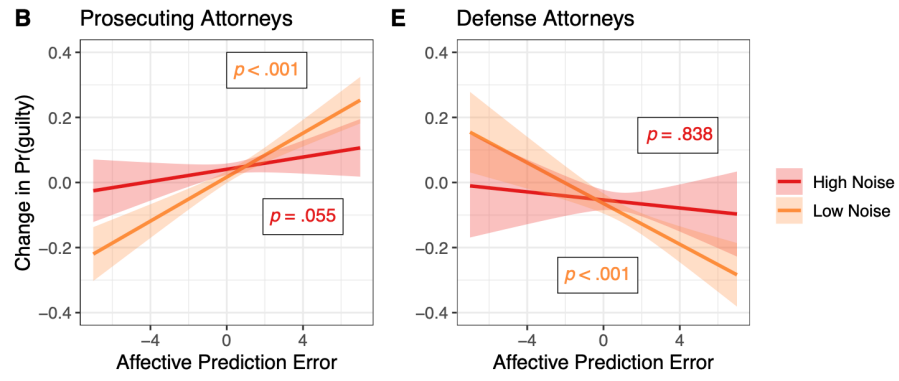
"Projection bias". Eg being full. Eg winning lottery.

As opposed to valence of affect, as predicted by motivated reasoning.

How diagnostic they expect affect to be



Noise hypothesis:



Measures of noise: (1) conditional on innocence/guilt, how what would you feel (and how confident); and (2) Unconditionally, how confident in your prediction?

With an extra trial to learn about affective errors, reduce both affective prediction errors and belief updating.

